Aus dem Deutschen Krebsforschungszentrum (DKFZ), Heidelberg,
in der Helmholtz-Gemeinschaft
(Wissenschaftlicher Stiftungsvorstand: Prof. Dr. med. Michael Baumann)
Abteilung Klinische Epidemiologie der Krebsfrüherkennung
(Leiter: Dr. Sigrid V. Carlsson)

# Associations of Metabolomic and Proteomic Profiles with Frailty and Mortality in Population-Based Cohorts

**Inauguraldissertation**
zur Erlangung des Doctor scientiarum humanarum (Dr. sc. hum.)
an der
Medizinischen Fakultät Heidelberg
der
Ruprecht-Karls-Universität

vorgelegt von
Lei Peng
aus
Sichuan, Volksrepublik China

2026

Dekan: Prof. Dr. Hans-Georg Kräusslich

Doktorvater: Prof. Dr. Ben Schöttker

# TABLE OF CONTENTS

**LIST OF ABBREVIATIONS** ..... **V**
**LIST OF FIGURES** ..... **VII**
**LIST OF TABLES** ..... **IX**
**1 INTRODUCTION** ..... **1**
1.1 Biological ageing: concepts and challenges ..... 1
1.1.1 Population ageing and the burden of age-related diseases ..... 1
1.1.2 Chronological age versus biological ageing ..... 1
1.1.3 Biological mechanisms of ageing ..... 2
1.2 Frailty as a clinical manifestation of biological ageing ..... 2
1.2.1 Concepts and models of frailty ..... 3
1.2.2 Clinical relevance of frailty ..... 4
1.2.3 Frailty assessment instruments ..... 4
1.2.4 Challenges in frailty measurement ..... 5
1.3 Molecular basis of frailty and biological ageing ..... 6
1.3.1 Inflammaging ..... 6
1.3.2 Metabolic dysregulation ..... 6
1.3.3 Tissue remodeling and inter-organ communication ..... 7
1.3.4 Molecular biomarkers of frailty ..... 7
1.4 Multi-omics approaches to biological ageing ..... 8
1.4.1 From phenotype to molecular signatures ..... 8
1.4.2 Metabolomics ..... 8
1.4.3 Proteomics ..... 9
1.4.4 Integrating metabolomics and proteomics ..... 10
1.5 Mortality as an integrative outcome of biological ageing ..... 10
1.5.1 Mortality as the ultimate ageing outcome ..... 10
1.5.2 Risk prediction beyond chronological age ..... 11
1.5.3 Molecular prediction of mortality ..... 11
1.6 Molecular measures of biological ageing ..... 12
1.6.1 Biological ageing clocks ..... 12
1.6.2 Metabolomics-based ageing clocks ..... 12
1.6.3 Proteomics-based ageing clocks ..... 13
1.6.4 Current knowledge gaps ..... 13
1.7 Integrative research framework of the dissertation ..... 14

1.8 Overall objectives and specific research questions ..... 15

**2 MATERIALS AND METHODS ..... 17**

2.1 Study 1: Comparison of four frailty measurements ..... 17

2.1.1 Study design and participant selection ..... 17
2.1.2 Frailty measurements and Short Physical Performance Battery (SPPB) ..... 18
2.1.3 Assessment of covariates ..... 23
2.1.4 Statistical analyses ..... 24

2.2 Study 2: Integrated proteomic and metabolomic profiling for frailty ..... 25

2.2.1 Study design and populations ..... 25
2.2.2 Assessment of frailty ..... 26
2.2.3 Proteomics profiling ..... 27
2.2.4 Metabolomics profiling ..... 27
2.2.5 Assessment of covariates ..... 28
2.2.6 Statistical analyses ..... 28

2.3 Study 3: Metabolomics-based biological ageing clocks for mortality ..... 30

2.3.1 Study design and populations ..... 30
2.3.2 Metabolomics profiling ..... 31
2.3.3 Mortality ascertainment ..... 32
2.3.4 Assessment of covariates ..... 32
2.3.5 Statistical analyses ..... 32

2.4 Study 4: Inflammaging and proteomic ageing clocks for mortality ..... 35

2.4.1 Study design and populations ..... 35
2.4.2 Mortality ascertainment ..... 37
2.4.3 Inflammation-related protein and proteomic profiling ..... 37
2.4.4 Assessment of covariates ..... 37
2.4.5 Statistical analyses ..... 37
2.4.6 Ethics approval and consent to participate ..... 39

**3 RESULTS ..... 40**

3.1 Study 1: Comparison of four frailty measurements ..... 40

3.1.1 Baseline characteristics of the study population ..... 40
3.1.2 Concordance of frailty measures ..... 41
3.1.3 Increasing frailty prevalence with age ..... 42
3.1.4 Baseline association between frailty scales and physical performance ..... 42
3.1.5 Longitudinal association of frailty status and all-cause mortality ..... 43
3.1.6 Sensitivity analyses ..... 44

3.2 Study 2: Integrated proteomic and metabolomic profiling for frailty ..... 49
3.2.1 Baseline characteristics of study populations ..... 49
3.2.2 Identification of proteomic and metabolomic features associated with prevalent and incident frailty ..... 50
3.2.3 TRIM models for prevalent and incident frailty ..... 52
3.2.4 Clinical utility of the TRIM models for prevalent and incident frailty in community-based populations ..... 58
3.2.5 Sensitivity analyses using incident frailty-derived omics biomarkers ..... 58
3.3 Study 3: Metabolomics-based biological ageing clocks for mortality ..... 61
3.3.1 Distribution of traditional risk factors and their associations with 10-year mortality in UKB and ESTHER ..... 61
3.3.2 Associations of metabolomic biomarkers with all-cause mortality ..... 66
3.3.3 Metabolic pathways associated with all-cause mortality ..... 67
3.3.4 Selected metabolomic biomarkers for predicting 10-year all-cause mortality ..... 71
3.3.5 Discrimination and reclassification analyses of 10-year mortality prediction models including metabolomic biomarkers ..... 74
3.3.6 Development and testing of metabolomic ageing clocks ..... 80
3.4 Study 4: Inflammaging and proteomic ageing clocks for mortality ..... 88
3.4.1 Baseline characteristics and conventional mortality risk factors ..... 88
3.4.2 Sex-specific inflammatory and proteomic biomarkers associated with 10-year all-cause mortality risk ..... 92
3.4.3 Biological pathways represented by mortality-associated protein signatures ..... 93
3.4.4 Inflammatory and proteomic biomarkers improve prediction of mortality risk ..... 94
3.4.5 Development and validation of inflammation- and proteomics-based ageing clocks ..... 102
3.4.6 Inflammaging and proteomic age accelerations are associated with all-cause and cause-specific mortality ..... 104

**4 DISCUSSION ..... 107**

4.1 Study 1: Comparison of four frailty measurements ..... 107
4.1.1 Summary of the key results ..... 107
4.1.2 Agreement of frailty classifications ..... 107
4.1.3 Ageing and frailty ..... 107
4.1.4 Frailty and physical performance ..... 108
4.1.5 Frailty and mortality ..... 109
4.1.6 Strengths and limitations ..... 110

4.2 Study 2: Integrated proteomic and metabolomic profiling for frailty ..... 110
4.2.1 Main findings and comparison with previous studies ..... 110
4.2.2 Biological interpretation of identified biomarkers ..... 111
4.2.3 Clinical and public health implications ..... 112
4.2.4 Strengths and limitations ..... 113
4.3 Study 3: Metabolomics-based biological ageing clocks for mortality ..... 113
4.3.1 Main findings ..... 113
4.3.2 Biological interpretation of mortality-associated metabolites ..... 114
4.3.3 Incremental value of metabolomics for mortality prediction ..... 115
4.3.4 Metabolomics-based biological ageing clocks ..... 116
4.3.5 Strengths and limitations ..... 117
4.4 Study 4: Inflammaging and proteomic ageing clocks for mortality ..... 117
4.4.1 Summary of the key results ..... 117
4.4.2 Biological basis of the proteomic and inflammaging ageing clocks ..... 118
4.4.3 Sex differences and comparison with previous molecular ageing clocks ..... 119
4.4.4 Clinical implications and practical applications ..... 120
4.4.5 Strengths and limitations ..... 121
4.5 Integrated interpretation across the four studies ..... 121
4.6 Overall strengths and limitations ..... 123
4.7 Overall implications and future directions ..... 124
4.8 Overall conclusions ..... 126

**5 SUMMARY AND ZUSAMMENFASSUNG ..... 127**
5.1 English summary ..... 127
5.2 Deutsche Zusammenfassung ..... 129

**6 REFERENCES ..... 131**
**7 PUBLICATIONS AND PERSONAL CONTRIBUTIONS ..... 149**
**8 APPENDIX ..... 152**
**CURRICULUM VITAE ..... 179**
**ACKNOWLEDGEMENTS ..... 181**
**EIDESSTATTLICHE VERSICHERUNG ..... 183**
**ANGABEN ZU VERWENDETEN KI-BASIERTEN ELEKTRONISCHEN HILFSMITTEL ..... 185**

# LIST OF ABBREVIATIONS

| | |
|---|---|
| ADL | Activities of Daily Living |
| AUC | Area Under the Receiver Operating Characteristic Curve |
| BMI | Body Mass Index |
| CFS | Clinical Frailty Scale |
| CI | Confidence Interval |
| CIRS-G | Cumulative Illness Rating Scale–Geriatric |
| CV | Cross-Validation |
| CVD | Cardiovascular Disease |
| DCA | Decision Curve Analysis |
| DNA | Deoxyribonucleic Acid |
| EDTA | Ethylenediaminetetraacetic Acid |
| ESTHER | Epidemiological Study on the Chances of Prevention, Early Detection and Optimized Treatment of Chronic Diseases in an Elderly Population *(German: Epidemiologische Studie zu Chancen der Verhütung, Früherkennung und optimierten Therapie chronischer Erkrankungen in der älteren Bevölkerung)* |
| FDR | False Discovery Rate |
| FI | Frailty Index |
| FFP | Fried Frailty Phenotype |
| FS | FRAIL Scale |
| GP | General Practitioner |
| HbA1c | Glycated Hemoglobin |
| HMDB | Human Metabolome Database |
| HR | Hazard Ratio |
| ICD-10 | International Classification of Diseases, Tenth Revision |
| IDI | Integrated Discrimination Improvement |
| IL | Interleukin |
| InflamAA | Inflammaging Age Acceleration |
| IPAQ | International Physical Activity Questionnaire |
| KEGG | Kyoto Encyclopedia of Genes and Genomes |
| LASSO | Least Absolute Shrinkage and Selection Operator |
| LOD | Limit of Detection |
| LR | Likelihood Ratio |
| MAE | Mean Absolute Error |

| | |
|---|---|
| MASH | Metabolic Dysfunction-Associated Steatohepatitis |
| MetAA | Metabolomic Age Acceleration |
| MetaboMR | Metabolomics-Based Mortality Risk |
| MREC | Multi-centre Research Ethics Committee |
| NHS | National Health Service |
| NMR | Nuclear Magnetic Resonance |
| NPX | Normalized Protein eXpression |
| NRI | Net Reclassification Improvement |
| OR | Odds Ratio |
| PEA | Proximity Extension Assay |
| ProtAA | Proteomic Age Acceleration |
| RAP | Research Analysis Platform |
| ROC | Receiver Operating Characteristic |
| SD | Standard Deviation |
| sHR | Subdistribution Hazard Ratio |
| SPPB | Short Physical Performance Battery |
| STRING | Search Tool for the Retrieval of Interacting Genes/Proteins |
| TCA | Tricarboxylic Acid |
| TRIM-FI | Traditional Risk factor, Inflammation, and Metabolism based Frailty Identification Model |
| TRIM-FP | Traditional Risk factor, Inflammation, and Metabolism based Frailty Prediction Model |
| UKB | UK Biobank |
| WGCNA | Weighted Gene Co-expression Network Analysis |
| WHO | World Health Organization |

# LIST OF FIGURES

Figure 1. Overview of the dissertation: Four complementary studies from clinical assessment to molecular ageing clocks and mortality prediction ..... 15

Figure 2. Flowchart of study participant selection ..... 18

Figure 3. Definitions and classification criteria of different frailty measures and physical performance instrument ..... 23

Figure 4. Flowchart of participant selection and analytical design in the UK Biobank and ESTHER study ..... 26

Figure 5. Flow chart of study population from the UK Biobank and ESTHER study ..... 31

Figure 6. Study design and analytical workflow for development and validation of inflammatory and proteomic mortality prediction models and ageing clocks ..... 36

Figure 7. Inter-instrument agreement between the four frailty instruments in the classification of non-frail, pre-frail and frail individuals ..... 41

Figure 8. Age trend of frailty prevalence according to the four frailty measures in the total population and specific subgroups ..... 42

Figure 9. Longitudinal association of the four frailty measures with all-cause mortality in the overall population and specific subgroups ..... 46

Figure 10. Longitudinal association of the four frailty measures with all-cause mortality in the overall population and specific subgroups with extended model adjustment ..... 48

Figure 11. Associations of LASSO-selected proteomic and metabolomic biomarkers with prevalent frailty and incident frailty during 8-year follow-up in the UK Biobank and ESTHER study ..... 51

Figure 12. Heatmap of Pearson correlations between 17 proteomic and 9 metabolomic biomarkers in the UK Biobank ..... 52

Figure 13. ROC curves for discrimination of the TRIM-FI and TRIM-FP models for prevalent and incident frailty, respectively ..... 56

Figure 14. Calibration of the TRIM models for prevalent and incident frailty ..... 57

Figure 15. Decision curve analysis in the external validation cohort (ESTHER study) ..... 58

Figure 16. Overlap of proteomic and metabolomic biomarkers selected at baseline and follow-up in the UK Biobank ..... 59

Figure 17. Metabolome-wide discovery and validation of biomarkers associated with 10-year all-cause mortality in the UK Biobank and ESTHER study ..... 66

Figure 18. Topological overlap map of 68 metabolomic biomarkers ..... 68

Figure 19. Sex- and age-specific metabolomic biomarkers associated with 10-year all-cause mortality ..... 72

Figure 20. Effect sizes of standardized β-coefficients of selected biomarkers used to derive mortality risk scores in the training set (70% of the UK Biobank) ..... 73

Figure 21. Spearman correlations between the conventional risk factors and 26 metabolomic biomarkers in the derivation set (70% UK Biobank)..... 75

Figure 22. Improved mortality prediction accuracy with identified metabolomic biomarkers in sex- and age-specific groups, as well as the total study population..... 78

Figure 23. Calibration curves of the models only including conventional risk factors and their combinations with the metabolites for risk prediction in the internal and external validation cohorts..... 79

Figure 24. Distributions of chronological age and metabolomics-based mortality-related ageing clocks in the UK Biobank and ESTHER cohorts..... 81

Figure 25. Linear correlation between chronological age and the metabolomic age based on 26 selected metabolomic biomarkers only (metabolomics-based mortality risk clock 1 (MetaboMR clock1))..... 82

Figure 26. Linear correlation between chronological age and metabolomic age based on 26 selected metabolomic biomarkers and traditional mortality risk factors (metabolomics-based mortality risk clock 2 (MetaboMR clock2))..... 83

Figure 27. Comparison of the C-statistics of chronological age and the MetaboMR clocks for 10-year mortality prediction in the UK Biobank and ESTHER study..... 87

Figure 28. Distribution of metabolomic age acceleration (MetAA) derived from MetaboMR clocks in the UK Biobank and ESTHER cohorts..... 87

Figure 29. Sex-specific overlap of inflammatory and proteomic biomarkers associated with 10-year all-cause mortality..... 92

Figure 30. Sex-specific inflammatory and proteomic biomarkers associated with 10-year all-cause mortality..... 93

Figure 31. Biological characteristics of sex-specific inflammatory and proteomic biomarkers associated with 10-year all-cause mortality..... 94

Figure 32. Spearman correlations between conventional risk factors and selected inflammatory and proteomic biomarkers in the derivation set (70% UKB)..... 96

Figure 33. Incremental improvements in discrimination (ΔC-statistics) for 10-year all-cause mortality after addition of individual inflammatory and proteomic biomarkers in internal and external validation sets..... 97

Figure 34. Distributions of chronological age and mortality risk ageing clocks in the UK Biobank and ESTHER cohorts..... 102

Figure 35. Linear correlations between chronological age and inflammatory age estimated by the inflammaging clocks derived from the Olink Target Inflammation Panel..... 103

Figure 36. Linear correlations between chronological age and proteomic age estimated by the proteomic ageing clocks derived from the Olink Explore Panels..... 104

# LIST OF TABLES

Table 1. Items used to calculate the FI, FFP, CFS, FS, and SPPB in the ESTHER study ..... 20

Table 2. Cut-off values for the Frailty Index (FI) selected to approximate the prevalence of frailty categories defined by the Fried Frailty Phenotype (FFP) in the ESTHER baseline population ..... 22

Table 3. Baseline characteristics of the study population (n = 2,000) ..... 40

Table 4. Baseline associations of frailty status assessed with different instruments with poor physical performance ..... 45

Table 5. Associations of frailty status assessed with different instruments with poor physical performance with extended model adjustment ..... 47

Table 6. Baseline characteristics of study participants in the UK Biobank and ESTHER study ..... 49

Table 7. Predictive value of proteomic and metabolomic biomarkers for the assessment of baseline frailty and prediction of incident frailty among older adults in the UK Biobank and ESTHER study ..... 54

Table 8. Sensitivity analysis: predictive value of proteomic and metabolomic biomarkers selected only based on incident frailty in the UK Biobank and validated in the ESTHER study ..... 60

Table 9. Baseline characteristics of study participants in the UK Biobank and ESTHER study ..... 62

Table 10. Longitudinal association between traditional risk factors and 10-year all-cause mortality in the UK Biobank and ESTHER study ..... 64

Table 11. Summary of potential KEGG pathways associated with 68 validated metabolic biomarkers identified from WGCNA-derived metabolite clusters ..... 69

Table 12. C-statistics of prediction models for 10- and 5-year all-cause mortality in the UK Biobank and ESTHER cohorts ..... 76

Table 13. Reclassification improvements in 10- and 5-year mortality prediction by adding metabolomic biomarkers to conventional risk models: evidence from the UK Biobank and ESTHER study ..... 77

Table 14. Association of metabolomic age acceleration based on metabolomic biomarkers only (MetAA1) and based on both biomarkers and traditional risk factors (MetAA2) with 10-year all-cause mortality ..... 84

Table 15. Association of metabolomic age acceleration based on metabolomic biomarkers only (MetAA1) and based on both biomarkers and traditional risk factors (MetAA2) with 10-year all-cause mortality, excluding participants on lipid-lowering therapy ..... 85

Table 16. Association of metabolomic age acceleration based on metabolomic biomarkers only (MetAA1) and based on both biomarkers and traditional risk factors (MetAA2) with 10-year CVD and cancer mortality in the total UK Biobank ..... 86

Table 17. Baseline characteristics of study participants in the UK Biobank and ESTHER Study ..... 89

Table 18. Longitudinal association between traditional risk factors and 10-year all-cause mortality in the UK Biobank and ESTHER study ..... 91

Table 19. Inflammatory and proteomic panels differentially improve prediction of 10-year all-cause mortality in the overall population of UK Biobank and ESTHER cohorts .. 98

Table 20. Inflammatory and proteomic panels differentially improve prediction of 5-year all-cause mortality in the overall population of UK Biobank and ESTHER cohorts ..... 99

Table 21. Inflammatory and proteomic panels differentially improve prediction of 10-year all-cause mortality in men from UK Biobank and ESTHER cohorts ..... 100

Table 22. Inflammatory and proteomic panels differentially improve prediction of 10-year all-cause mortality in women from UKB and ESTHER cohorts ..... 101

Table 23. Association of inflammaging/proteomic age acceleration based on proteins only and based on both proteins and traditional risk factors with 10-year all-cause mortality ..... 105

Table 24. Association of inflammaging/proteomic age acceleration based on proteomic biomarkers only and on both biomarkers and traditional risk factors with 10-year CVD and cancer mortality in the overall study population of UK Biobank ..... 105

# 1 INTRODUCTION

## 1.1 Biological ageing: concepts and challenges

### 1.1.1 Population ageing and the burden of age-related diseases

Population ageing is one of the defining demographic changes of the twenty-first century and poses major challenges for healthcare systems worldwide. Driven by declining fertility and increasing life expectancy, the global population aged 60 years and older is projected to rise from approximately 1 billion in 2020 to more than 2 billion by 2050, accounting for around one-fifth of the world's population (Amuthavalli Thiyagarajan et al. 2022). This demographic shift is expected to substantially increase the burden of chronic diseases, disability, long-term care needs, and healthcare expenditure (Keating 2022).

Ageing is the strongest risk factor for most chronic non-communicable diseases, including cardiovascular disease, cancer, type 2 diabetes, chronic kidney disease, osteoporosis, and neurodegenerative disorders (López-Otín et al. 2013). As life expectancy increases, multimorbidity has become increasingly common, often accompanied by declines in physical function, cognitive performance, and quality of life. Extending lifespan has therefore become insufficient as a public health goal; increasing emphasis is now placed on healthy ageing and healthspan, defined as the years lived free from major disease, disability, and functional impairment (Beard et al. 2016). However, individuals of the same chronological age often differ markedly in health status and functional capacity (Campisi et al. 2019; Kennedy et al. 2014; Sudlow et al. 2015; Sun et al. 2018), highlighting the need for measures that better reflect biological ageing than chronological age alone.

### 1.1.2 Chronological age versus biological ageing

Chronological age, defined as the time elapsed since birth, has long been the principal measure of ageing in clinical practice and epidemiological research (Kennedy et al. 2014). Although it is strongly associated with chronic diseases, functional decline, and mortality (Campisi et al. 2019; Jia et al. 2024; Jia et al. 2026; Ma et al. 2023; Wang et al. 2023; Xiong et al. 2025), chronological age alone does not explain the marked variation in health status observed among individuals of the same age. Older adults with identical chronological ages may differ substantially in physiological function, resilience, disease susceptibility, and life expectancy (Ding et al. 2025; Johnson et al. 2020; López-Otín et al. 2023). This heterogeneity has led to the concept of biological ageing, which reflects the progressive decline in molecular integrity, cellular function, tissue homeostasis,

and physiological reserve resulting from lifelong interactions between genetic, environmental, and lifestyle factors (Coenen et al. 2023; Johnson et al. 2020; Lehallier et al. 2019; Lehallier et al. 2020; Ma et al. 2026; Macdonald-Dunlop et al. 2022; Oh et al. 2023; Tanaka et al. 2020). Biological ageing therefore provides a more informative measure of functional capacity and vulnerability to age-related diseases than chronological age alone.

Objective measures of biological ageing have become a major focus of ageing research because molecular changes often precede overt functional decline (Ayala-Lopez 2020; Balasubramanian et al. 2020; Deelen et al. 2019; Fischer et al. 2014; Huang et al. 2018; Lian and Vardhanabhuti 2024; Wang et al. 2023; Yan et al. 2022; Zhang et al. 2024). At the same time, biological ageing cannot be defined by molecular markers alone. Clinical phenotypes remain indispensable because they reflect the integrated consequences of molecular dysfunction at the whole-body level. Among these, frailty is widely regarded as one of the most representative clinical manifestations of biological ageing and provides a direct link between biological mechanisms and clinically relevant outcomes (Eldjarn et al. 2023; Lau et al. 2024; Reicher et al. 2024).

### 1.1.3 Biological mechanisms of ageing

Biological ageing is a multisystem process characterized by the progressive decline of molecular integrity, cellular function, and tissue homeostasis (Campisi et al. 2019; Kennedy et al. 2014). Rather than resulting from a single pathway, it arises through coordinated alterations across immune, metabolic, endocrine, and tissue maintenance systems, leading to progressive loss of physiological reserve and increasing susceptibility to age-related diseases (Campisi et al. 2019; Kennedy et al. 2014).

Among the biological processes implicated in ageing, chronic low-grade inflammation (inflammaging), metabolic dysregulation, and impaired inter-organ communication are considered central mechanisms linking molecular alterations to functional decline (Franceschi et al. 2018; Sayed et al. 2021). Together, these processes contribute to declining resilience, frailty, and other adverse ageing-related outcomes (Peng et al. 2026). The following sections discuss frailty as a major clinical manifestation of biological ageing before introducing the molecular mechanisms and multiomics approaches investigated in this dissertation.

## 1.2 Frailty as a clinical manifestation of biological ageing

Biological ageing cannot be directly measured in routine clinical practice but is reflected by changes in physiological function, resilience, and the capacity to maintain homeostasis under stress

(Beard et al. 2016; Kennedy et al. 2014; López-Otín et al. 2013). Among the clinical phenotypes used to characterize biological ageing, frailty is the most widely studied and best established (Clegg et al. 2013; Fried et al. 2001; Rockwood and Mitnitski 2007). Frailty is characterized by reduced physiological reserve and increased vulnerability to internal and external stressors, reflecting the cumulative decline of multiple physiological systems (Clegg et al. 2013; Fried et al. 2001). It is consistently associated with falls, disability, hospitalization, and mortality and has become a key concept in geriatric medicine and epidemiological research (Dent et al. 2019a; Hoogendijk et al. 2019; Rockwood and Mitnitski 2007).

### 1.2.1 Concepts and models of frailty

Frailty has no universally accepted definition but is generally considered a state of increased biological vulnerability resulting from diminished physiological reserve and a reduced capacity to maintain homeostasis when exposed to internal or external stressors (Clegg et al. 2013; Dent et al. 2019a; Hoogendijk et al. 2019). Although closely associated with ageing, frailty is not an inevitable consequence of growing older. Older adults of the same chronological age may differ substantially in resilience, functional capacity, and susceptibility to adverse health outcomes, highlighting frailty as a measure of biological rather than chronological ageing (Dent et al. 2019a; Hoogendijk et al. 2019).

Two conceptual models have shaped contemporary frailty research. The physical phenotype, proposed by Fried and colleagues, defines frailty as a biological syndrome identified by five measurable criteria: unintentional weight loss, self-reported exhaustion, weakness, slow walking speed, and low physical activity. Individuals meeting three or more criteria are classified as frail, whereas those with one or two criteria are considered prefrail (Fried et al. 2001). In contrast, the deficit accumulation model, developed by Rockwood and Mitnitski, conceptualizes frailty as the cumulative burden of health deficits across multiple physiological domains and quantifies vulnerability using a frailty index (Rockwood and Mitnitski 2007). These models emphasize different aspects of biological ageing—the physical phenotype focuses on functional performance and physiological reserve, whereas the frailty index reflects the accumulation of health deficits across multiple systems. Their distinct conceptual foundations have led to the development of several frailty assessment instruments, which may classify individuals differently and produce varying estimates of frailty prevalence (Cesari et al. 2014).

### 1.2.2 Clinical relevance of frailty

Frailty is an important predictor of adverse health outcomes in older adults and has become an integral component of geriatric assessment. Longitudinal studies have consistently shown that frailty is associated with increased risks of falls, disability, hospitalization, institutionalization, and all-cause mortality, independent of chronological age and chronic disease burden (Clegg et al. 2013; Dent et al. 2019a; Fried et al. 2004; Hoogendijk et al. 2019). Because frailty reflects diminished physiological reserve, even relatively minor stressors such as acute illness or surgery may lead to marked declines in functional status and loss of independence. Frailty assessment is therefore increasingly incorporated into clinical practice across multiple specialties, including cardiology, oncology, nephrology, and surgery, where it supports risk stratification, treatment planning, and prediction of recovery following medical interventions (Afilalo et al. 2014; Cesari et al. 2016; Hoogendijk et al. 2019). As frailty often develops gradually through a prefrail stage, early identification provides an opportunity to implement interventions that may delay functional decline and help maintain independence (Fried et al. 2004; Xue 2011).

Current frailty assessment largely relies on clinical characteristics, functional performance, or self-reported information, which describe the clinical manifestation of frailty but provide limited insight into its underlying biological mechanisms. Identifying objective molecular biomarkers may improve risk stratification, facilitate earlier detection, and enhance understanding of the biological processes driving frailty (Franceschi et al. 2018; Mak et al. 2023; Sayed et al. 2021; Wang et al. 2025; Yao et al. 2024). Proteomic and metabolomic profiling have therefore become valuable approaches for investigating the molecular basis of frailty and identifying biological pathways associated with ageing-related functional decline.

### 1.2.3 Frailty assessment instruments

A wide range of instruments has been developed to assess frailty in clinical practice and epidemiological research. Because these instruments are based on different conceptual models and emphasize different domains of health, they may identify different individuals as frail and produce varying estimates of frailty prevalence (Clegg et al. 2013; Dent et al. 2019a; Hoogendijk et al. 2019).

The Frailty Phenotype (FP), developed by Fried and colleagues, assesses frailty using five physical criteria: unintentional weight loss, exhaustion, weakness, slow walking speed, and low physical activity. Individuals meeting three or more criteria are classified as frail, whereas those meeting one or two criteria are considered prefrail (Fried et al. 2001). In contrast, the Frailty Index (FI),

derived from the deficit accumulation model, quantifies frailty as the proportion of accumulated health deficits across multiple physiological domains, providing a continuous measure of biological vulnerability (Mitnitski et al. 2001; Rockwood and Mitnitski 2007). The Clinical Frailty Scale (CFS) is a judgment-based instrument that classifies individuals according to mobility, functional independence, comorbidity, and cognition, making it suitable for rapid clinical assessment (Rockwood et al. 2005). The FRAIL Scale (FS) is a brief screening questionnaire based on fatigue, resistance, ambulation, illnesses, and weight loss, offering a simple approach for frailty assessment in both clinical and community settings (Morley et al. 2012). Although these instruments have all demonstrated clinical utility, they differ in their conceptual basis, required data, and intended application, which contributes to variation in frailty classification and may influence comparisons across studies (Clegg et al. 2013; Dent et al. 2019a; Hoogendijk et al. 2019).

### 1.2.4 Challenges in frailty measurement

Despite the widespread use of frailty assessment, no consensus exists regarding the optimal instrument for measuring frailty. Available instruments are based on different conceptual frameworks and emphasize different dimensions of health, resulting in variation in frailty prevalence, risk classification, and prediction of adverse outcomes across studies (Cesari et al. 2014; Clegg et al. 2013; Dent et al. 2019a; Hoogendijk et al. 2019). The performance of these instruments may also vary according to age, sex, healthcare setting, and population characteristics, making direct comparisons between studies challenging (Cesari et al. 2014; Dent et al. 2019a).

These methodological differences highlight the need to evaluate commonly used frailty instruments within the same study population (Arteaga et al. 2021; Belga et al. 2016; Checa-Lopez et al. 2023; Chong et al. 2018; Li et al. 2021; Malmstrom et al. 2014; Op Het Veld et al. 2019; Oviedo-Briones et al. 2022; Ravindrarajah et al. 2013; Ritt et al. 2015; Shang et al. 2022; Theou et al. 2013; Widagdo et al. 2015; Woo et al. 2012; Zhang et al. 2022). Such comparisons improve understanding of the strengths and limitations of individual instruments, facilitate appropriate instrument selection for specific research questions, and support interpretation of epidemiological findings. They also provide a foundation for integrating conventional frailty assessment with molecular biomarkers of biological ageing, which is the focus of the studies presented in this dissertation (Mak et al. 2023; Xu et al. 2025).

## 1.3 Molecular basis of frailty and biological ageing

Frailty reflects the cumulative effects of biological ageing across multiple physiological systems and is characterized by progressive loss of physiological reserve and resilience (Clegg et al. 2013; Hoogendijk et al. 2019). Rather than resulting from a single pathological process, frailty arises through the interaction of chronic low-grade inflammation, metabolic dysregulation, impaired tissue maintenance, and disrupted inter-organ communication, all of which contribute to increased vulnerability to stressors and adverse health outcomes (Clegg et al. 2013; Hoogendijk et al. 2019). Advances in proteomic and metabolomic technologies have enabled comprehensive characterization of these biological processes and the identification of circulating biomarkers associated with frailty and biological ageing (Bountziouka et al. 2022; Di et al. 2024; Jian et al. 2024). The following sections summarize the major molecular mechanisms implicated in frailty and discuss their relevance to molecular biomarkers of biological ageing.

### 1.3.1 Inflammaging

Chronic low-grade inflammation, or inflammaging, is a hallmark of biological ageing and one of the key mechanisms contributing to frailty (Li et al. 2023; López-Otín et al. 2023; Mak et al. 2025). Unlike acute inflammation, inflammaging is characterized by persistent systemic inflammatory activity that develops with age in the absence of overt infection. It is thought to arise from multiple age-related processes, including cellular senescence, mitochondrial dysfunction, impaired autophagy, and alterations in innate and adaptive immunity (Ferrucci and Fabbri 2018; Franceschi and Campisi 2014; Franceschi et al. 2018; Kennedy et al. 2014).

Persistent inflammation disrupts tissue homeostasis and impairs the function of multiple organ systems, contributing to muscle loss, endothelial dysfunction, reduced regenerative capacity, and declining physiological reserve. Consequently, inflammaging has been consistently associated with frailty, sarcopenia, cardiovascular disease, cognitive impairment, and mortality (Clegg et al. 2013; Di et al. 2024; Franceschi and Campisi 2014). Circulating inflammatory proteins are also associated with frailty severity and adverse clinical outcomes, supporting their potential as biomarkers of biological ageing (Franceschi et al. 2018; Sayed et al. 2021; Xu et al. 2025).

### 1.3.2 Metabolic dysregulation

Metabolic dysregulation is a fundamental feature of biological ageing and involves coordinated alterations in energy metabolism, mitochondrial function, glucose homeostasis, lipid metabolism, and amino acid turnover (Kennedy et al. 2014; López-Otín et al. 2023; Zhang et al. 2024). These

changes impair cellular energy production, reduce metabolic flexibility, and compromise adaptive responses to physiological stress. Because metabolic pathways are closely linked to immune regulation, endocrine signaling, and tissue maintenance, metabolic dysfunction contributes to the progressive decline in physiological reserve that characterizes ageing and frailty (Kennedy et al. 2014; López-Otín et al. 2023; Zhang et al. 2024).

Evidence from experimental and epidemiological studies indicates that metabolic abnormalities are associated with impaired muscle function, reduced physical performance, mobility limitation, disability, and mortality in older adults (Kennedy et al. 2014; Rattray et al. 2019). Metabolomic profiling has identified circulating metabolic signatures associated with frailty, biological ageing, and long-term health outcomes, providing a systems-level view of metabolic alterations across multiple biochemical pathways (Deelen et al. 2019; Liu et al. 2025a; Mak et al. 2023). These molecular signatures may improve understanding of the biological mechanisms underlying frailty and support the identification of biomarkers for risk stratification.

### 1.3.3 Tissue remodeling and inter-organ communication

Biological ageing is accompanied by progressive tissue remodeling, including impaired regeneration, extracellular matrix reorganization, and loss of structural integrity, which together contribute to declining organ function and reduced physiological reserve (Ferrucci and Fabbri 2018; Kennedy et al. 2014; López-Otín et al. 2013). Age-related changes also disrupt communication between organs through alterations in endocrine signaling, inflammatory mediators, metabolic regulation, and other circulating factors. As a result, dysfunction in one organ system can propagate to others, accelerating functional decline and increasing susceptibility to frailty (Ferrucci and Fabbri 2018; Mak et al. 2025).

Because ageing affects multiple tissues simultaneously, circulating proteins and metabolites provide an integrated view of biological processes occurring throughout the body rather than within individual organs. Molecular profiling of these circulating biomarkers therefore offers a systems-level approach to investigating tissue remodeling, inter-organ communication, and the biological mechanisms underlying frailty and biological ageing (Fried et al. 2001; Jia et al. 2026; Kim and Rockwood 2024; Mak et al. 2023; Takeda et al. 2020).

### 1.3.4 Molecular biomarkers of frailty

Frailty arises from the cumulative dysregulation of multiple physiological systems, making it unlikely that a single biomarker can fully reflect its biological complexity (Cao et al. 2025; Jia et al.

2026; Pu et al. 2025; Pu et al. 2024; Zhang et al. 2024). Although clinical frailty instruments provide important information on functional status and vulnerability, they offer limited insight into the molecular processes underlying frailty development (Clegg et al. 2013; Hoogendijk et al. 2019). Numerous circulating biomarkers, including inflammatory proteins, metabolic intermediates, hormones, oxidative stress markers, and molecules involved in tissue remodeling and immune regulation, have been associated with frailty (Ferrucci and Fabbri 2018; Zhang et al. 2024). These findings support the use of molecular biomarkers to complement clinical assessment and improve understanding of biological ageing. The following chapter introduces proteomics and metabolomics as complementary approaches for characterizing these molecular processes at a systems level.

## 1.4 Multi-omics approaches to biological ageing

### 1.4.1 From phenotype to molecular signatures

Traditional measures of ageing, including frailty, disability, multimorbidity, and mortality, describe the clinical consequences of biological ageing but provide limited insight into the molecular processes underlying functional decline (Clegg et al. 2013; Ferrucci and Fabbri 2018; Hoogendijk et al. 2019). Individuals with similar clinical phenotypes may therefore differ substantially in their underlying biological state, highlighting the need for molecular approaches that better capture biological heterogeneity.

Molecular signatures integrate information from multiple biological pathways, including inflammation, metabolism, immune function, and tissue remodeling, providing a systems-level view of biological ageing (Mak et al. 2023; Wang et al. 2025; Yao et al. 2024). High-throughput omics technologies have enabled the simultaneous measurement of thousands of circulating proteins and metabolites, facilitating the identification of molecular signatures associated with frailty, biological ageing, and mortality (Sayed et al. 2021; Zhang et al. 2024). Compared with individual biomarkers, these approaches more comprehensively characterize the biological heterogeneity of ageing and are well suited to large population-based studies (Xiong et al. 2025). The following sections introduce proteomics and metabolomics, two complementary approaches that form the molecular basis of the studies presented in this dissertation.

### 1.4.2 Metabolomics

Metabolomics is the large-scale analysis of small molecules (metabolites) in biological samples such as blood, urine, and tissues. Because metabolites represent the downstream products of gene

expression, protein activity, and environmental influences, they provide a dynamic snapshot of physiological function and biological state (Johnson et al. 2016; Patti et al. 2012). Unlike static genomic information, the metabolome responds rapidly to ageing, lifestyle, disease, and environmental exposures, making metabolomics well suited to investigating biological ageing.

Metabolomic profiling has identified molecular signatures associated with chronological age, physical function, frailty, and mortality, reflecting coordinated alterations across multiple metabolic pathways rather than isolated biochemical changes (Wang et al. 2023; Zhang et al. 2026). This systems-level perspective is particularly relevant to frailty, which arises from dysregulation across multiple physiological systems. In population-based studies, metabolomics has become an important tool for identifying biomarkers, characterizing biological heterogeneity, and improving understanding of the metabolic mechanisms underlying ageing (Guo et al. 2024; Zhang et al. 2026). Although metabolite profiles are influenced by factors such as diet, medication use, and circadian rhythms, continued improvements in analytical platforms and computational methods have enhanced the reproducibility and interpretability of metabolomic data, supporting its application in ageing research (Wang et al. 2023; Wishart 2019).

### 1.4.3 Proteomics

Proteomics is the large-scale analysis of proteins, including their abundance, structure, and biological function. Because proteins mediate most cellular processes, the circulating proteome provides a dynamic representation of physiological and pathological states. Unlike genomic information, which reflects inherited susceptibility, proteomic profiles capture biological responses to ageing, environmental exposures, lifestyle, and disease, making proteomics a valuable approach for investigating biological ageing (Aebersold and Mann 2016; Argentieri et al. 2024; Geyer et al. 2016).

Proteomic profiling has identified proteins associated with chronological age, frailty, physical function, and mortality, reflecting coordinated changes in inflammation, immune regulation, extracellular matrix remodeling, tissue repair, and cellular signaling (Argentieri et al. 2024; Ferrucci and Fabbri 2018). High-throughput platforms now enable the simultaneous quantification of thousands of circulating proteins from small-volume blood samples, facilitating the identification of molecular signatures associated with ageing and age-related outcomes (Franceschi et al. 2018; Sayed et al. 2021). Although circulating protein concentrations are influenced by physiological and environmental factors, continued improvements in analytical technologies and computational methods have enhanced the reproducibility of proteomic studies (Ding et al. 2025; Macdonald-

Dunlop et al. 2022). Together with metabolomics, proteomics provides complementary information on the molecular mechanisms underlying frailty and biological ageing.

### 1.4.4 Integrating metabolomics and proteomics

Proteomics and metabolomics provide complementary information on the biological processes underlying ageing and frailty. Proteomic profiling primarily reflects changes in cellular signaling, immune regulation, tissue remodeling, and other functional pathways, whereas metabolomic profiling captures alterations in energy metabolism and biochemical homeostasis (Hasin et al. 2017; Subramanian et al. 2020). Because proteins regulate metabolic processes and metabolites represent their downstream products, integrating these molecular layers provides a more comprehensive view of physiological function than either approach alone (Hasin et al. 2017; Subramanian et al. 2020).

Multi-omics analyses enable the characterization of interconnected biological pathways that contribute to frailty and biological ageing and better capture the systemic nature of ageing than single biomarkers or individual omics platforms (Cao et al. 2025; Jia et al. 2026; Pu et al. 2025; Pu et al. 2024). In population-based studies, integrating proteomic and metabolomic data facilitates the identification of molecular signatures associated with frailty and mortality, improves biological interpretation, and supports biomarker discovery and risk prediction (Cao et al. 2025; Jia et al. 2026; Karczewski and Snyder 2018; Xiong et al. 2025). These complementary approaches form the molecular basis of the analyses presented in this dissertation.

## 1.5 Mortality as an integrative outcome of biological ageing

### 1.5.1 Mortality as the ultimate ageing outcome

Mortality is the ultimate outcome of biological ageing and the most comprehensive endpoint in ageing research. Unlike individual age-related diseases, which reflect dysfunction within specific organs or physiological systems, mortality results from the cumulative decline in physiological reserve and resilience across the whole organism (Clegg et al. 2013; Hoogendijk et al. 2019; López-Otín et al. 2013). It therefore reflects the combined effects of genetic susceptibility, environmental exposures, lifestyle factors, chronic diseases, and the biological processes that drive ageing. Although chronological age is a strong predictor of mortality, survival varies considerably among individuals of the same age, reflecting substantial heterogeneity in biological ageing. Clinical measures such as frailty improve mortality prediction by capturing deficits in physiological

reserve that are not reflected by chronological age alone, but they remain downstream manifestations of the ageing process and do not fully explain the underlying biological mechanisms (Clegg et al. 2013; Peng et al. 2026; Rockwood and Mitnitski 2007).

### 1.5.2 Risk prediction beyond chronological age

Chronological age provides an important estimate of population-level mortality risk but does not adequately reflect individual differences in biological ageing (Argentieri et al. 2024; Ganna and Ingelsson 2015; Lehallier et al. 2019). Older adults of the same age often differ substantially in physiological reserve, functional capacity, disease burden, and survival. Measures that more closely reflect biological ageing, including frailty, physical performance, and multimorbidity, improve prediction beyond chronological age by capturing deficits in physiological reserve (Argentieri et al. 2024; Lau et al. 2024). However, these phenotype-based measures generally identify ageing-related decline only after functional impairment has become clinically apparent and provide limited insight into the biological processes underlying disease progression. This has driven increasing interest in molecular approaches that complement conventional clinical assessment and provide a more direct representation of biological ageing (Aebersold and Mann 2016; Johnson et al. 2016; Mutz et al. 2024).

### 1.5.3 Molecular prediction of mortality

Molecular profiling provides complementary information for mortality prediction by capturing biological processes that precede clinical deterioration (Argentieri et al. 2024; Ganna and Ingelsson 2015; Lehallier et al. 2019). Population-based proteomic and metabolomic studies have identified circulating proteins and metabolites associated with all-cause mortality, particularly those involved in inflammation, immune regulation, extracellular matrix remodeling, lipid and amino acid metabolism, and energy homeostasis (Argentieri et al. 2024; Franceschi et al. 2018; Yan et al. 2022). Integrating multiple molecular markers better reflects the systemic nature of biological ageing than individual biomarkers and provides insight into the biological pathways contributing to survival (Aebersold and Mann 2016; Ma et al. 2026; Wang et al. 2023). Although challenges remain regarding analytical platforms, study populations, and biomarker reproducibility, continued advances in molecular technologies have strengthened the application of proteomic and metabolomic profiling in population-based ageing research (Aebersold and Mann 2016; Johnson et al. 2016; Zhang et al. 2024).

## 1.6 Molecular measures of biological ageing

### 1.6.1 Biological ageing clocks

Biological ageing clocks estimate biological rather than chronological age by integrating molecular information that reflects physiological function and the ageing process (Belsky et al. 2022; Jylhävä et al. 2017; Levine 2013). Unlike chronological age, which measures time since birth, these models quantify inter-individual differences in biological ageing and aim to detect age-related changes before they become clinically apparent, thereby improving the assessment of health status and future disease risk (Belsky et al. 2022; Jylhävä et al. 2017; Levine 2013).

Early biological ageing measures were based on clinical biomarkers and physiological characteristics, followed by molecular clocks derived from telomere length, DNA methylation, transcriptomics, proteomics, and metabolomics (Belsky et al. 2022; Horvath 2013). Among these, epigenetic clocks were the first widely validated molecular measures of biological ageing, while recent proteomic and metabolomic clocks have extended this concept by capturing dynamic changes in inflammation, metabolism, immune regulation, and tissue homeostasis (Belsky et al. 2022; Horvath 2013; Tanaka et al. 2020). Biological age estimates often differ substantially among individuals of the same chronological age, and accelerated biological ageing has consistently been associated with frailty, chronic disease, functional decline, and mortality (Kennedy et al. 2014; Lehallier et al. 2019). The following sections focus on proteomic and metabolomic ageing clocks, which are the molecular approaches investigated in this dissertation.

### 1.6.2 Metabolomics-based ageing clocks

Metabolomics-based ageing clocks estimate biological age by integrating information from large numbers of circulating metabolites using statistical and machine-learning approaches (Franck et al. 2025; Robinson and Lau 2020). Rather than relying on individual metabolites, these models capture coordinated alterations across multiple metabolic pathways to quantify biological ageing.

In population-based studies, metabolomic age correlates strongly with chronological age while simultaneously capturing substantial inter-individual variation in ageing trajectories. Individuals with accelerated metabolomic ageing consistently show increased risks of frailty, cardiometabolic disease, functional decline, and all-cause mortality (Franck et al. 2025; Robinson and Lau 2020). These findings suggest that metabolomic ageing captures biological information beyond chronological age and may improve risk stratification in ageing research.

Although metabolomic ageing clocks remain influenced by differences in analytical platforms and study populations, ongoing improvements in metabolomic profiling and model development continue to enhance their robustness and reproducibility (Liu et al. 2025a).

### 1.6.3 Proteomics-based ageing clocks

Proteomic ageing clocks estimate biological age by integrating information from large numbers of circulating proteins using statistical and machine-learning approaches (Ferrucci et al. 2020; Johnson et al. 2020; Oh et al. 2023). Rather than relying on individual proteins, these models combine multiple protein signatures to capture biological ageing at the systems level.

In population-based cohorts, proteomic age correlates strongly with chronological age while revealing substantial inter-individual variation in ageing trajectories. Accelerated proteomic ageing has consistently been associated with frailty, multimorbidity, functional decline, and all-cause mortality, supporting its value as a measure of biological ageing (Ferrucci et al. 2020; Johnson et al. 2020; Oh et al. 2023). Compared with conventional clinical indicators, proteomic ageing clocks provide quantitative measures of biological ageing that may improve risk stratification and facilitate investigation of ageing-related mechanisms (Lehallier et al. 2020; Ma et al. 2026).

Although differences in analytical platforms, protein panels, and model development strategies remain important challenges, ongoing advances in proteomic technologies and increasing availability of large population-based datasets continue to improve the robustness and reproducibility of proteomic ageing clocks (Coenen et al. 2023; Johnson et al. 2020; Lehallier et al. 2019; Lehallier et al. 2020; Ma et al. 2026; Macdonald-Dunlop et al. 2022; Oh et al. 2023; Tanaka et al. 2020).

### 1.6.4 Current knowledge gaps

Despite considerable advances in molecular ageing research, important knowledge gaps remain. Proteomic and metabolomic studies have improved understanding of the biological processes underlying ageing, and molecular ageing clocks have shown promising associations with frailty, chronic disease, and mortality. However, biological ageing is a complex multisystem process that cannot be fully explained by individual biomarkers or single molecular layers.

First, most previous studies have investigated metabolomic or proteomic markers separately, despite evidence that biological ageing arises from interactions among multiple interconnected biological pathways. Integrating complementary molecular information may therefore provide a more comprehensive representation of physiological ageing than analyses based on a single omics platform.

Second, many molecular biomarker studies have been conducted in selected patient populations or relatively small cohorts, limiting the generalizability and reproducibility of their findings. Large, well-characterized population-based cohorts with detailed clinical phenotyping are needed to evaluate the robustness and clinical relevance of molecular biomarkers across heterogeneous ageing populations.

Finally, frailty, mortality, and biological ageing have largely been investigated as separate research topics. Few population-based studies have examined clinical manifestations of ageing, long-term survival, and molecular measures of biological ageing within a unified analytical framework. Integrating these outcomes may improve understanding of the biological heterogeneity of ageing and facilitate the identification of clinically relevant biomarkers.

Together, these knowledge gaps support the need for integrative population-based studies combining proteomic and metabolomic profiling with clinically relevant ageing phenotypes and outcomes. Such approaches may improve understanding of the molecular mechanisms underlying frailty and mortality while contributing to the development and validation of molecular measures of biological ageing.

## 1.7 Integrative research framework of the dissertation

This dissertation is based on the premise that biological ageing is a multidimensional process spanning molecular alterations, functional decline, and adverse clinical outcomes. Frailty, mortality, and molecular measures of biological ageing represent complementary manifestations of the same underlying biological processes and should therefore be investigated within an integrated framework rather than as isolated outcomes.

Guided by this concept, the four studies included in this dissertation examine biological ageing across different levels of organization (**Figure 1**). The first study focuses on frailty as a clinical manifestation of biological ageing. The second study investigates molecular signatures underlying frailty using circulating proteomic and metabolomic profiles. The third and fourth studies extend this framework by evaluating metabolomics- and proteomics-based measures of biological ageing and their relationships with mortality and other ageing-related outcomes. Together, these studies integrate clinical epidemiology with molecular epidemiology to improve understanding of the biological heterogeneity of ageing in population-based cohorts.

This research framework provides the rationale for the research aims presented in the following chapter and illustrates how the individual studies collectively contribute to a more comprehensive understanding of biological ageing.

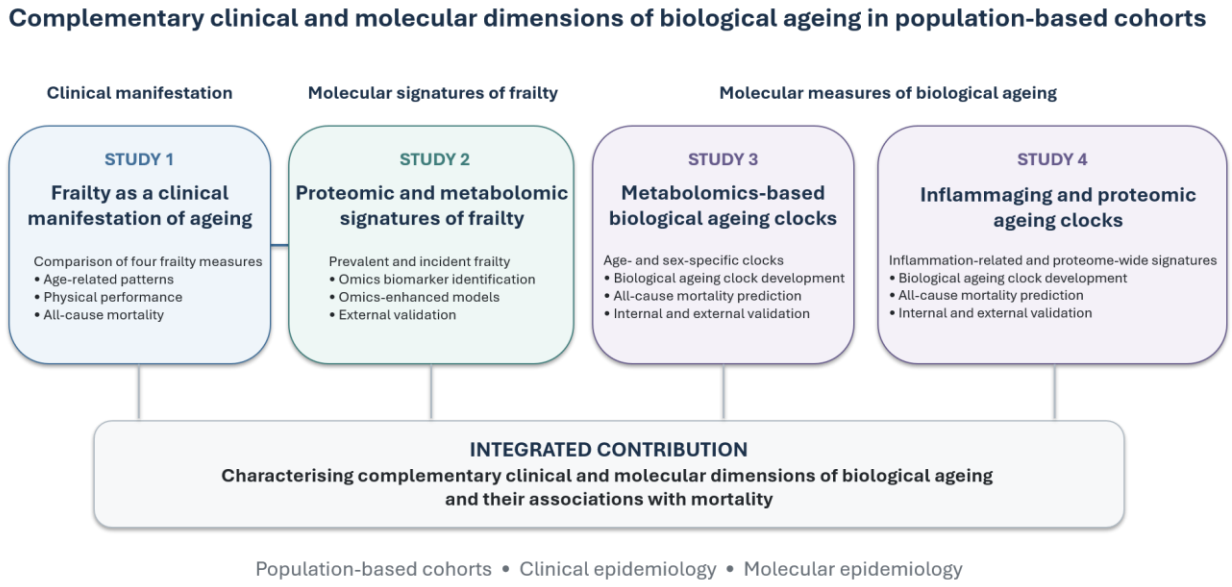



**Figure 1. Overview of the dissertation: Four complementary studies from clinical assessment to molecular ageing clocks and mortality prediction**

## 1.8 Overall objectives and specific research questions

The overall aim of this dissertation was to improve the assessment of biological ageing by integrating clinical frailty measures with proteomic and metabolomic profiling in large population-based cohorts, and to evaluate their complementary value for characterizing biological ageing and predicting adverse ageing-related outcomes.

To address this overall aim, the dissertation comprised four complementary studies:

*Study 1: Comparison of Four Frailty Measurements in Relation to Ageing, Physical Performance and All-Cause Mortality*

This study aimed to compare four commonly used frailty measurements with respect to their prevalence, age-related patterns, associations with physical performance, and prediction of all-cause mortality.

*Study 2: Integrated Proteomic and Metabolomic Profiling for Frailty Identification and Prediction*

This study aimed to identify proteomic and metabolomic signatures associated with prevalent and incident frailty and to develop and externally validate omics-enhanced models for frailty identification and prediction.

*Study 3: Age- and Sex-Specific Metabolomics-Based Biological Ageing Clocks for 10-Year Mortality Prediction*

This study aimed to develop age- and sex-specific metabolomics-based biological ageing clocks and to evaluate their associations with and predictive performance for 10-year mortality.

*Study 4: Novel Inflammaging and Proteomic Ageing Clocks as Surrogates of Biological Age Predict All-Cause Mortality*

This study aimed to characterize inflammaging and proteome-wide signatures of biological ageing, develop proteomic ageing clocks, and evaluate their associations with all-cause mortality.

Collectively, these four studies examine complementary clinical and molecular dimensions of biological ageing, progressing from the assessment of frailty as a clinical phenotype to the identification of molecular signatures of frailty and the development of metabolomic and proteomic measures of biological ageing for mortality prediction.

# 2 MATERIALS AND METHODS

## 2.1 Study 1: Comparison of four frailty measurements

### 2.1.1 Study design and participant selection

The analyses utilize data from the ESTHER study (German name: *Epidemiologische Studie zu Chancen der Verhütung Früherkennung und optimierten Therapie chronischer Erkrankungen in der älteren Bevölkerung*), a comprehensive population-based cohort study conducted in Saarland, Germany (Saum et al. 2014). Baseline recruitment occurred between July 2000 and December 2002 during routine general health check-ups conducted by general practitioners (GPs), resulting in an initial cohort of 9,940 participants.

Follow-up assessments were conducted at the 2-, 5-, 8-, 11-, 14-, 17-, and 20-year follow-up waves (2002–2004, 2004–2007, 2007–2010, 2010–2013, 2014–2016, 2017–2019, and 2020–2022, respectively). At each follow-up wave, participants completed standardized questionnaires on sociodemographic characteristics, medical history, and lifestyle factors.

In addition to questionnaire-based follow-up, participants were invited to consent to in-home visits by study physicians at the 8-, 11-, and 14-year follow-up waves. These home visits included detailed geriatric and pharmacological assessments. A total of 3,124 participants completed the first home visit (8-year follow-up), 2,761 completed the second home visit (11-year follow-up), and 2,342 completed the third home visit (14-year follow-up).

The present study uses data from the third home visit as the analytic baseline, which was conducted between October 2014 and September 2016. This follow-up wave was selected as the analytic baseline because it was the earliest time point at which complete data on all four frailty instruments and related assessments required for the present analyses were available.

The only inclusion criterion was participation in the third home visit (conducted from 10/2014 to 09/2016), to which all ESTHER participants were invited. This time point served as the baseline for the present analysis. Of 2,342 home-visit participants, 342 were excluded due to age outside 63–89 years or completely missing frailty data, leaving 2,000 for final analysis (**Figure 2**).

Details of ethics approval and consent to participate are provided in Section 2.2.1.

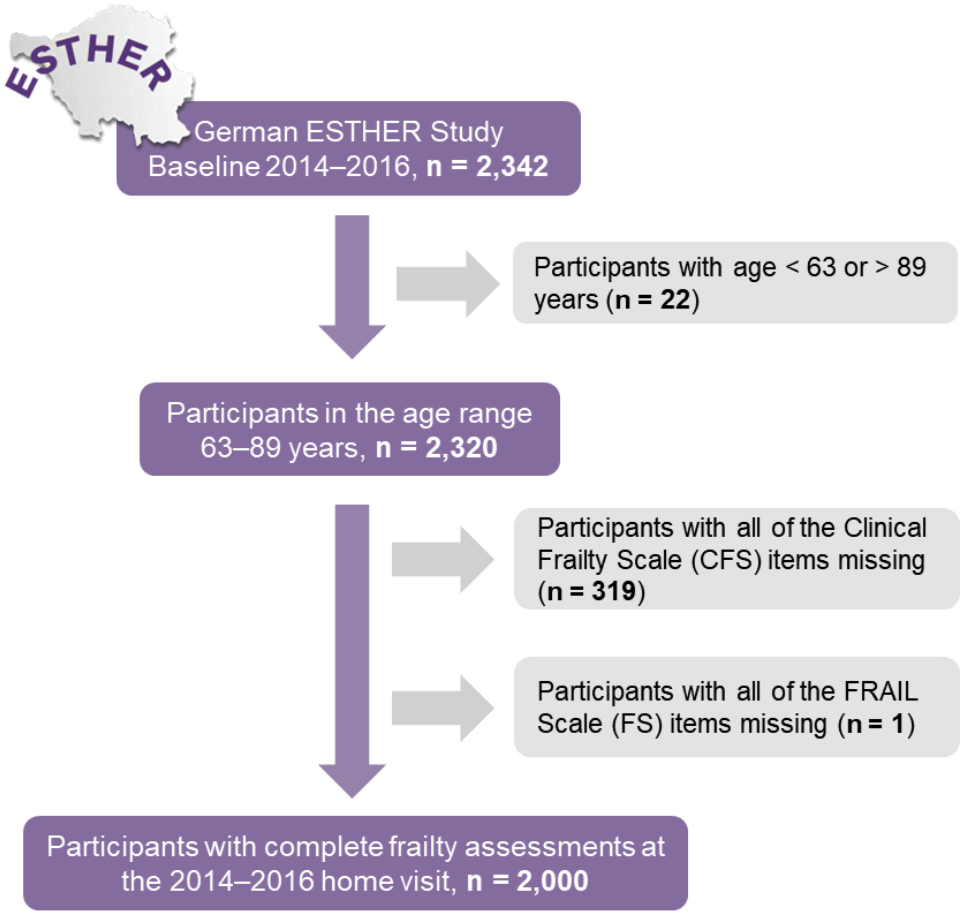



**Figure 2. Flowchart of study participant selection**

### 2.1.2 Frailty measurements and Short Physical Performance Battery (SPPB)

(1) Fried Frailty Phenotype (FFP)

The FFP was proposed and validated by Fried and colleagues in the Cardiovascular Health Study (Fried et al. 2001). The measurement scale includes five dimensions: unintentional weight loss (10 lbs in the past year), self-reported exhaustion, weakness (grip strength), slow walking speed, and low physical activity (Fried et al. 2001). I used population-independent cut-offs for grip strength, walking speed and physical activity as described previously (Saum et al. 2012). If three or more criteria are met, subjects are considered "frail", if one or two criteria are met as "pre-frail", and if no criterion is met as "non-frail" (Schöttker et al. 2014). Frailty (≥ 3), pre-frailty (1–2), and non-frailty (0) classifications followed standard criteria.

(2) Clinical Frailty Scale (CFS)

The 7-point CFS was developed and validated by Rockwood and colleagues in the Canadian Study of Health and Aging (CSHA) (Rockwood et al. 2005). In this study, I labeled the category "very fit" as "non-frail", I combined the categories "well" and "well, with treated comorbid disease" as

the “pre-frail” group; and I combined the categories “apparently vulnerable”, “mildly frail”, “moderately frail”, and “severely frail” as the “frail” group. As the CFS requires clinical assessment, frailty status was rated by the participants' GPs based on their clinical judgement according to the following information: “Frailty is a geriatric syndrome characterized by reduced resilience to external stressors. For those affected, it is associated with an increased risk of loss of independence, hospitalization and mortality.” This categorization of the CFS into three levels (non-frail, pre-frail, frail) has been used in previous studies (Church et al. 2020; Theou et al. 2013).

(3) FRAIL Scale (FS)

The FS comprises five components: fatigue, resistance, ambulation, illness, and weight loss (Abellan van Kan et al. 2008). Fatigue was defined as feeling tired “all of the time” or “most of the time” in the past four weeks. Resistance and ambulation were assessed by reported difficulty climbing 10 steps or walking several hundred yards, respectively. Illness was defined as having five or more of the following conditions: hypertension, diabetes, cancer, chronic lung disease, myocardial infarction, heart failure, coronary artery disease, pulmonary embolism, arthritis, stroke, or renal failure. Weight loss was defined as an unintentional loss of ≥5% of body weight in the past 12 months and was scored as 1 point. Each item scores 1 if present (0 otherwise); total 0–5 with frail 3–5, pre-frail 1–2, non-frail 0 (Morley et al. 2012). In this study, the FS was obtained from the GP-administered questionnaire.

(4) Frailty Index (FI)

The FI was developed by Mitnitski and Rockwood and defines frailty as an accumulation of age-related deficits (Mitnitski et al. 2001). These deficits can be symptoms, disabilities, diseases or disease markers and at least 30 deficits need to be included, which need to fulfil the following criteria: accumulation with age, association with the general health status, not having a high prevalence at younger ages, not having a low prevalence at older age, and not being repetitive (Searle et al. 2008). In this study, the continuous FI was based on 31 deficits (**Table 1**) following a standardized procedure described previously for the ESTHER study (Schöttker et al. 2014). Each item was recoded such that a value of 0 indicates its absence and 1 indicates its presence, with scaled variables assigned intermediate values as appropriate. The FI is expressed as the ratio of deficits present to the total number of deficits considered. To harmonize frailty classification across instruments, I applied empirical cut-offs for the FI that approximated the distribution of frailty and pre-frailty categories defined by the FFP (**Table 2**). The following cut-offs were obtained: frail ($FI > 0.31$), pre-frail ($0.09 < FI \leq 0.31$), and non-frail ($FI \leq 0.09$).

**Table 1. Items used to calculate the FI, FFP, CFS, FS, and SPPB in the ESTHER study**

| No. | Variable | Coding with assigned value | Frequency (%) | Proportion of missing values (%) [a] |
|---|---|---|---|---|
| **The chosen 31 deficits for the Rockwood frailty index** | | | | |
| | **Chronic diseases** | | | |
| 1 | Coronary artery disease | "Yes" = 1; "No" = 0 | 523 (26.2) | 0.1 |
| 2 | Heart failure | "Yes" = 1; "No" = 0 | 372 (18.7) | 0.6 |
| 3 | Diabetes | "Yes" = 1; "No" = 0 | 560 (28.1) | 0.2 |
| 4 | Cancer | "Yes" = 1; "No" = 0 | 397 (19.9) | 0.0 |
| 5 | Glaucoma | "Yes" = 1; "No" = 0 | 149 (7.5) | 0.5 |
| 6 | Cataract | "Yes" = 1; "No" = 0 | 671 (33.7) | 0.3 |
| 7 | Parkinson's disease | "Yes" = 1; "No" = 0 | 94 (4.7) | 0.9 |
| | **History of major disease events** | | | |
| 8 | Myocardial infarction | "Yes" = 1; "No" = 0 | 152 (7.8) | 2.2 |
| 9 | Stroke | "Yes" = 1; "No" = 0 | 243 (12.4) | 2.2 |
| 10 | Femur fracture | "Yes" = 1; "No" = 0 | 53 (2.7) | 2.4 |
| 11 | Joint replacement | "Yes" = 1; "No" = 0 | 355 (17.8) | 0.3 |
| | **Use of drugs** | | | |
| 12 | Anti-hypertensives | "Yes" = 1; "No" = 0 | 1,468 (74.1) | 0.9 |
| 13 | Lipid lowering drugs | "Yes" = 1; "No" = 0 | 714 (36.0) | 0.9 |
| 14 | Vasodilators | "Yes" = 1; "No" = 0 | 132 (6.7) | 0.9 |
| 15 | Cardiac glycosides | "Yes" = 1; "No" = 0 | 33 (1.7) | 0.9 |
| 16 | Prescribed aspirin | "Yes" = 1; "No" = 0 | 612 (30.9) | 0.9 |
| 17 | Anti-osteoporotic drugs | "Yes" = 1; "No" = 0 | 233 (11.8) | 0.9 |
| 18 | Anxiolytics | "Yes" = 1; "No" = 0 | 49 (2.5) | 0.9 |
| 19 | Sedatives | "Yes" = 1; "No" = 0 | 82 (4.1) | 0.9 |
| 20 | Antidementia drugs | "Yes" = 1; "No" = 0 | 60 (3.0) | 0.9 |
| 21 | Drugs for prostatic hyperplasia or urinary incontinence | "Yes" = 1; "No" = 0 | 223 (11.3) | 0.9 |
| | **Difficulties in Activities of Daily Living (ADL)** | | | |
| 22 | Problems with averagely difficult tasks | "Yes, severely" = 1; | 153 (8.1) | 5.5 |
| | | "Yes, a little" = 0.5; | 621 (32.9) | |
| | | "No, none" = 0 | 1,116 (59.0) | |
| 23 | Problems climbing several flights of stairs | "Yes, severely" = 1; | 211 (11.0) | 4.1 |
| | | "Yes, a little" = 0.5; | 704 (36.8) | |
| | | "No, none" = 0 | 997 (52.1) | |
| 24 | Impairments due to pain | "Yes, severely" = 1; | 82 (4.2) | 2.4 |
| | | "Yes, rather severely" = 0.75; | 257 (13.2) | |
| | | "Yes, moderately" = 0.5 | 358 (18.3) | |
| | | "Yes, a little" = 0.25; | 652 (33.4) | |
| | | "No, not at all" = 0 | 603 (30.9) | |
| 25 | Impairments affecting social contacts | "Always" = 1; | 22 (1.2) | 6.4 |
| | | "Mostly" = 0.75; | 85 (4.5) | |
| | | "Sometimes" = 0.5; | 320 (17.1) | |

| No. | Variable | Coding with assigned value | Frequency (%) | Proportion of missing values (%) [a] |
|---|---|---|---|---|
| | | "Rarely" = 0.25; | 435 (23.2) | |
| | | "Never" = 0 | 1,010 (54.0) | |
| 26 | Managed only specific tasks due to physical impairments | "Yes" = 1; "No" = 0 | 609 (33.0) | 7.6 |
| 27 | Finished less tasks than intended due to mental problems | "Yes" = 1; "No" = 0 | 461 (24.2) | 4.7 |
| | **General health** | | | |
| 28 | General self-rated health | "Very poor" = 1 | 41 (2.1) | 2.3 |
| | | "Poor" = 0.75 | 441 (22.6) | |
| | | "Fair" = 0.5 | 1,176 (60.2) | |
| | | "Good" = 0.25 | 260 (13.3) | |
| | | "Excellent" = 0 | 36 (1.8) | |
| | **Life-style related factors** | | | |
| 29 | Underweight | "BMI < 20" = 1; "BMI >= 20" = 0 | 239 (12.0) | 0.1 |
| 30 | Overweight | "BMI >= 35" = 1 | 97 (4.9) | 0.1 |
| | | "BMI 30-35" = 0.5 | 343 (17.2) | |
| | | "BMI < 30" = 0 | 1,558 (78.0) | |
| 31 | Lack of vigorous physical activity | "Yes" = 1; "No" = 0 | 816 (41.1) | 7.0 |
| **Five Dimensions of the Fried Frailty Phenotype (FFP)** | | | | |
| 1 | Unintentional weight loss (10 lbs in the past year) | "Yes" = 1; "No" = 0 | 114 (5.7) | 0.0 |
| 2 | Self-reported exhaustion | "Yes" = 1; "No" = 0 | 165 (8.3) | 0.9 |
| 3 | Weakness (grip strength) | "Yes" = 1; "No" = 0 | 781 (39.1) | 0.1 |
| 4 | Slow walking speed | "Yes" = 1; "No" = 0 | 833 (41.7) | 0.0 |
| 5 | Low physical activity | "Yes" = 1; "No" = 0 | 1,111 (55.9) | 0.6 |
| **The 7-point Clinical Frailty Scale (CFS)** | | | | |
| | Non-frail | "Very fit" = 1 | 438 (21.9) | 0.0 |
| | Pre-frail | "Well" = 2 | 407 (20.4) | 0.0 |
| | | "Well, with treated comorbid disease" = 3 | 752 (37.6) | |
| | Frail | "Apparently vulnerable" = 4 | 259 (12.9) | 0.0 |
| | | "Mildly frail" = 5 | 101 (5.1) | |
| | | "Moderately frail" = 6 | 32 (1.6) | |
| | | "Severely frail" = 7 | 11 (0.5) | |
| **Five Components of the FRAIL Scale (FS)** | | | | |
| 1 | Fatigue | "Yes" = 1; "No" = 0 | 203 (10.4) | 2.4 |
| 2 | Resistance | "Yes" = 1; "No" = 0 | 46 (2.4) | 2.2 |
| 3 | Ambulation | "Yes" = 1; "No" = 0 | 907 (46.9) | 3.4 |
| 4 | Illness | "Yes" = 1; "No" = 0 | 367 (19.1) | 3.8 |
| 5 | Loss of weight | "Yes" = 1; "No" = 0 | 534 (27.1) | 1.5 |
| **Three Components of the Short Physical Performance Battery (SPPB) (12 Points)** | | | | |
| 1 | **Balance Tests (4 Points)** | | | |

| No. | Variable | Coding with assigned value | Frequency (%) | Proportion of missing values (%) [a] |
|---|---|---|---|---|
| 1.1 | Side-by-side-stand (1 Points) | "Held for 10 sec" = 1; "Not held for 10 sec" = 0 | 1,906 (96.4) | 1.1 |
| 1.2 | Semi-Tandem Stand (1 Points) | "Held for 10 sec" = 1; "Not held for 10 sec" = 0 | 1,774 (89.1) | 1.0 |
| 1.3 | Tandem Stand (2 Points) | "Held for 10 sec" = 2 | 1,422 (71.9) | 1.1 |
| | | "Held for 3 to 9.99 sec" = 1 | 338 (17.1) | |
| | | "Held for less than 3 sec" = 0 | 218 (11.0) | |
| 2 | **Gait Speed Test (3-Meter Walk) (4 Points)** | "If time is more than 6.52 sec" = 1 | 77 (3.9) | 1.2 |
| | | "If time is 4.66 to 6.52 sec" = 2 | 254 (12.8) | |
| | | "If time is 3.62 to 4.65 sec" = 3 | 515 (26.1) | |
| | | "If time is less than 3.62 sec" = 4 | 1,130 (57.2) | |
| 3 | **Chair Stand Test (4 Points)** [b] | "Participant unable to complete 5 chair stands or completes stands in >60 sec" = 0 | 101 (5.5) | 8.2 |
| | | "If chair stand time is 16.70 sec or more" = 1 | 289 (15.7) | |
| | | "If chair stand time is 13.70 to 16.69 sec" = 2 | 357 (19.4) | |
| | | "If chair stand time is 11.20 to 13.69 sec" = 3 | 415 (22.6) | |
| | | "If chair stand time is 11.19 sec or less" = 4 | 674 (36.7) | |

Notes: Missing values were imputed using a random forest-based method (*missRanger* in R). The imputation model primarily included age, sex, comorbidity, and mobility indicators, which directly influence the probability of non-assessment. Education, smoking status and body mass index were additionally considered as auxiliary predictors, but their contribution was secondary.

[a] The proportion of missing values was calculated based on the number of non-respondents to the participant questionnaire or the physician-conducted home visit among the 2,000 participants selected at the 2014–2016 follow-up, which served as the baseline for this study.

[b] The chair-stand item had higher missingness due to safety considerations during home visits.

**Table 2. Cut-off values for the Frailty Index (FI) selected to approximate the prevalence of frailty categories defined by the Fried Frailty Phenotype (FFP) in the ESTHER baseline population**

| Frailty status | Fried Frailty Phenotype (FFP) [a] | | | Frailty Index (FI) [a] | | |
|---|---|---|---|---|---|---|
| | Scores | n | Proportion (%) | Cut-offs | n | Proportion (%) |
| Non-frail | 0 | 390 | 19.5 | ≤0.09 | 399 | 19.9 |
| Pre-frail | 1–2 | 1,248 | 62.4 | >0.09–≤0.31 | 1,227 | 61.4 |
| Frail | 3–5 | 362 | 18.1 | >0.31 | 374 | 18.7 |

[a] The FI cut-offs were empirically set to match FFP category prevalences at baseline; they are data-driven and may not generalize to other populations without recalibration. Missing values for items from the FI and FFP, as detailed in **Table 1** of the ESTHER study, were imputed using random forest imputation with the R package *missRanger* (version 2.6.0).

(5) Short Physical Performance Battery (SPPB)

Physical performance at 2014–2016 follow-up was assessed by study physicians during the home visit with the SPPB, which is a robust measure of lower-extremity function predictive of clinical outcomes (Pavasini et al. 2016; Saum et al. 2015). It comprises three components: balance tests, a 3-meter gait speed test, and a chair stand test. The balance tests assess the ability to maintain side-by-side, semi-tandem, and tandem stands, with scores assigned based on the duration each position is held (maximum 4 points). The gait speed test is scored from 1 to 4 points according to the time required to complete a 3-meter walk. The chair stand test is also scored from 1 to 4 points based on the time taken to perform five consecutive stands. The cumulative SPPB score, ranging from 0 to 12, classifies performance as poor (0–6), suboptimal (7–10), or optimal (11–12).

**Figure 3** shows the cut-offs used to define pre-frailty and frailty for all four frailty instruments and for the SPPB to define poor, sub-optimal and optimal physical performance.

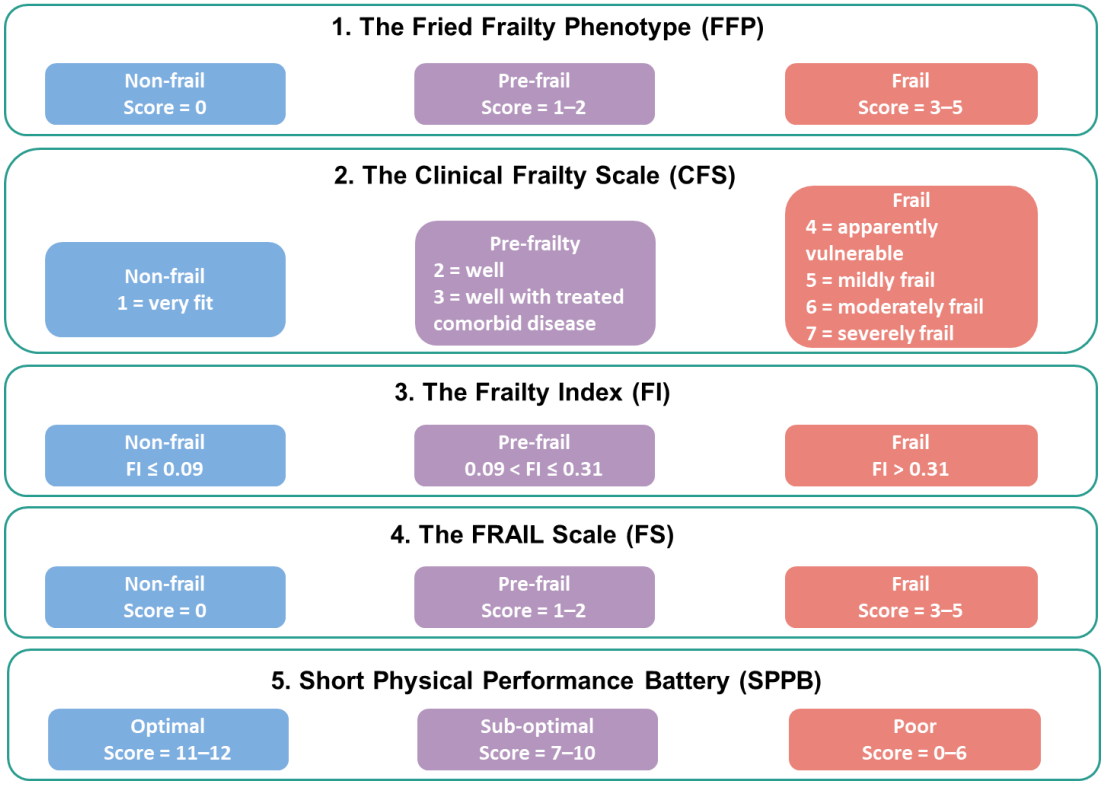



**Figure 3. Definitions and classification criteria of different frailty measures and physical performance instrument**

### 2.1.3 Assessment of covariates

Covariates considered were age, sex, education, smoking status, body mass index (BMI), and the Cumulative Illness Rating Scale-Geriatric (CIRS-G) score. The study participant's age was the time from the date of birth to the date of the baseline assessment. Sex and education were self-reported at baseline. The education level was classified as ≤9 years, 10–11 years, and ≥12 years of

school attendance. Smoking status at baseline was categorized as never, former, or current smoker. BMI was calculated from measured height and weight and further stratified into five clinically relevant categories: <21.5, 21.5–<25, 25–<30, 30–<35, and ≥35 kg/m². Comorbidity burden was quantified using the CIRS-G, which aggregates the severity of chronic diseases across 14 organ systems, rated by the study participant's general practitioners (Miller et al. 1992).

### 2.1.4 Statistical analyses

Weighted kappa coefficients were computed to assess the inter-instrument agreement between frailty instruments to define pre-frailty and frailty. I used the interpretation of kappa coefficients of Checa-López et al.: < 0.01, no agreement; 0.01–0.20, poor agreement; 0.21–0.40, fair agreement; 0.41–0.60, moderate agreement; and 0.61–0.80, substantial agreement (Checa-Lopez et al. 2023). To assess the association of age with (pre-)frailty, unadjusted ordinal logistic regression models were employed with frailty status (non-frail, pre-frail, and frail) as the dependent variable and age as a continuous independent variable.

Logistic regression models, adjusted for age (continuous), sex, education (categorical: ≤9, 10-11, and ≥12 years of school attendance), and smoking status (categorical: never, former, current smoker), were used to estimate odds ratios (ORs) and 95% confidence intervals (CIs) for the associations of pre-frailty and frailty (assessed using the four instruments) with poor physical performance. Similarly, Cox proportional hazards regression models, adjusted for the same covariates, were used to estimate hazard ratios (HRs) and 95% CIs for the associations of pre-frailty and frailty with all-cause mortality. The proportional hazards assumption was assessed using Schoenfeld residuals (Grambsch and Therneau 1994). In sensitivity analyses, the models were extended by additional adjustment for BMI (categorical: <21.5, 21.5–<25, 25–<30, 30–<35, and ≥35 kg/m$^2$) and the CIRS-G score (continuous). These variables were not part of the main model because low BMI and morbidity can be highly correlated with frailty and are also items of some frailty instruments.

The predictive performance of the frailty measures was evaluated using discrimination metrics. For logistic regression models (Hanley and McNeil 1982), the area under the receiver operating characteristic curve (AUC) was calculated and for Cox proportional hazards models, Harrell's C-statistic was employed (Antolini et al. 2004).

All statistical analyses were performed in R, version 4.3.1 (R Foundation for Statistical Computing, Vienna, Austria). The following R packages were used: *missRanger* (version 2.6.0) for random forest imputation, *irr* (version 0.84.1) for Cohen's kappa, *stats* (base R) for logistic regression, *survival* (version 3.5-8) for Cox regression, *pROC* (version 1.18.5) for AUC estimation, *survcomp*

(version 1.52.0) for C-statistic estimation, *boot* (version 1.3-28.1) for bootstrap testing, and *ggplot2* (version 3.5.1) for visualization.

## 2.2 Study 2: Integrated proteomic and metabolomic profiling for frailty

### 2.2.1 Study design and populations

The UKB is a population-based prospective study of 502,414 participants aged 40–69 years, recruited from 22 assessment centers across England, Scotland, and Wales between March 2006 and October 2010 (Bycroft et al. 2018; Sudlow et al. 2015). Each center enrolled individuals living within approximately 40 km of the site. Baseline evaluations included standardized questionnaires and clinical, biochemical, and biomarker assessments covering sociodemographic, behavioral, and health-related domains (Sudlow et al. 2015). Vital status was ascertained via linkage to national death registries (Sudlow et al. 2015), enabling the consideration of death as a competing event.

The ESTHER study (German acronym for *Epidemiologische Studie zu Chancen der Verhütung Früherkennung und optimierten Therapie chronischer Erkrankungen in der älteren Bevölkerung*) is a population-based prospective cohort study conducted in Saarland, Germany (Saum et al. 2014; Schöttker et al. 2013; Schöttker et al. 2022), which was described in Section 2.1.1. Details of ethics approval and consent to participate are provided in Section 2.4.6.

Participants were selected according to predefined inclusion and exclusion criteria in both UKB and ESTHER (**Figure 4**). From the UKB cohort (proteomics measured in a random subset), 19,151 participants aged 60–69 years with available proteomic and metabolomic data and without systemic steroid use were included in baseline analyses. For longitudinal analyses, participants with frailty at baseline (n=1,863, 9.7%) or without follow-up frailty assessments (n=14,742) were excluded, leaving 2,546 non-frail participants, of whom 250 (9.8%) developed incident frailty over a median follow-up of 7.8 years.

Applying similar eligibility criteria in ESTHER yielded 5,031 eligible older adults with available omics data, including 494 (9.8%) who were frail at baseline. For longitudinal analyses, 2,593 non-frail participants were included, among whom 255 developed incident frailty (9.8%) over a median follow-up of 7.7 years.

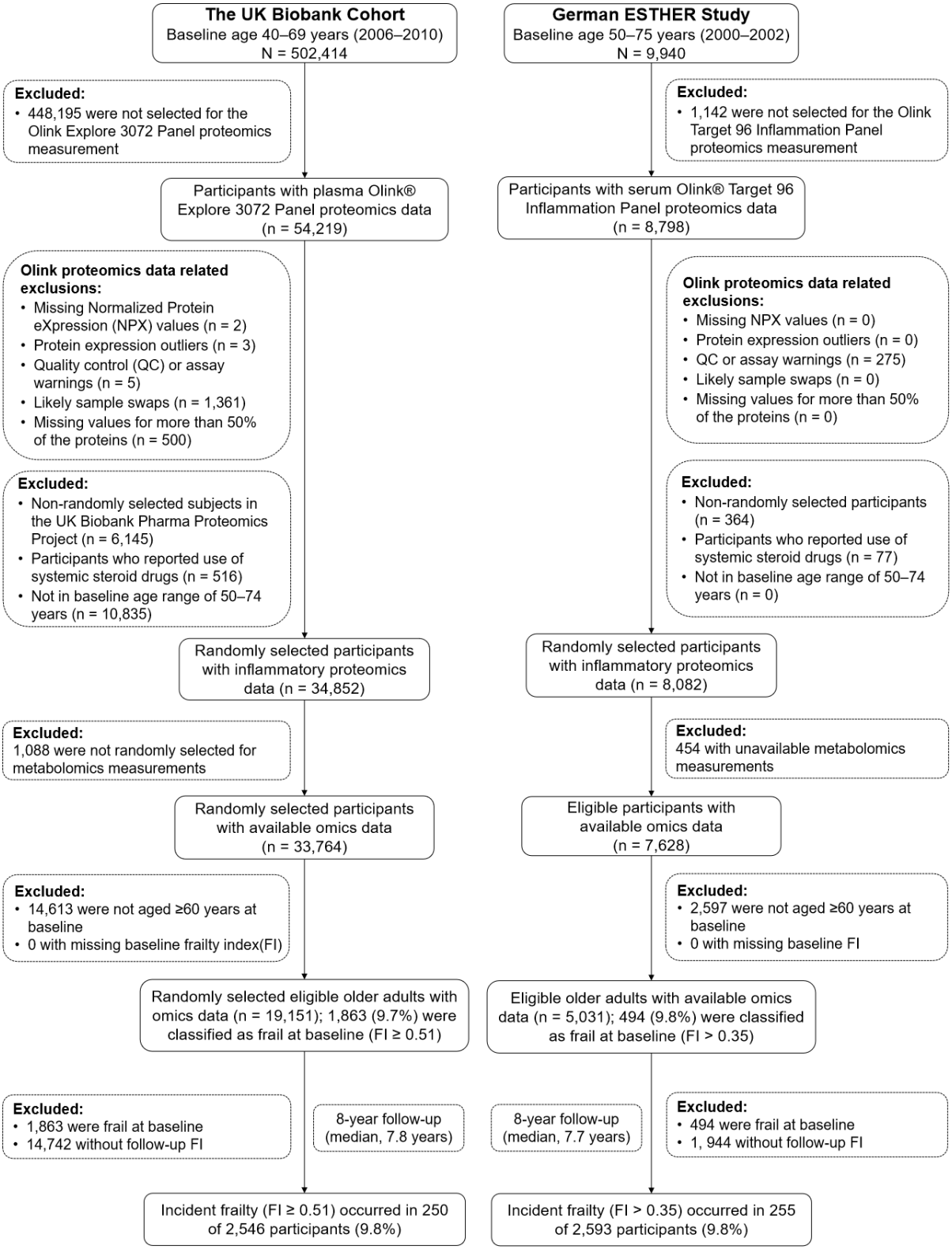



**Figure 4. Flowchart of participant selection and analytical design in the UK Biobank and ESTHER study**

### 2.2.2 Assessment of frailty

Frailty was assessed using a deficit accumulation approach based on the frailty index (FI), as proposed by Rockwood and colleagues and constructed according to established guidelines

(Rockwood and Mitnitski 2007; Searle et al. 2008). The same 31 health deficits were included in the FI in both cohorts, covering multiple physiological domains, including chronic diseases and major disease events, medication use, functional limitations, self-rated health, and lifestyle-related factors. Individual deficits were coded as binary or ordinal variables and rescaled to range from 0 to 1, in accordance with standard FI methodology (Rockwood and Mitnitski 2007; Rockwood et al. 2005; Searle et al. 2008). The FI was calculated for each participant by summing the same 31 deficit scores in UKB and ESTHER and dividing the sum by the total number of items (31), yielding a continuous measure ranging from 0 (no deficits) to 1 (maximum deficit burden), with higher values indicating greater frailty. The same FI construction procedure was applied at baseline and during follow-up to ensure longitudinal consistency.

Frailty status was defined using cohort-specific FI thresholds, as previously validated: FI>0.35 in ESTHER (Schöttker et al. 2014) and FI≥0.51 in UKB (Chen et al. 2024). Incident frailty was defined as the first onset of frailty among non-frail participants (robust and pre-frail) at baseline. Frailty was assessed at baseline, 2, 5, and 8 years in ESTHER, whereas in UKB it was reassessed once at 8 years post-baseline.

### 2.2.3 Proteomics profiling

In UKB, baseline ethylenediaminetetraacetic acid (EDTA) plasma samples were analyzed using the Olink® Explore 3072 platform between May 2021 and November 2022. Protein abundances were reported as normalized protein expression (NPX, $\log_2$ scale) following standard quality control and normalization procedures (Eldjarn et al. 2023; Sun et al. 2023).

In ESTHER, baseline serum samples were analyzed using the Olink® Target 96 Inflammation panel between March 2018 and December 2024. The assay is based on proximity extension assay (PEA) technology, and protein abundances were reported as NPX values following standard quality control and normalization procedures (Trares et al. 2022; Trares et al. 2024).

A total of 67 inflammation-related proteins measured on both platforms and passing quality control (>75% above the limit of detection and <20% missing) in both cohorts were included in the main analyses (**Appendix, Table A1**). NPX values were standardized to z-scores within each cohort to enable comparison of effect estimates across proteins and cohorts.

### 2.2.4 Metabolomics profiling

Baseline circulating metabolomic biomarkers were quantified using a high-throughput nuclear magnetic resonance (NMR) platform (Nightingale Health®, Helsinki, Finland) (Julkunen et al.

2023) in EDTA plasma samples from UKB and serum samples from ESTHER. Measurements were conducted using standard platform-specific quality control and normalization procedures between July 2021 and November 2025 and between September 2022 and May 2023, respectively.

The platform quantified 250 biomarkers, including lipids, lipoprotein subclasses, apolipoproteins, fatty acids, amino acids, ketone bodies, glycolysis-related metabolites, albumin, creatinine, and glycoprotein acetyls (GlycA). Detailed descriptions of the platform have been reported previously (Soininen et al. 2009; Würtz et al. 2017). Glycerol was excluded due to more than 10% missing values in both cohorts, leaving 249 biomarkers for the main analyses. Metabolomic biomarker concentrations were $\log_2$-transformed and standardized to z-scores within each cohort (Altman and Bland 2005; Bland and Altman 1996).

### 2.2.5 Assessment of covariates

Baseline demographic characteristics, lifestyle factors, and medical history were obtained using standardized questionnaires and clinical examinations in both cohorts. In UKB, data were collected at assessment centers through interviews and touchscreen questionnaires, whereas in ESTHER, standardized health examinations and medication documentation were conducted by general practitioners and complemented by self-administered questionnaires.

Education level was classified according to the highest attained qualification. Smoking status was categorized as never, former, or current. Physical activity was assessed using cohort-specific questionnaires (the International Physical Activity Questionnaire in UKB and a structured self-developed questionnaire in ESTHER). Alcohol consumption was assessed from weekly intake of beer, wine, and spirits and converted to grams of pure ethanol per day. BMI was calculated as weight divided by height squared ($kg/m^2$). Medical history (medically treated dyslipidemia, hypertension, diabetes, cardiovascular disease, and cancer) was derived from self-reported physician diagnoses and medication use. Hypertension was additionally defined as systolic/diastolic blood pressure ≥140/90 mm Hg, and diabetes as glycated hemoglobin ≥6.5%.

### 2.2.6 Statistical analyses

Continuous variables are presented as means (standard deviations), and categorical variables as counts and percentages. Missing values were imputed using a random forest-based imputation method (100 trees and 5 iterations) (Stekhoven and Bühlmann 2012), performed separately within each cohort.

Frailty status was assessed at discrete visits, and the exact time of onset was not observed. Time to incident frailty was therefore treated as interval-censored, with the subject-specific event interval defined by the last assessment at which frailty was absent and the first assessment at which frailty was identified.

To identify a parsimonious set of proteomic and metabolomic biomarkers associated with prevalent frailty at baseline, a bootstrap-based least absolute shrinkage and selection operator (LASSO) approach (Wang et al. 2011) was applied in UKB, with proteomic and metabolomic biomarkers jointly entered into a single model for simultaneous feature selection. A total of 1,000 bootstrap resampled datasets were generated, and the LASSO procedure was fitted in each resample. The optimal penalty parameter (λ) was determined using the one-standard-error rule in 10-fold cross-validation (CV) to mitigate overfitting. Biomarkers selected in at least 95% of the bootstrap samples were retained to ensure selection stability. Omics biomarkers were $\log_2$-transformed and standardized to z-scores before model fitting (Altman and Bland 2005; Bland and Altman 1996). The LASSO model was always adjusted for the covariates in the reference model, which included age, sex, education, smoking status, physical activity, alcohol consumption, BMI, dyslipidemia, hypertension, diabetes, cardiovascular disease, and history of cancer. These LASSO-selected biomarkers for prevalent frailty were also used in the longitudinal analyses in both cohorts. As a sensitivity analysis, the same bootstrap-enhanced LASSO feature-selection procedure was repeated using incident frailty cases as the outcome in UKB to evaluate the robustness of biomarker selection. For external validation, the regression coefficients estimated in UKB were directly applied to ESTHER without refitting.

For prevalent frailty at baseline, cross-sectional associations between selected biomarkers and frailty were estimated using a multivariable logistic regression model (TRIM-FI) with effect estimates reported as odds ratios (ORs) and 95% confidence intervals (CIs). Pairwise correlations among selected biomarkers were examined to assess potential multicollinearity.

For incident frailty during follow-up, a multivariable interval-censored competing risks regression model (TRIM-FP) was used (Park et al. 2019), with death treated as a competing event. Effect estimates are presented as subdistribution hazard ratios (sHRs) and 95% CIs. Model assumptions were evaluated by including time-by-covariate interaction terms (Grambsch and Therneau 1994), and no evidence of violation was observed.

Model performance of the TRIM models was compared to the reference model as follows: For prevalent frailty, model discrimination was quantified using the area under the receiver operating characteristic curve (AUC) (Hanley and McNeil 1982). For incident frailty, discrimination was

assessed using the time-dependent Harrell's C-statistic (Antolini et al. 2004). Differences in AUC were tested using DeLong's method (Delong et al. 1988), whereas differences in time-dependent C-statistics were evaluated using bootstrap resampling with 1,000 iterations (Uno et al. 2011). Calibration was evaluated for both cross-sectional and longitudinal analyses using calibration plots comparing observed and model-predicted risks, with observed proportions (for prevalent frailty) and observed risks (for incident frailty) estimated using a locally weighted scatterplot smoothing (loess) approach (Van Calster et al. 2019). Furthermore, the re-classification statistics "continuous net reclassification improvement (NRI)" (Kerr et al. 2014), "categorical NRI", and "integrated discrimination improvement (IDI)" (Pencina et al. 2011) were estimated. The predefined risk categories for the categorical NRI were <5%, 5–10%, and >10% (Ersbøll et al. 2013; Peng et al. 2026). Differences in NRI and IDI were assessed using likelihood ratio-based inference as described by Pencina and colleagues (Pencina et al. 2008; Pencina et al. 2011).

Clinical utility was assessed using decision curve analysis (DCA) (Vickers and Elkin 2006), which quantifies net benefit across a range of threshold probabilities for decision-making. Net benefit by the TRIM models compared to the reference models was assessed across clinically plausible thresholds for both prevalent and incident frailty.

All analyses were performed using R software (version 4.3.1; R Foundation for Statistical Computing, Vienna, Austria). Primary analyses were implemented using the following R packages: *missRanger* (random forest imputation), *glmnet* (LASSO regression), *stats* (multivariable logistic regression), *intccr* (interval-censored competing risks regression), *pROC* (AUC), *timeROC* (time-dependent C-statistic), *nricens* (reclassification metrics), *rms* (calibration), and *rmda* (decision curve analysis). Furthermore, to improve the usability and accessibility of the TRIM models, an interactive web-based application was developed using the R package *shiny*. All statistical tests were two-sided, and statistical significance was defined as a false discovery rate (FDR)-adjusted *P* value <0.05 according to the Benjamini-Hochberg procedure (Benjamini and Hochberg 1995), unless otherwise specified.

## 2.3 Study 3: Metabolomics-based biological ageing clocks for mortality

### 2.3.1 Study design and populations

This work used data from the UKB as the derivation cohort and the ESTHER study as the external validation cohort, which were described in Section 2.2.1 and Section 2.1.1, respectively. Details of ethics approval and consent to participate are provided in Section 2.4.6.

Participants were selected according to predefined inclusion and exclusion criteria in both UKB and ESTHER (**Figure 5**). In UKB, individuals not selected for plasma metabolomics analysis (n=228,017) and those outside the baseline age range of 50–69 years (n=65,205) were excluded, resulting in a final cohort of 209,144 participants. This cohort was randomly divided into a derivation set (70%, n=146,400) and an internal test set (30%, n=62,744) and stratified by sex and age. Finally, both the derivation and internal test set were stratified by sex and age groups (50–59 and 60–69 years). For external validation, the ESTHER was processed using the same inclusion and exclusion criteria, yielding 6,820 eligible participants, who were similarly stratified into four subgroups.

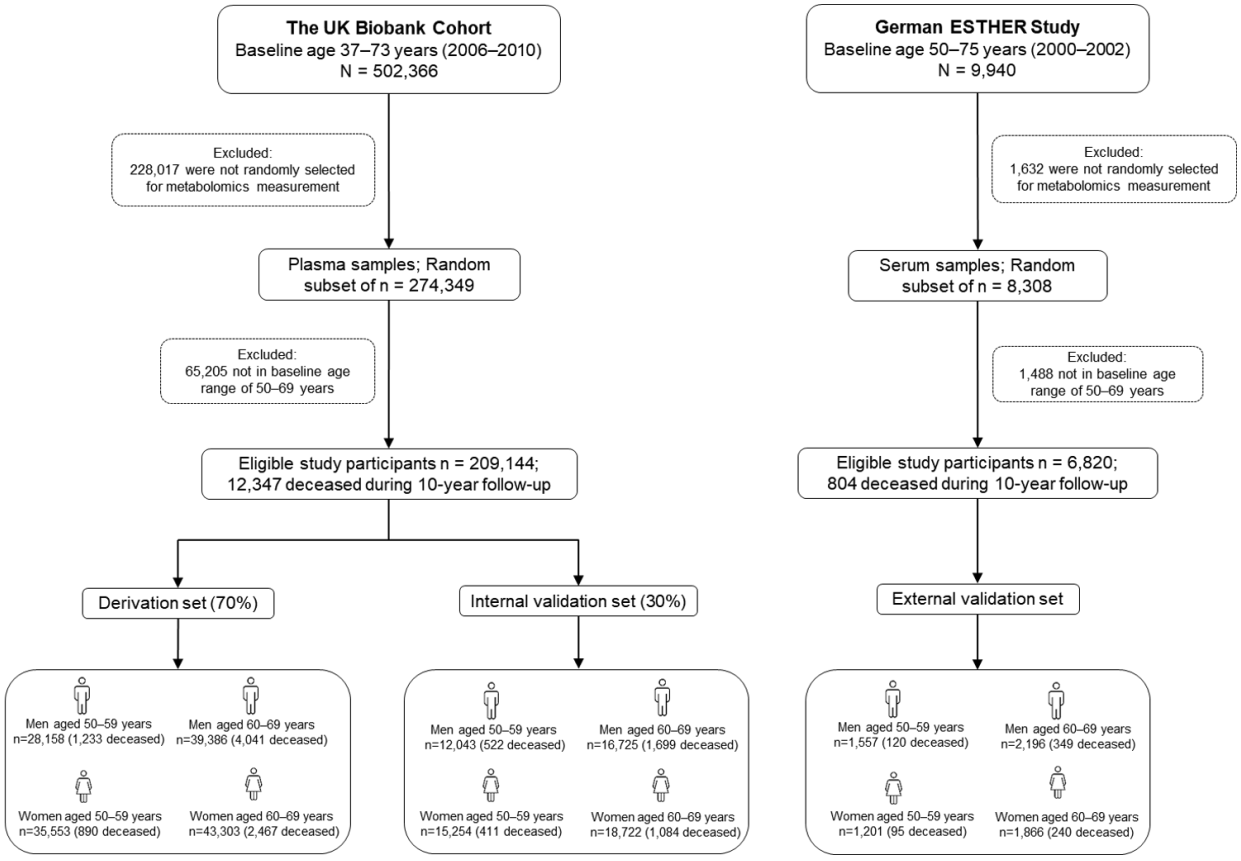



**Figure 5. Flow chart of study population from the UK Biobank and ESTHER study**

### 2.3.2 Metabolomics profiling

An NMR platform (Nightingale Health Ltd., Helsinki, Finland) (Julkunen et al. 2023) was employed to quantify circulating metabolomic biomarkers at baseline from ethylenediaminetetraacetic acid (EDTA) plasma samples collected from study participants in UKB and from serum samples in ESTHER. Nightingale Health's NMR technology ensures high reproducibility and reliability, supported by a certified quality management system in accordance with international standards such as EN ISO 13485:2016 and SFS-EN ISO/IEC 17025. This method enables the

simultaneous quantification of 250 metabolomic biomarkers, including lipids, lipoprotein subclasses, apolipoproteins, fatty acids, amino acids, ketone bodies, glycolysis related metabolites, albumin, the kidney function biomarker creatinine and the inflammation biomarker glycoprotein acetyls (GlycA). Details of the experimentation and applications of the NMR metabolomics platform have been described previously (Soininen et al. 2009; Würtz et al. 2017). Glycerol was excluded due to a proportion of missing values exceeding 10% in both cohorts. The names of the metabolomic biomarkers are indicated by representative abbreviations, and the full names and distributions are detailed in **Appendix, Table A2**.

### 2.3.3 Mortality ascertainment

In UKB, mortality was ascertained through linkage to national health registries (NHS Digital for England and Wales and the NHS Central Register for Scotland) (Sudlow et al. 2015), with follow-up through November 30, 2022. In ESTHER, mortality was determined through linkage to local population registries in Saarland, Germany (Gao et al. 2022), with follow-up through December 31, 2022. Causes of death were coded according to the International Classification of Diseases, Tenth Revision (ICD-10), with cardiovascular disease (CVD) mortality defined as codes I00–I99 and cancer mortality as codes C00–C97, excluding non-melanoma skin cancer (C44). Participants were followed from baseline until death or 10 years of follow-up.

### 2.3.4 Assessment of covariates

In both UKB and the ESTHER, baseline sociodemographic characteristics, lifestyle factors, and medical history were collected using standardized questionnaires and health examinations. Covariates included age, sex, education, smoking status, physical activity, alcohol consumption, BMI, dyslipidemia, hypertension, diabetes, cardiovascular disease, and history of cancer. Detailed definitions of all covariates are provided in Section 2.2.5.

### 2.3.5 Statistical analyses

Missing values for all variables were imputed using a random forest-based imputation method (Stekhoven and Bühlmann 2012). This iterative approach predicts missing values by constructing random forest models for each variable with missing data, leveraging observed values from other variables as predictors (Stekhoven and Bühlmann 2012). All metabolomic biomarker concentrations were log-transformed and standardized to z-scores by mean-centering and scaling to a SD of 1, allowing for direct comparisons of effect sizes across metabolomic biomarkers (Altman and Bland 2005; Bland and Altman 1996).

In the total study population of both cohorts, Cox proportional hazards models, adjusted for conventional risk factors (age, sex, education level, smoking status, physical activity, alcohol consumption, BMI, treated dyslipidemia, hypertension, diabetes, cardiovascular disease, and cancer), were used to determine hazard ratios (HRs) per 1-SD increment and 95% confidence intervals (CIs) for the associations of all 249 metabolomic biomarkers with 10-year all-cause mortality. To account for multiple testing, the Benjamini-Hochberg procedure (Benjamini and Hochberg 1995) was applied to keep the FDR at <5%. Schoenfeld residuals were utilized to examine the proportional hazards assumption (Sjölander and Dickman 2024; Stensrud and Hernán 2020). I searched for metabolite clusters among the validated biomarkers using WGCNA, which constructs a scale-free network based on hierarchical clustering (Langfelder and Horvath 2008; Zeleznik et al. 2020). Furthermore, metabolic pathway analysis based on the KEGG database (http://www.kegg.jp/ or http://www.genome.jp/kegg/) (Kanehisa et al. 2014) was used for biological interpretation of the associations of validated metabolites with mortality.

The LASSO-regularized Cox regression model, including the covariates of the traditional risk factor model, was used to identify promising metabolomic biomarkers for predicting 10-year all-cause mortality within age- (50–59/60–69 years) and sex-specific strata of the derivation set (Gafita et al. 2021; Tibshirani 1997). A bootstrap LASSO approach was employed (Wang et al. 2011), involving 1,000 bootstrap resamples. Within each resample, the penalty parameter (λ) was selected using the one-standard-error rule in 10-fold cross-validation. Metabolomic biomarkers selected by the LASSO procedure in at least 95% of the bootstrap resamples were retained for inclusion in the final model.

I used three metrics to compare and evaluate the improvement in predictive performance between models that included metabolites and those based solely on traditional risk factors: (1) Harrell's C-statistic (Antolini et al. 2004), (2) NRI (Kerr et al. 2014) and (3) IDI (Pencina et al. 2011). The Delong test was used to compare the C-statistics (Delong et al. 1988). The categorical NRI was evaluated across three risk categories (≤ 5%, >5%–10%, >10%) based on conventional risk factors and selected metabolomic biomarkers (Ersbøll et al. 2013). Model calibration was assessed by plotting observed versus predicted mortality probabilities using a nonparametric locally weighted scatterplot smoothing (loess) approach (Van Calster et al. 2019). In an additional sensitivity analysis, the model performance metrics were assessed for short-term mortality prediction (5 years).

To construct two age- and sex-specific MetaboMR clocks, I used the LASSO-selected mortality-associated metabolites from the derivation set as predictors in elastic net regression models (Horvath and Raj 2018), with chronological age as the dependent variable. The MetaboMR clock1

was based on the selected metabolites only and the MetaboMR clock2 combined them with the traditional risk factors. To improve the comparability with the chronological age, the raw metabolomic age was subsequently calibrated using the intercept ($\beta_0$) and slope ($\beta_1$) from a linear regression model with chronological age as the dependent variable and the raw metabolomic age as the independent variable. The calibrated metabolomic age was then defined as: MetaboMR clock = $\beta_0 + \beta_1 \times$ raw metabolomic age. The distribution of the MetaboMR clocks was plotted against the distribution of the chronological age in histograms to show how well the two ageing metrics align. Subsequently, a Shiny-based web calculator for the MetaboMR clocks was developed for simplified use of the algorithms in clinical routine.

The performance of these clocks was evaluated by examining their correlations with the chronological age using Pearson correlation coefficients. To evaluate the discriminatory ability of the MetaboMR clocks for mortality prediction, I also calculated their C-statistics for 10-year all-cause mortality in comparison with the C-statistic of the chronological age. Subsequently, I assessed HRs with 95% CI for the associations of each year of the metabolomic age acceleration (MetAA) with 10-year all-cause mortality. The MetAA is the residual from a linear regression model, in which the MetaboMR clock was the dependent and the chronological age was the independent variable. These residuals represent the portion of the difference between the MetaboMR clocks and chronological age that is not explained by the chronological age, or in other words, the MetAAs reflect inter-individual variation in biological ageing beyond what is expected for a given inter-individual variation of the chronological age. The distribution of the MetAAs was visualized using violin plots. Finally, associations of the MetAAs with all-cause mortality, CVD mortality, and cancer mortality were assessed with proportional hazards regression models.

In a sensitivity analysis, participants on lipid-lowering therapy at baseline were excluded and the models' discrimination (C-statistics) and associations of MetAAs with mortality were re-assessed.

The dataset was created using SAS software, version 9.4 (Cary, North Carolina, USA). Continuous variables are reported as mean (SD) and categorical variables as count (percent). All statistical tests were two-sided, with significance defined as $P<0.05$, and were performed in R, version 4.3.1. The main R packages included '*missRanger* 2.1.5' for random forest imputation (100 trees and 5 iterations), '*survival* 3.5-7' for Cox regression and model calibration, '*WGCNA* 1.72-5' for metabolite cluster derivation, '*glmnet* 4.1-8' for LASSO and elastic net regressions, '*pROC* 1.18.5' for discrimination, '*nricens* 1.6' for reclassification, '*ggplot2* 3.5.1' for visualization, and '*shiny* 1.8.1.1' for development of the practical web calculator for MetaboMR clocks.

## 2.4 Study 4: Inflammaging and proteomic ageing clocks for mortality

### 2.4.1 Study design and populations

This work used data from the UKB as the derivation cohort and the ESTHER study as the external validation cohort, which were described in Section 2.2.1 and Section 2.1.1, respectively. Details of ethics approval and consent to participate are provided in Section 2.4.6.

Participant selection is shown in **Figure 6**. In UKB, proteomic profiling was performed in a subset of 54,219 participants. A total of 34,852 individuals from the subset, aged 50–69 years, with available Olink Explore 3072 proteomic data that passed quality control and who did not use systemic steroids were included in the analyses. Participants were randomly assigned, stratified by sex, to a derivation set (70%; n=24,396) and an internal validation set (30%; n=10,456). For external validation, similar eligibility criteria were applied in ESTHER. From 8,798 participants with Olink Target 96 Inflammation panel measurements, 8,082 randomly selected participants, aged 50-74 years, were included in external validation subset 1. In addition, 100 randomly selected participants from the ESTHER study (aged 50–74 years) who died within the first 10 years of follow-up and 308 who survived the first 10 years were selected for Olink Explore HT proteomic measurements and were included as external validation subset 2 (n=408). None of the Olink Explore HT measurements were excluded for quality control reasons.

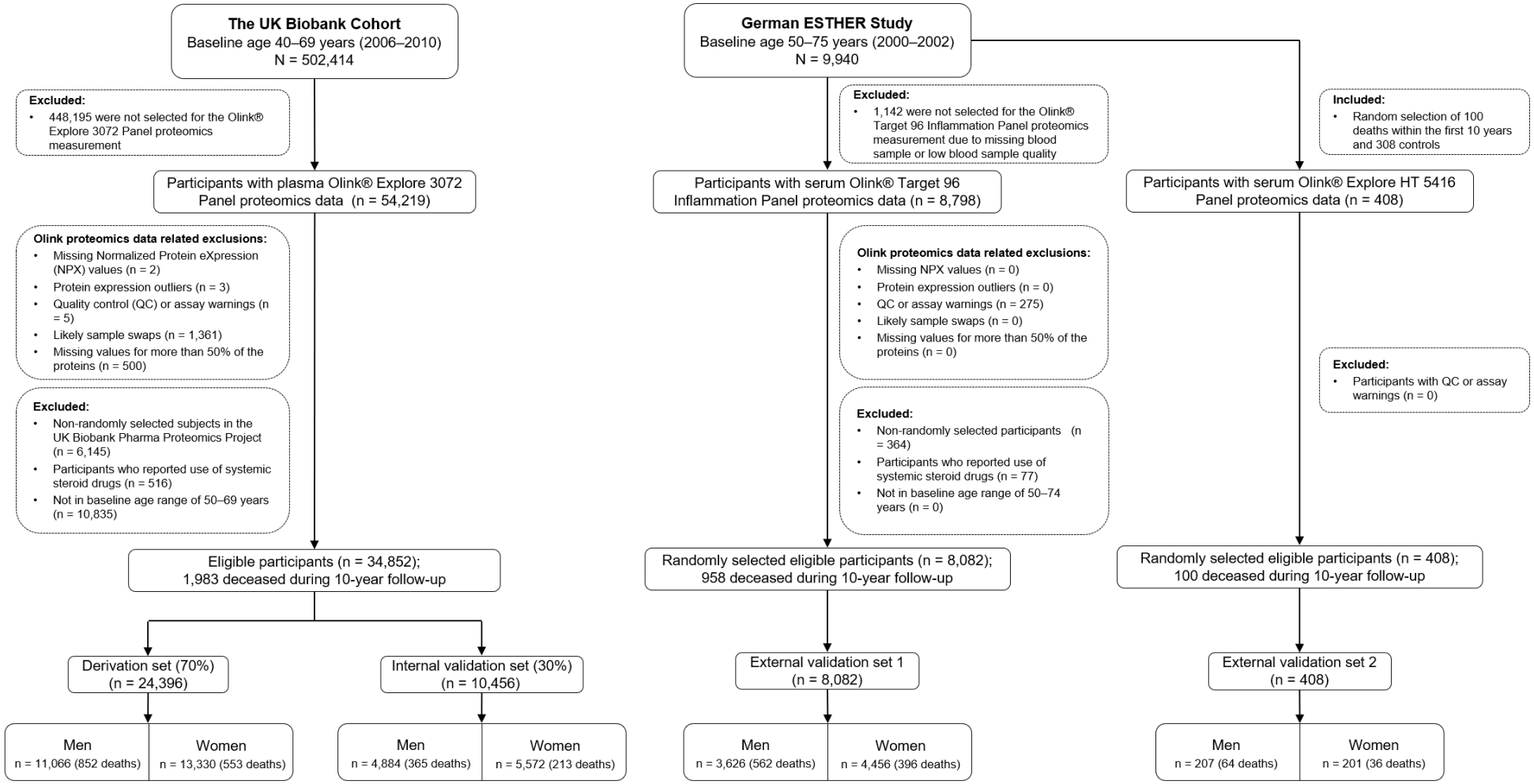



**Figure 6. Study design and analytical workflow for development and validation of inflammatory and proteomic mortality prediction models and ageing clocks**

### 2.4.2 Mortality ascertainment

The ascertainment of all-cause and cause-specific mortality was as described in Section 2.2.1.

### 2.4.3 Inflammation-related protein and proteomic profiling

In UKB, baseline EDTA plasma samples were analyzed using the Olink® Explore 3072 panel (Olink Proteomics, Uppsala, Sweden) between May 2021 and November 2022 (Eldjarn et al. 2023; Sun et al. 2023). The platform uses proximity extension assay (PEA) technology with next-generation sequencing (NGS) readout and quantifies 2,923 unique proteins (Assarsson et al. 2014; Lundberg et al. 2011). Protein abundances were reported as normalized protein expression (NPX) values on a $\log_2$ scale following standard quality control and normalization procedures. A total of 2,040 overlapping proteins passing quality control (>75% above the limit of detection and <20% missing values) in both cohorts were included in the main analyses.

In ESTHER, baseline serum samples were analyzed using the Olink® Target 96 Inflammation panel (92 proteins) and the Olink® Explore HT 5416 panel (5,401 proteins). Both assays are based on PEA technology, and protein abundances were reported as NPX values following Olink's standard quality control and normalization procedures (Trares et al. 2022; Trares et al. 2024). After quality control, 67 proteins from the Olink® Target 96 Inflammation panel and 2,040 proteins measured on the Olink® Explore HT panel that overlapped with the UKB Olink® Explore 3072 panel were retained for external validation.

### 2.4.4 Assessment of covariates

In both UKB and the ESTHER, baseline sociodemographic characteristics, lifestyle factors, and medical history were collected using standardized questionnaires and health examinations. Covariates included age, sex, education, smoking status, physical activity, alcohol consumption, BMI, dyslipidemia, hypertension, diabetes, cardiovascular disease, and history of cancer. Detailed definitions of all covariates are provided in Section 2.2.5.

### 2.4.5 Statistical analyses

Continuous variables are presented as means (standard deviations [SDs]), and categorical variables as counts and percentages. Missing values for all covariates included in the analyses were imputed separately within each cohort using a random forest-based imputation approach (Stekhoven and Bühlmann 2012). No variable had more than 9% of missing values in either of the studies. NPX

values were standardized to z-scores (mean=0 and SD=1) prior to analysis (Altman and Bland 2005; Bland and Altman 1996).

Separate analyses were performed for the inflammation-related protein panel and the proteome-wide panels. Within each panel, sex-specific protein selection was conducted in the derivation cohort using LASSO-regularized Cox regression, with conventional risk factors (age, education level, smoking status, physical activity, alcohol consumption, BMI, treated dyslipidemia, hypertension, diabetes, cardiovascular disease, and history of cancer) included as unpenalized covariates (Gafita et al. 2021; Tibshirani 1997). The optimal penalty parameter (λ) was determined by 10-fold cross-validation as the value that minimized the cross-validated partial-likelihood deviance ($\lambda_{min}$). Selection stability was assessed using 1,000 bootstrap resamples (Wang et al. 2011), and proteins selected in at least 95% of bootstrap samples were retained. Associations between selected proteins and 10-year all-cause mortality were subsequently evaluated using multivariable Cox proportional hazards models adjusted for conventional risk factors. Hazard ratios (HRs) and 95% confidence intervals (CIs) were reported per 1-SD increment. Multiple testing was controlled using the Benjamini-Hochberg procedure (FDR-adjusted *P* value <5%) (Benjamini and Hochberg 1995). Proportional hazards assumptions were assessed using Schoenfeld residuals (Grambsch and Therneau 1994; Sjölander and Dickman 2024; Stensrud and Hernán 2020).

Sex-specific protein-protein interaction networks were constructed separately for the inflammation-related protein panel and the proteome-wide panels using the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database (Burke et al. 2023). Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were subsequently performed on the union of proteins selected in men and women within each panel using the clusterProfiler package (Wu et al. 2021). Enrichment was assessed using a hypergeometric test, and pathways with an FDR-adjusted *P* value <0.05 were considered significantly enriched.

For each panel, the incremental predictive value of adding the LASSO-selected sex-specific protein biomarkers to conventional risk factor models for predicting short- and long-term all-cause mortality was evaluated using Harrell's C-statistic (Antolini et al. 2004), net reclassification improvement (NRI) (Kerr et al. 2014), and integrated discrimination improvement (IDI) (Pencina et al. 2011). C-statistics were estimated using 1,000 bootstrap resamples (Uno et al. 2011). Categorical NRI was assessed across three risk categories (≤5%, >5–10%, and >10%) (Ersbøll et al. 2013). Continuous NRI and IDI were calculated according to Pencina's method (Pencina et al. 2008; Pencina et al. 2011), with 95% confidence intervals derived from 1,000 bootstrap resamples. Calibration was assessed by comparing observed and predicted mortality risks using locally weighted

scatterplot smoothing (loess) (Van Calster et al. 2019). In an additional sensitivity analysis, the same performance metrics were additionally evaluated for 5-year all-cause mortality.

For each panel, sex-specific mortality-risk ageing clocks were developed in the derivation set using elastic net regression, with candidate proteins preselected by LASSO (Horvath and Raj 2018). Clock 1 incorporated only the LASSO-selected proteins, whereas Clock 2 additionally incorporated conventional risk factors. Raw biological ages were subsequently calibrated against chronological age using linear regression, with calibrated biological ages calculated as: Ageing Clock = $\beta_0 + \beta_1 \times$ Raw Biological Age. This approach yielded eight ageing clocks: Inflammaging Clock 1 (InflamAge1), Inflammaging Clock 2 (InflamAge2), Proteomic Ageing Clock 1 (ProtAge1), and Proteomic Ageing Clock 2 (ProtAge2), each developed separately for men and women.

The validity of the inflammaging and proteomic ageing clocks was evaluated by assessing their correlations with chronological age using Pearson correlation coefficients and their associations with 10-year all-cause mortality using Cox proportional hazards models. HRs and 95% CIs were estimated per one-year increment in inflammaging age acceleration (InflamAA) and proteomic age acceleration (ProtAA), respectively. InflamAAs and ProtAAs were defined as the residuals from linear regression models of the corresponding inflammaging/proteomic ageing clocks on chronological age.

All analyses were performed using R software (version 4.3.1; R Foundation for Statistical Computing, Vienna, Austria). Primary analyses were implemented using the following R packages: missRanger (random forest imputation), glmnet (LASSO and elastic net regression), survival (Cox regression and proportional hazards diagnostics), riskRegression (discrimination analyses), nricens (reclassification analyses), clusterProfiler (pathway enrichment analyses), and ggplot2 (data visualization and calibration plots). Protein-protein interaction networks were visualized using Cytoscape software (version 3.10.1; National Institute of General Medical Sciences, Bethesda, MD, USA). All statistical tests were two-sided, and statistical significance was defined as $P$ value <0.05 unless otherwise specified.

### 2.4.6 Ethics approval and consent to participate

The UKB was approved by the North West Multi-center Research Ethics Committee (MREC) as a Research Tissue Bank (RTB) approval (2021: 21/NW/0157) and all participants provided written informed consent. The ESTHER study was approved by the ethics committee of the Heidelberg Medical Faculty of Heidelberg University and the State Medical Board of Saarland, Germany. All participants provided written informed consent.

# 3 RESULTS

## 3.1 Study 1: Comparison of four frailty measurements

### 3.1.1 Baseline characteristics of the study population

The study included 2,000 participants with a mean (SD) age of 74.3 (6.1) years (**Table 3**). Slightly more than half of the participants (53.6%) were women. Almost two thirds of the study participants had low education ≤ 9 years (64.4%), while 18.6% had attended school for 10–11 years, and 17.0% for ≥ 12 years. The majority were never smokers (55.3%), while the others were either former (38.7%) or current smokers (6.0%). In terms of BMI, nearly half of the participants (44.3%) were classified as overweight (25–<30 kg/m²), whereas 23.9% met the criteria for obesity (≥30 kg/m²). The burden of comorbidity, as measured by the CIRS-G, showed a mean (SD) score of 7.49 (5.54).

The proportion of study participants classified as pre-frail or frail were generally comparable across the four frailty measures, with approximately 60% classified as pre-frail and 20% classified as frail. As assessed by the SPPB, 28.5% of participants had optimal, 51.6% sub-optimal, and 19.9% poor physical performance.

**Table 3. Baseline characteristics of the study population (n = 2,000)**

| Variables | Data [a] |
|---|---|
| **Age (years)** | 74.3 (6.1) |
| **Female, n (%)** | 1,072 (53.6) |
| **Education, n (%)** [b] | |
| ≤9 years | 1,289 (64.4) |
| 10–11 years | 371 (18.6) |
| ≥12 years | 340 (17.0) |
| **Smoking Status, n (%)** [b] | |
| Never smoker | 1,107 (55.3) |
| Former smoker | 773 (38.7) |
| Current smoker | 120 (6.0) |
| **BMI, n (%)** [b] | |
| < 21.5 kg/m$^2$ | 132 (6.6) |
| 21.5 – < 25 kg/m$^2$ | 503 (25.2) |
| 25 – < 30 kg/m$^2$ | 887 (44.3) |
| 30 – < 35 kg/m$^2$ | 374 (18.7) |
| ≥ 35 kg/m$^2$ | 104 (5.2) |
| **Total CIRS-G Morbidity Score** [b] | 7.49 (5.54) |
| **Fried Frailty Phenotype (FFP), n (%)** [b] | |
| Non-frail | 390 (19.5) |
| Pre-frail | 1,248 (62.4) |
| Frail | 362 (18.1) |
| **Clinical Frailty Scale (CFS), n (%)** | |
| Non-frail | 438 (21.9) |

| Variables | Data [a] |
|---|---|
| Pre-frail | 1,159 (57.9) |
| Frail | 403 (20.2) |
| **Frailty Index (FI), n (%)** [b] | |
| Non-frail | 399 (19.9) |
| Pre-frail | 1,227 (61.4) |
| Frail | 374 (18.7) |
| **FRAIL Scale (FS), n (%)** [b] | |
| Non-frail | 712 (35.6) |
| Pre-frail | 919 (45.9) |
| Frail | 369 (18.5) |
| **Short Physical Performance Battery (SPPB), n (%)** [b] | |
| Optimal | 569 (28.5) |
| Sub-optimal | 1,033 (51.6) |
| Poor | 398 (19.9) |

Abbreviations: BMI, body mass index; CIRS-G, Cumulative Illness Rating Scale-Geriatric; SD, standard deviation.
[a] Data are presented as mean (SD) or n (%).
[b] Missing values for education (1.3%), smoking status (1.9%), BMI (1.7%), the 14 somatic organ systems of the CIRS-G (0.1–0.5%), and items from the FI, FFP, FS, and SPPB (0.1-8.2%) were imputed using random forest imputation implemented with the *missRanger* package in R (version 2.6.0).

### 3.1.2 Concordance of frailty measures

The agreement between the four frailty measures in classifying individuals as pre-frail and frail ranged from fair to moderate (**Figure 7A**–**E**). Moderate agreement was noted in the total sample between the CFS and FI (kappa: 0.47), whereas the agreement of the comparisons between other frailty measures was fair (kappa values: 0.26–0.40). The agreement of the instruments did not differ substantially by sex. However, the kappa values were generally higher among the older (75–89 years) than the younger (63–74 years) subgroup.

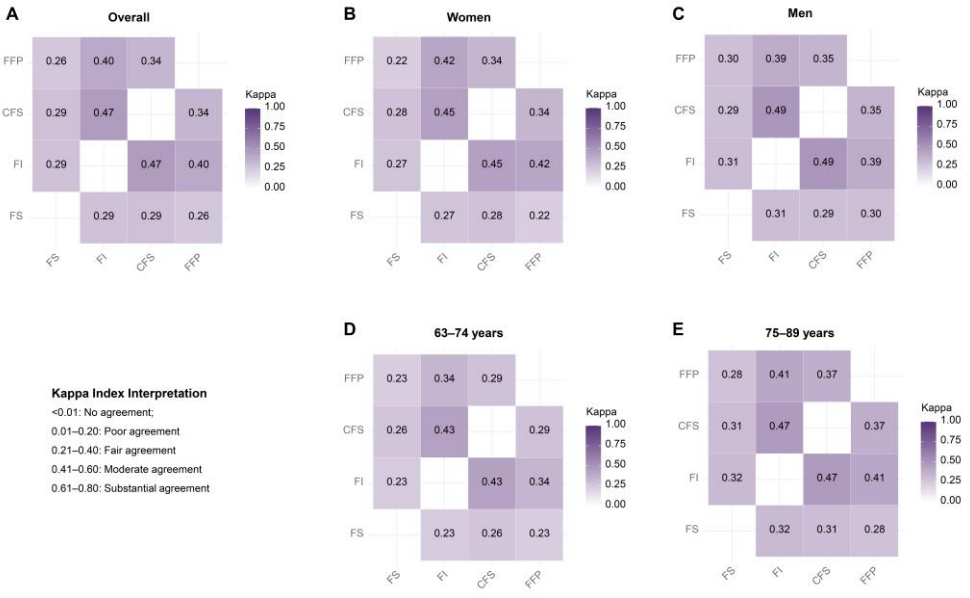



**Figure 7. Inter-instrument agreement between the four frailty instruments in the classification of non-frail, pre-frail and frail individuals**

(**A**) Overall. (**B**) Women. (**C**) Men. (**D**) 63–74 years. (**E**) 75–89 years.

### 3.1.3 Increasing frailty prevalence with age

As illustrated in **Figure 8A–D**, the proportion of frail individuals increased progressively with age in the total population for all frailty instruments ($P < 0.001$ in ordinal logistic regression model) with the exception of the FS ($P > 0.05$). The age trend was clearest for the FI and FFP. This trend was consistently observed across sex- and age-specific groups.

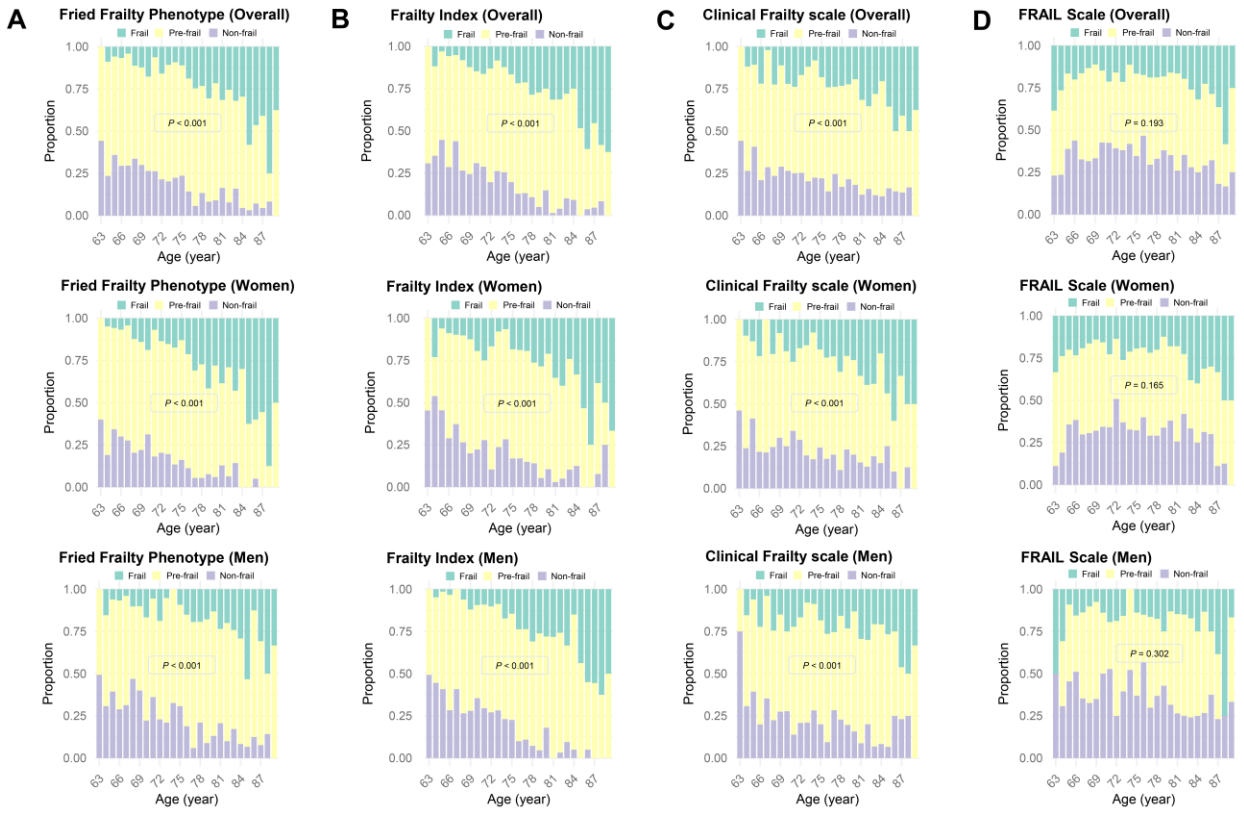



**Figure 8. Age trend of frailty prevalence according to the four frailty measures in the total population and specific subgroups**

(**A**) Fried Frailty Phenotype. (**B**) Frailty Index. (**C**) Clinical Frailty scale. (**D**) FRAIL Scale. The *P* values were derived from ordinal logistic regression models to test the association between the study participant's age (continuous) and frailty status (non-frail / pre-frail / frail).

### 3.1.4 Baseline association between frailty scales and physical performance

Every fifth study participant had poor physical performance (n = 398, 19.9%). In the total study population, frailty was significantly associated with poor physical performance across all frailty instruments, with frailty defined by the FFP showing the strongest association (OR: 9.19), followed by the other three instruments, which were each associated with approximately 4-fold higher odds of poor physical performance (**Table 4**). Pre-frail individuals had more than 4-fold odds for

poor physical performance when assessed by the FFP and approximately 2-fold odds when assessed by the other frailty instruments.

Stratified analyses revealed notable differences in the strength of associations by sex and age (**Table 4**). However, across all subgroups, the FFP consistently demonstrated the strongest association with poor physical performance, and the strength of the association was generally consistent, except for a slightly weaker association among younger participants (63–74 years). In women, frailty classified by the FI showed the second strongest association, whereas in men, the FS ranked second. Among younger individuals, frailty classified by the FS had the second strongest association, whereas in older adults, the FI showed the second strongest association.

The observed ranking of model discrimination based on AUC for poor physical performance was FFP > FS > FI > CFS in the total population, FFP > FI > FS > CFS among females, FFP > FS > FI > CFS among males, FFP > FS > FI > CFS among younger individuals, and FFP > FI > FS > CFS among older individuals (**Table 4**).

### 3.1.5 Longitudinal association of frailty status and all-cause mortality

During a median follow-up of 6.2 years, 368 of the 2,000 participants died. Cardiovascular disease and cancer were the leading causes of death, accounting for 30.4% (n = 112) and 24.2% (n = 89) of all deaths, respectively. In adjusted Cox proportional hazard regression models, frailty was significantly associated with increased all-cause mortality across all frailty measures and all populations (**Figure 9**). The Schoenfeld residual test indicated no violation of the proportional hazards assumption (all $P > 0.05$). Among all participants, frailty was associated with a 2.6- to 3.8-fold increased mortality across the different frailty measures with the FI showing the strongest association. Pre-frailty was associated with a 1.3- to 1.7-fold increased risk with again the FI showing the strongest association. Only the association of pre-frailty classified by the CFS was not statistically significant.

Frailty classified by the FI showed the highest mortality among women, younger and older individuals (**Figure 9**). Only among men did the FI rank second after the FS. The FS was also the only frailty instrument with a stronger association of frailty with mortality among males than females whereas the opposite pattern was observed for all other instruments. The frailty instruments had comparable associations with mortality in the younger and older subgroup with exception of the FFP, which showed a stronger association of frailty with mortality among younger than older individuals.

The observed ranking of model discrimination based on C-statistic for all-cause mortality was FI > CFS > FS > FFP in the total population, FI > FFP > CFS > FS among females, FS > FI > FFP > CFS among males, FI > FFP > CFS > FS among younger individuals, and FI > FS > CFS > FFP among older individuals (**Figure 9**).

### 3.1.6 Sensitivity analyses

In sensitivity analyses, additionally adjusting for BMI and CIRS-G, the results were highly consistent with the main analyses for both poor physical performance **(Table 5)** and all-cause mortality **(Figure 10**). Although all HRs and ORs mostly decreased and AUCs increased in all models, the top ranked frailty instrument always remained the same as in the main analysis.

**Table 4. Baseline associations of frailty status assessed with different instruments with poor physical performance**

| Frailty status | OR (95% CI) / AUC (95% CI) [a] | | | | |
|---|---|---|---|---|---|
| | **Total study population** | **Women** | **Men** | **63–74 years old** | **75–89 years old** |
| | **$n_{total}$ = 2,000** | **$n_{total}$ = 1,072** | **$n_{total}$ = 928** | **$n_{total}$ = 978** | **$n_{total}$ = 1,022** |
| | **$n_{poor}$ = 398 (19.9%)** | **$n_{poor}$ = 225 (21.0%)** | **$n_{poor}$ = 173 (18.6%)** | **$n_{poor}$ = 118 (12.1%)** | **$n_{poor}$ = 280 (27.4%)** |
| **Fried Frailty Phenotype (FFP)** | | | | | |
| Non-frail | Ref. | Ref. | Ref. | Ref. | Ref. |
| Pre-frail | **4.80 (2.94–7.82)** | **3.71 (1.74–7.90)** | **5.73 (3.02–10.86)** | **4.39 (2.45–7.85)** | **6.58 (2.51–17.28)** |
| Frail | **9.19 (4.87–14.30)** | **8.16 (2.12–16.42)** | **7.94 (2.99–14.07)** | **5.86 (2.26–10.06)** | **8.94 (2.91–15.20)** |
| AUC (95% CI) | **0.804 (0.780–0.828)** | **0.812 (0.779–0.845)** | **0.791 (0.755–0.827)** | **0.755 (0.709–0.800)** | **0.794 (0.763–0.825)** |
| **Clinical Frailty Scale (CFS)** | | | | | |
| Non-frail | Ref. | Ref. | Ref. | Ref. | Ref. |
| Pre-frail | **1.87 (1.18–2.96)** | **2.20 (1.17–4.15)** | 1.48 (0.75–2.90) | 1.83 (0.86–3.89) | **1.84 (1.01–3.33)** |
| Frail | **4.06 (2.32–7.11)** | **4.68 (2.14–10.22)** | **3.34 (1.50–7.47)** | **5.74 (2.39–13.76)** | **3.11 (1.49–6.45)** |
| AUC (95% CI) | **0.745 (0.717–0.773)** | **0.752 (0.716–0.788)** | **0.737 (0.694–0.780)** | **0.671 (0.620–0.721)** | **0.749 (0.716–0.788)** |
| **Frailty Index (FI)** | | | | | |
| Non-frail | Ref. | Ref. | Ref. | Ref. | Ref. |
| Pre-frail | **1.78 (1.17–2.71)** | 1.35 (0.76–2.41) | **2.30 (1.24–4.28)** | **1.90 (1.12–3.23)** | 1.42 (0.70–2.88) |
| Frail | **3.77 (2.15–6.64)** | **2.83 (1.29–6.18)** | **5.00 (2.18–11.48)** | **4.16 (1.68–10.27)** | **3.00 (1.33–6.79)** |
| AUC (95% CI) | **0.762 (0.734–0.789)** | **0.775 (0.740–0.809)** | **0.746 (0.703–0.789)** | **0.684 (0.630–0.737)** | **0.770 (0.737–0.802)** |
| **FRAIL Scale (FS)** | | | | | |
| Non-frail | Ref. | Ref. | Ref. | Ref. | Ref. |
| Pre-frail | **2.02 (1.39–2.92)** | **2.33 (1.39–3.91)** | **1.83 (1.06–3.15)** | 1.62 (0.85–3.08) | **2.29 (1.44–3.67)** |
| Frail | **3.97 (1.83–8.64)** | **6.15 (2.33–16.27)** | 1.74 (0.41–7.45) | **4.54 (1.37–14.91)** | **4.02 (1.42–11.31)** |
| AUC (95% CI) | **0.764 (0.738–0.790)** | **0.760 (0.724–0.796)** | **0.774 (0.736–0.813)** | **0.714 (0.667–0.760)** | **0.758 (0.725–0.791)** |

Abbreviations: AUC, area under the curve; CI, confidence interval; OR, odds ratio. Printed in bold: Statistically significant ($P < 0.05$).

[a] Poor physical performance was defined by ≤ 6 points for the short physical performance battery (SPPB). All logistic regression models included the covariates age (years), sex, education and smoking status.

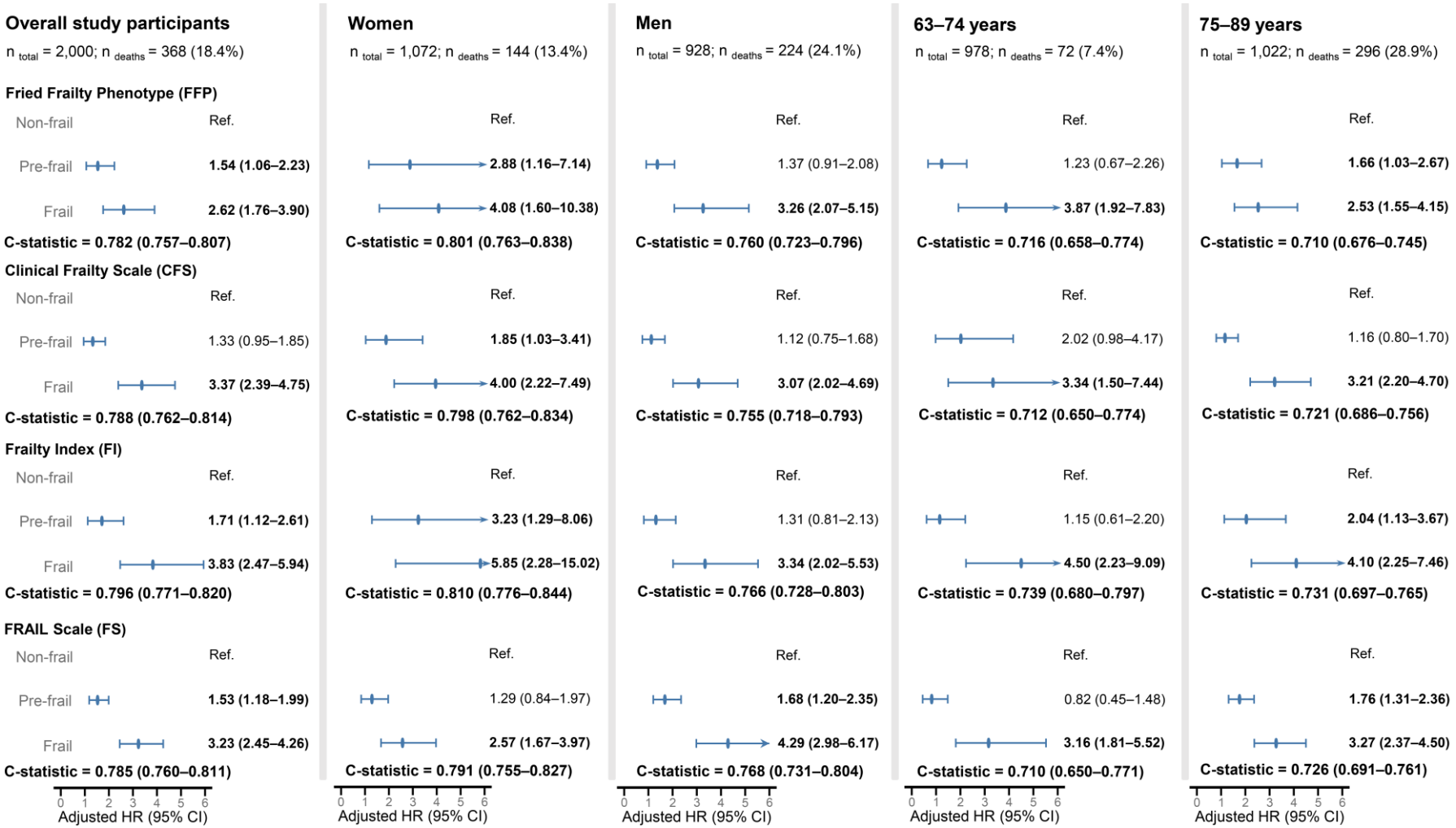



**Figure 9. Longitudinal association of the four frailty measures with all-cause mortality in the overall population and specific subgroups**

All Cox regression models were adjusted for age (years), sex, education, and smoking status. Statistically significant results ($P < 0.05$) are highlighted in bold. Abbreviations: CI, confidence interval; HR, hazard ratio.

**Table 5. Associations of frailty status assessed with different instruments with poor physical performance with extended model adjustment**

| Frailty status | OR (95% CI) / AUC (95% CI) [a] | | | | |
|---|---|---|---|---|---|
| | **Total study population** | **Women** | **Men** | **63–74 years old** | **75–89 years old** |
| | **$n_{total}$ = 2,000** | **$n_{total}$ = 1,072** | **$n_{total}$ = 928** | **$n_{total}$ = 978** | **$n_{total}$ = 1,022** |
| | **$n_{poor}$ = 398 (19.9%)** | **$n_{poor}$ = 225 (21.0%)** | **$n_{poor}$ = 173 (18.6%)** | **$n_{poor}$ = 118 (12.1%)** | **$n_{poor}$ = 280 (27.4%)** |
| **Fried Frailty Phenotype (FFP)** | | | | | |
| Non-frail | Ref. | Ref. | Ref. | Ref. | Ref. |
| Pre-frail | **4.25 (2.60–7.08)** | **3.63 (1.32–6.80)** | **5.60 (2.44–9.64)** | **4.21 (2.57–6.87)** | **5.97 (2.43–15.11)** |
| Frail | **8.63 (4.22–12.47)** | **7.67 (2.03–14.19)** | **7.22 (2.15–13.58)** | **5.43 (2.04–9.43)** | **8.01 (2.25–12.44)** |
| AUC (95% CI) | **0.818 (0.795–0.842)** | **0.826 (0.795–0.857)** | **0.817 (0.783–0.851)** | **0.784 (0.743–0.825)** | **0.815 (0.787–0.844)** |
| **Clinical Frailty Scale (CFS)** | | | | | |
| Non-frail | Ref. | Ref. | Ref. | Ref. | Ref. |
| Pre-frail | **1.68 (1.11–2.37)** | **1.99 (1.08–3.91)** | 1.33 (0.72–2.60) | 1.53 (0.94–2.51) | 1.79 (0.96–2.55) |
| Frail | **3.69 (2.11–5.89)** | **4.48 (2.07–9.35)** | **3.28 (1.39–7.01)** | **5.13 (2.09–12.34)** | **2.72 (1.34–5.64)** |
| AUC (95% CI) | **0.754 (0.727–0.781)** | **0.765 (0.730–0.800)** | **0.753 (0.711–0.796)** | **0.683 (0.634–0.732)** | **0.761 (0.728–0.793)** |
| **Frailty Index (FI)** | | | | | |
| Non-frail | Ref. | Ref. | Ref. | Ref. | Ref. |
| Pre-frail | **1.75 (1.12–2.57)** | 1.18 (0.72–1.89) | **2.07 (1.13–4.02)** | **1.86 (1.10–3.06)** | 1.30 (0.64–2.57) |
| Frail | **3.53 (2.06–5.63)** | **2.45 (1.13–5.45)** | **4.92 (2.03–9.55)** | **4.07 (1.18–9.48)** | **2.87 (1.21–5.12)** |
| AUC (95% CI) | **0.767 (0.740–0.793)** | **0.787 (0.752–0.821)** | **0.761 (0.720–0.802)** | **0.698 (0.647–0.750)** | **0.784 (0.753–0.815)** |
| **FRAIL Scale (FS)** | | | | | |
| Non-frail | Ref. | Ref. | Ref. | Ref. | Ref. |
| Pre-frail | **1.98 (1.29–2.47)** | **2.03 (1.11–3.53)** | **1.71 (1.22–2.44)** | 1.34 (0.85–2.16) | **2.09 (1.10–3.20)** |
| Frail | **3.84 (1.77–7.68)** | **5.92 (2.14–14.66)** | 1.58 (0.56–6.67) | **4.32 (1.29–13.62)** | **3.96 (1.25–9.40)** |
| AUC (95% CI) | **0.786 (0.761–0.811)** | **0.781 (0.747–0.816)** | **0.799 (0.763–0.834)** | **0.738 (0.693–0.782)** | **0.775 (0.743–0.806)** |

Abbreviations: AUC, area under the curve; BMI, body mass index; CI, confidence interval; OR, odds ratio. Printed in bold: Statistically significant ($P < 0.05$).

[a] Poor physical performance was defined by ≤ 6 points for the short physical performance battery (SPPB). All logistic regression models included the covariates age (years), sex, education, smoking status, continuous CIRS-G score, and categorical BMI.

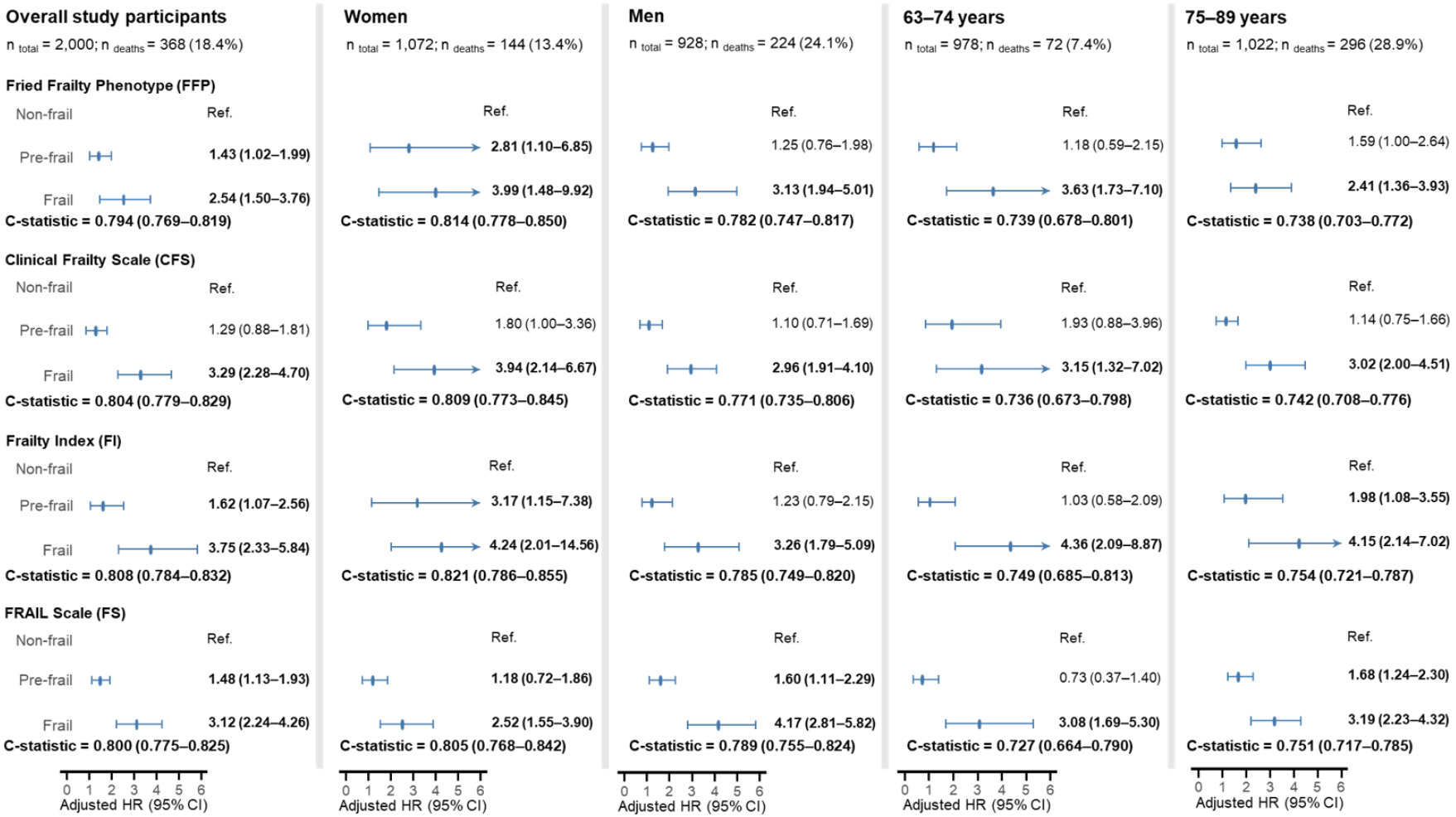



**Figure 10. Longitudinal association of the four frailty measures with all-cause mortality in the overall population and specific subgroups with extended model adjustment**

All Cox regression models were adjusted for age, sex, education, smoking status, continuous CIRS-G score, and categorical BMI. Statistically significant results ($P < 0.05$) are highlighted in bold. Abbreviations: BMI, body mass index; CI, confidence interval; HR, hazard ratio.

## 3.2 Study 2: Integrated proteomic and metabolomic profiling for frailty

### 3.2.1 Baseline characteristics of study populations

Participant selection in UKB and ESTHER is shown in **Figure 4**. The baseline characteristics of the included participants in both cohorts were broadly comparable (**Table 6**). In both cohorts, frail participants were older, were less often women, and had lower educational attainment than non-frail individuals. They also had less favorable lifestyle profiles, including lower proportions of never smokers, lower physical activity, higher proportions of alcohol abstainers, and a higher prevalence of obesity (BMI ≥30 kg/m²). Medical history, including dyslipidemia, hypertension, diabetes, cardiovascular disease, and cancer, was more prevalent among frail participants in both cohorts.

**Table 6. Baseline characteristics of study participants in the UK Biobank and ESTHER study**

| Characteristics | UK Biobank (n=19,151) | | | ESTHER study (n=5,031) | | |
|---|---|---|---|---|---|---|
| | Non-Frailty (n=17,288) | Frailty (n=1,863) | FDR [i] | Non-Frailty (n=4,537) | Frailty (n=494) | FDR [i] |
| **Age, years** | 64.0 (2.8) | 65.1 (2.8) | <0.001 | 64.7 (4.0) | 66.6 (4.2) | <0.001 |
| **Sex, n (%)** | / | / | 0.001 | / | / | 0.001 |
| Women | 9,227 (53.4) | 920 (49.4) | | 2,500 (55.1) | 234 (47.4) | |
| Men | 8,061 (46.6) | 943 (50.6) | | 2,037 (44.9) | 260 (52.6) | |
| **Education, n (%)** [a, h] | / | / | 0.015 | / | / | <0.001 |
| Low | 12,443 (72.0) | 1,391 (74.7) | | 3,515 (77.5) | 429 (86.8) | |
| High | 4,845 (28.0) | 472 (25.3) | | 1,022 (22.5) | 65 (13.2) | |
| **Smoking status, n (%)** [h] | / | / | <0.001 | / | / | 0.006 |
| Never smoker | 8,940 (51.7) | 716 (38.4) | | 2,545 (56.1) | 246 (49.8) | |
| Former smoker | 6,995 (40.5) | 956 (51.3) | | 1,447 (31.9) | 193 (39.1) | |
| Current smoker | 1,353 (7.8) | 191 (10.3) | | 545 (12.0) | 55 (11.1) | |
| **Physical activity, n (%)** [h] | / | / | <0.001 | / | / | <0.001 |
| Low | 2,062 (12.0) | 328 (17.6) | | 1,010 (22.3) | 175 (35.4) | |
| Moderate | 8,946 (51.7) | 1,017 (54.6) | | 2,110 (46.5) | 268 (54.3) | |
| High | 6,280 (36.3) | 518 (27.8) | | 1,417 (31.2) | 51 (10.3) | |
| **Alcohol consumption, n (%)** [b, h] | / | / | <0.001 | / | / | <0.001 |
| Abstainer | 5,140 (29.7) | 735 (39.5) | | 1,354 (29.8) | 225 (45.5) | |
| Low | 7,175 (41.5) | 675 (36.2) | | 2,912 (64.2) | 254 (51.5) | |
| Medium or high | 4,973 (28.8) | 453 (24.3) | | 271 (6.0) | 15 (3.0) | |
| **BMI, n (%)** [c, g] | / | / | <0.001 | / | / | <0.001 |
| < 21.5 kg/m² | 1,059 (6.1) | 34 (1.8) | | 179 (3.9) | 16 (3.2) | |
| ≥ 21.5 – < 25 kg/m² | 4,407 (25.5) | 193 (10.4) | | 1,019 (22.5) | 63 (12.8) | |
| ≥ 25 – < 30 kg/m² | 8,010 (46.3) | 684 (36.7) | | 2,262 (49.8) | 195 (39.5) | |
| ≥ 30 – < 35 kg/m² | 2,966 (17.2) | 583 (31.3) | | 874 (19.3) | 155 (31.4) | |
| ≥ 35 kg/m² | 846 (4.9) | 369 (19.8) | | 203 (4.5) | 65 (13.1) | |
| **Dyslipidemia, n (%)** [d, h] | / | / | <0.001 | / | / | <0.001 |
| No | 13,115 (75.9) | 760 (40.8) | | 4,043 (89.1) | 311 (63.0) | |
| Yes | 4,173 (24.1) | 1,103 (59.2) | | 494 (10.9) | 183 (37.0) | |
| **Hypertension, n (%)** [e, h] | / | / | <0.001 | / | / | <0.001 |
| No | 6,160 (35.6) | 194 (10.4) | | 1,454 (32.0) | 32 (6.5) | |
| Yes | 11,128 (64.4) | 1,669 (89.6) | | 3,083 (68.0) | 462 (93.5) | |
| **Diabetes, n (%)** [f, h] | / | / | <0.001 | / | / | <0.001 |
| No | 15,710 (90.9) | 1,204 (64.6) | | 3,902 (86.0) | 278 (56.3) | |
| Yes | 1,578 (9.1) | 659 (35.4) | | 635 (14.0) | 216 (43.7) | |

| Characteristics | UK Biobank (n=19,151) | | | ESTHER study (n=5,031) | | |
|---|---|---|---|---|---|---|
| | Non-Frailty (n=17,288) | Frailty (n=1,863) | FDR [i] | Non-Frailty (n=4,537) | Frailty (n=494) | FDR [i] |
| **Cardiovascular disease, n (%)** [g, h] | / | / | <0.001 | / | / | <0.001 |
| No | 16,002 (92.6) | 1,381 (74.1) | | 4,266 (94.0) | 273 (55.3) | |
| Yes | 1,286 (7.4) | 482 (25.9) | | 271 (6.0) | 221 (44.7) | |
| **History of cancer, n (%)** [h] | / | / | <0.001 | / | / | <0.001 |
| No | 15,628 (90.4) | 1,526 (81.9) | | 4,149 (91.4) | 422 (85.4) | |
| Yes | 1,660 (9.6) | 337 (18.1) | | 388 (8.6) | 72 (14.6) | |

Abbreviations: BMI, body mass index; FDR, false discovery rate; SD, standard deviation.

Notes: Values are means (SDs) for continuous variables and percentages for categorical variables.

[a] In the UK Biobank, participants who reported "None of the above," "Certificate of Secondary Education (CSEs) or equivalent," "Ordinary Levels/General Certificate of Secondary Education (O levels/GCSEs) or equivalent," "Other professional qualifications (e.g., nursing, teaching)," or "Advanced Levels/Advanced Subsidiary Levels (A levels/AS levels) or equivalent" were categorized as having a low or medium education level. Those who reported "National Vocational Qualification (NVQ), Higher National Diploma (HND), Higher National Certificate (HNC) or equivalent," or a "College or University degree" were categorized as having a high education level. In the German ESTHER study, participants who reported "No education or 'Hauptschule'" or "'Realschule'/'Mittlere Reife'" were categorized as having a low education level, whereas those who reported "'Fachhochschulreife' or 'Abitur'" were categorized as having a high education level.

[b] Alcohol consumption was categorized as low for women consuming 0–19.99 g/day and men consuming 0–39.99 g/day. It was categorized as medium or high for women consuming 20–39.99 g/day or ≥40 g/day, and for men consuming 40–59.99 g/day or ≥60 g/day.

[c] Low weight was defined as a body mass index (BMI) <21.5 kg/m² because of the small number of participants with BMI <18.5 kg/m².

[d] Dyslipidemia was defined as use of lipid-lowering medication.

[e] Hypertension was defined as self-reported hypertension, the use of antihypertensive medication, a systolic blood pressure ≥140 mm Hg, or a diastolic blood pressure ≥90 mm Hg.

[f] Diabetes was defined as self-reported diabetes of any type, the use of glucose-lowering medication, or an $HbA_{1c}$ level of ≥6.5%.

[g] Cardiovascular disease was defined as a history of myocardial infarction or stroke.

[h] The proportions of missing values in the UK Biobank were as follows: education (1.6%), smoking status (0.1%), physical activity (8.4%), alcohol consumption (2.3%), BMI (1.4%), dyslipidemia (0.03%), hypertension (0.03%), diabetes (0.01%), cardiovascular disease (0.5%), and cancer (0.4%), and in the ESTHER study: education (3.2%), smoking status (3.5%), physical activity (0.5%), alcohol consumption (7.7%), BMI (1.6%), dyslipidemia (1.7%), hypertension (2.2%), diabetes (1.5%), cardiovascular disease (3.6%) and cancer (3.0%). Missing values were imputed using random forest imputation with the R package *missRanger* (version 2.1.5).

[i] Comparisons were made using independent-sample *t* tests for age and chi-square tests for categorical variables. *P* values were adjusted for multiple comparisons using the Benjamini-Hochberg FDR method.

### 3.2.2 Identification of proteomic and metabolomic features associated with prevalent and incident frailty

LASSO feature selection identified 17 proteomic and 9 metabolomic biomarkers associated with prevalent frailty in UKB after adjustment for conventional risk factors. Of the 26 selected features, 20 were significantly associated with prevalent frailty in both cohorts. The remaining biomarkers showed consistent directions of effect in ESTHER (**Figure 11A**). Furthermore, 9 biomarkers (6 proteins and 3 metabolites) were also significantly associated with incident frailty in both cohorts (**Figure 11B**). Several biomarkers, including *HGF*, *TGFA*, *FGF23*, *CDCP1*, *IL6*, *CCL20*, glucose, and albumin, showed consistent directions of association with both prevalent and incident frailty. Detailed effect estimates, 95% CIs and FDR-adjusted *P* values for prevalent frailty and incident frailty are shown in **Appendix, Table A3** and **A4**.

Pairwise correlations among the selected biomarkers are shown in **Figure 12**. No strong correlations were observed. A small number of moderate correlations ($0.3 < r < 0.6$) were observed among proteins, among metabolites, and between proteins and metabolites.

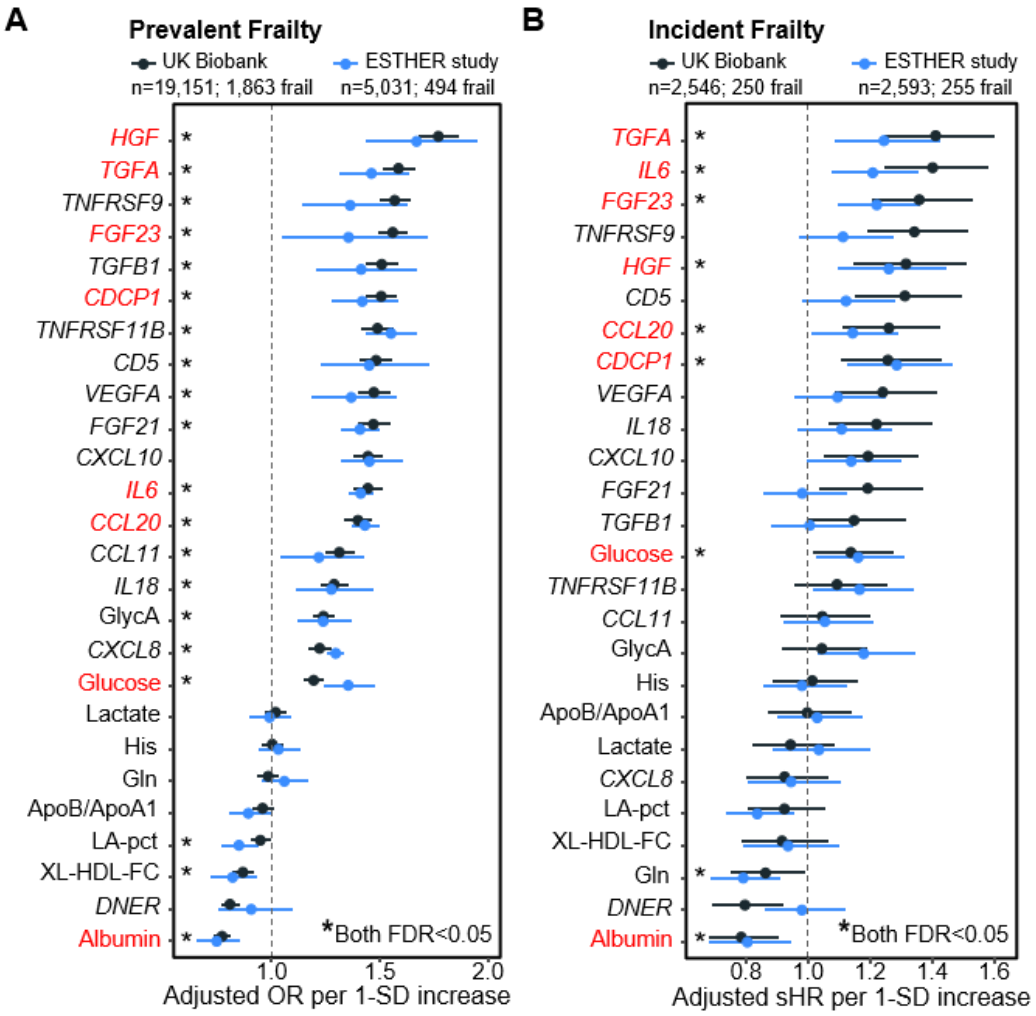



**Figure 11. Associations of LASSO-selected proteomic and metabolomic biomarkers with prevalent frailty and incident frailty during 8-year follow-up in the UK Biobank and ESTHER study**

**(A)** Multivariable-adjusted odds ratios (ORs) and 95% confidence intervals (CIs) for associations between LASSO-selected omics biomarkers and prevalent frailty at baseline. **(B)** Multivariable-adjusted subdistribution hazard ratios (sHRs) and 95% CIs for associations between LASSO-selected omics biomarkers and incident frailty during 8 years of follow-up. Models were adjusted for chronological age, sex, education, smoking status, physical activity, alcohol consumption, body mass index (BMI), dyslipidemia, hypertension, diabetes, cardiovascular disease, and history of cancer. All ORs and sHRs are shown per 1-standard-deviation (SD) increase of the biomarker.

Black points and horizontal lines indicate effect estimates and 95% CIs in the UK Biobank; blue points and horizontal lines indicate the corresponding estimates in the ESTHER study. Asterisks indicate biomarkers with false discovery rate (FDR)-adjusted *P* values <0.05 in both cohorts. Biomarker names shown in red indicate biomarkers that were statistically significantly associated with both prevalent and incident frailty. Corresponding numerical estimates and FDR-adjusted *P* values are provided for prevalent frailty in **Appendix, Table A3** and for incident frailty in **Appendix, Table A4**.

Abbreviations: BMI, body mass index; OR, odds ratio; sHR, subdistribution hazard ratio; CI, confidence interval; SD, standard deviation; FDR, false discovery rate; LASSO, least absolute shrinkage and selection operator; the abbreviations for proteins and metabolites are listed in **Appendix, Table A5**.

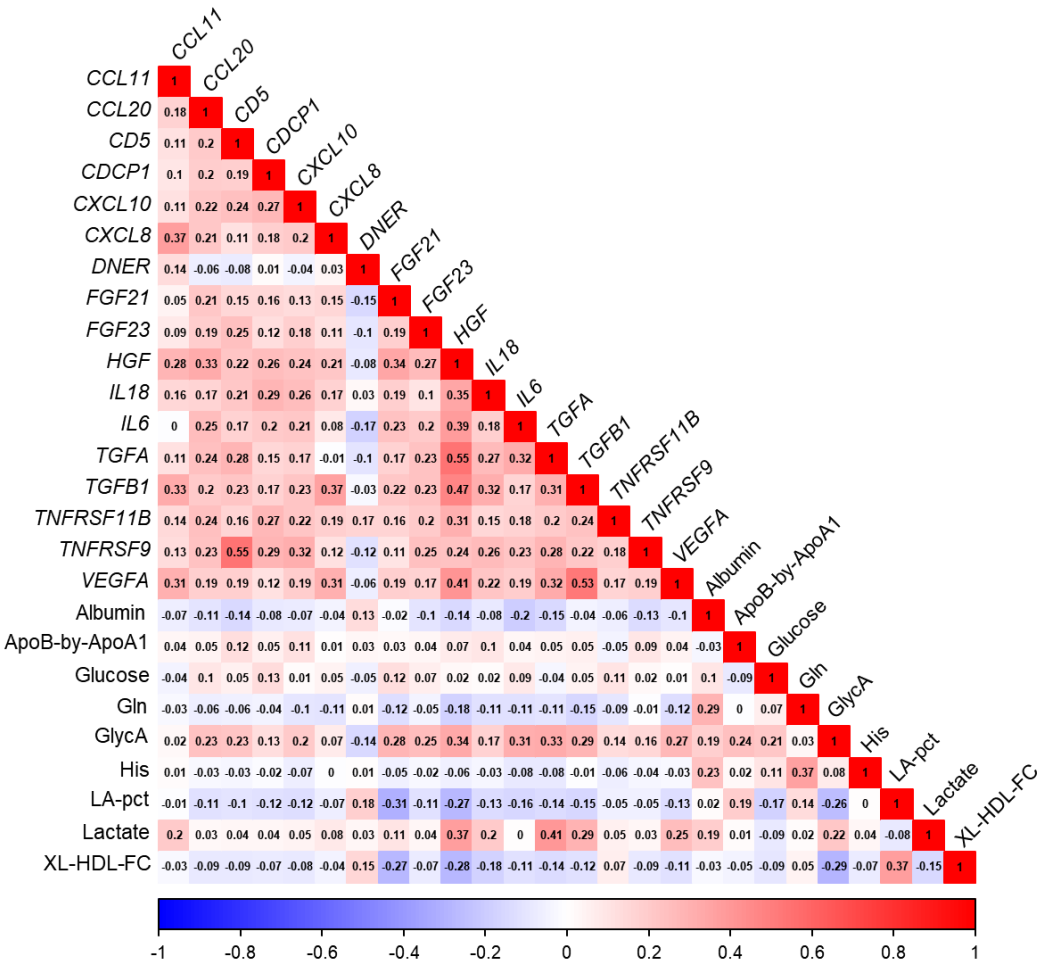



**Figure 12. Heatmap of Pearson correlations between 17 proteomic and 9 metabolomic biomarkers in the UK Biobank**

Pearson correlation coefficients were calculated to assess pairwise associations between the selected proteomic (n=17) and metabolomic (n=9) biomarkers in the UK Biobank, based on $\log_2$-transformed and z-standardized values. Correlation strength and direction are represented by the color scale.

### 3.2.3 TRIM models for prevalent and incident frailty

I compared reference models including conventional risk factors with the TRIM models, which additionally included the selected 17 proteins and 9 metabolites for prevalent frailty and incident frailty. The $\beta$ coefficients of the final models derived from UKB are provided in **Appendix, Table A6**.

For prevalent frailty, the reference model achieved AUCs of 0.818 (95% CI, 0.809–0.828) and 0.801 (95% CI, 0.781–0.822) in UKB and ESTHER, respectively (**Table 7**). Both proteomics and metabolomics alone improved the reference model's AUCs, but the TRIM-FI achieved the highest

AUCs: 0.864 (Δ0.046, $P$<0.001) in UKB and 0.834 (Δ0.033, $P$<0.001) in ESTHER. Reclassification was substantial by the TRIM-FI, with a continuous NRI of 33.3%, a categorical NRI of 11.6%, and an IDI of 4.0% (all $P$<0.001) in the external validation set (ESTHER) (**Table 7**).

For incident frailty, the reference model yielded C-statistics of 0.739 (95% CI, 0.707–0.771) and 0.699 (95% CI, 0.667–0.731) in UKB and ESTHER, respectively (**Table 7**). Addition of proteomics and metabolomics biomarkers (TRIM-FP) increased the C-statistic to 0.788 (Δ0.049, $P$<0.001) and 0.746 (Δ0.047, $P$<0.001) in UKB and ESTHER, respectively, which was more than adding only a single omics biomarker group. The continuous NRI (25.7%), categorical NRI (13.0%), and IDI (3.0%) were all statistically significant (all $P$<0.001) in the external validation set.

The Receiver Operating Characteristic (ROC) curves of the reference models and the TRIM models for prevalent and incident frailty are presented in **Figure 13**. The TRIM models demonstrated good calibration in the derivation and external validation cohorts (**Figure 14**).

**Table 7. Predictive value of proteomic and metabolomic biomarkers for the assessment of baseline frailty and prediction of incident frailty among older adults in the UK Biobank and ESTHER study**

| Outcome / Cohort / Model | AUC (95% CI) or C-statistic (95% CI) [f] | Δ AUC or Δ C-statistic [f] | Δ Continuous NRI (95% CI) | Δ Categorical NRI (95% CI) [g] | Δ IDI (95% CI) |
|---|---|---|---|---|---|
| **Baseline frailty** | | | | | |
| UK Biobank (derivation set) n=19,151; 1,863 (9.7%) frail | | | | | |
| Reference model [a] | 0.818 (0.809–0.828) | Reference | — | — | — |
| + Proteomics [b] | 0.846 (0.838–0.855) | +0.028, *P*<0.001 | 29.3% (24.5%–34.1%), *P*<0.001 | 15.3% (13.2%–17.3%), *P*<0.001 | 5.4% (2.8%–8.0%), *P*<0.001 |
| + Metabolomics [c] | 0.839 (0.830–0.848) | +0.021, *P*<0.001 | 22.5% (17.6%–27.3%), *P*<0.001 | 10.7% (8.2%–13.2%), *P*<0.001 | 3.9% (2.6%–5.1%), *P*<0.001 |
| TRIM-FI [d] | 0.864 (0.856–0.872) | +0.046, *P*<0.001 | 44.6% (40.1%–49.1%), *P*<0.001 | 20.4% (18.3%–22.5%), *P*<0.001 | 7.3% (6.6%–8.0%), *P*<0.001 |
| ESTHER study (validation set) n=5,031; 494 (9.8%) frail | | | | | |
| Reference model [a] | 0.801 (0.781–0.822) | Reference | — | — | — |
| + Proteomics [b] | 0.824 (0.805–0.843) | +0.023, *P*=0.007 | 20.7% (11.5%–29.9%), *P*<0.001 | 9.4% (5.8%–13.0%), *P*<0.001 | 2.8% (1.9%–3.8%), *P*<0.001 |
| + Metabolomics [c] | 0.819 (0.799–0.839) | +0.018, *P*=0.011 | 16.6% (7.2%–25.9%), *P*<0.001 | 6.8% (3.0%–10.5%), *P*<0.001 | 1.8% (1.0%–2.6%), *P*=0.004 |
| TRIM-FI [d] | 0.834 (0.815–0.852) | +0.033, *P*<0.001 | 33.3% (19.2%–47.4%), *P*<0.001 | 11.6% (7.5%–15.7%), *P*<0.001 | 4.0% (2.8%–5.3%), *P*<0.001 |
| **Incident frailty [e]** | | | | | |
| UK Biobank (derivation set) n=2,546; 250 (9.8%) frail | | | | | |
| Reference model [a] | 0.739 (0.707–0.771) | Reference | — | — | — |
| + Proteomics [b] | 0.775 (0.745–0.805) | +0.036, *P*<0.001 | 20.3% (6.8%–33.7%), *P*=0.003 | 9.2% (4.7%–13.7%), *P*<0.001 | 3.2% (1.9%–4.5%), *P*<0.001 |
| + Metabolomics [c] | 0.770 (0.739–0.801) | +0.031, *P*<0.001 | 18.1% (8.2%–27.9%), *P*<0.001 | 8.4% (6.7%–10.1%), *P*<0.001 | 2.7% (2.1%–3.3%), *P*<0.001 |
| TRIM-FP [d] | 0.788 (0.758–0.817) | +0.049, *P*<0.001 | 28.8% (19.7%–37.9%), *P*<0.001 | 16.9% (10.1%–23.6%), *P*<0.001 | 4.2% (3.1%–5.3%), *P*<0.001 |
| ESTHER study (validation set) n=2,593; 255 (9.8%) frail | | | | | |
| Reference model [a] | 0.699 (0.667–0.731) | Reference | — | — | — |
| + Proteomics [b] | 0.728 (0.697–0.759) | +0.029, *P*=0.006 | 17.7% (4.5%–30.9%), *P*=0.009 | 7.3% (1.9%–12.7%), *P*=0.008 | 2.2% (1.3%–3.1%), *P*<0.001 |
| + Metabolomics [c] | 0.721 (0.690–0.753) | +0.022, *P*=0.001 | 15.9% (3.3%–28.5%), *P*=0.013 | 7.0% (1.2%–12.8%), *P*=0.021 | 2.0% (1.0%–3.0%), *P*=0.001 |
| TRIM-FP [d] | 0.746 (0.715–0.777) | +0.047, *P*<0.001 | 25.7% (12.1%–39.3%), *P*<0.001 | 13.0% (4.6%–21.4%), *P*=0.002 | 3.0% (1.8%–4.2%), *P*<0.001 |

Abbreviations: AUC, area under the receiver operating characteristic curve; BMI, body mass index; C-statistic, concordance statistic; CI, confidence interval; IDI, integrated discrimination improvement; NRI, net reclassification improvement; TRIM-FI, Traditional Risk factor, Inflammation, and Metabolism based Frailty Identification model; TRIM-FP, Traditional Risk factor, Inflammation, and Metabolism based Frailty Prediction model.

Notes: All *P* values for model improvements were <0.05. Differences in AUC were evaluated using DeLong's method, differences in time-dependent C-statistics were assessed using bootstrap resampling, and NRI and IDI were tested using likelihood ratio-based inference.

[a] The reference model included the following conventional risk covariates: age, sex, education, smoking status, physical activity, alcohol consumption, BMI, dyslipidemia, hypertension, diabetes, cardiovascular disease, and history of cancer.

[b] The proteomics model included the variables of the reference model plus 17 proteins: *TNFRSF11B*, *DNER*, *FGF23*, *FGF21*, *HGF*, *VEGFA*, *CXCL8*, *CXCL10*, *CCL11*, *CCL20*, *CDCP1*, *CD5*, *TGFA*, *TGFB1*, *TNFRSF9*, *IL18*, and *IL6* (see **Appendix, Table A5** for abbreviations).

[c] The metabolomics model included the variables of the reference model plus 9 metabolites: His, Glucose, Lactate, Albumin, GlycA, Gln, LA-pct, ApoB-by-ApoA1, and XL-HDL-FC (see **Appendix, Table A5** for abbreviations).

[d] The TRIM model included the variables of the reference model plus the aforementioned 17 proteins and 9 metabolites.

[e] Incident frailty analysis during a follow-up period of 8 years (median of 7.8 years in the UK Biobank and 7.7 years in the ESTHER study).

[f] AUC for baseline frailty or Harrell's C-statistic for incident frailty.

[g] Categorical NRI was calculated based on predefined risk thresholds of <5%, 5%–10%, and >10%.

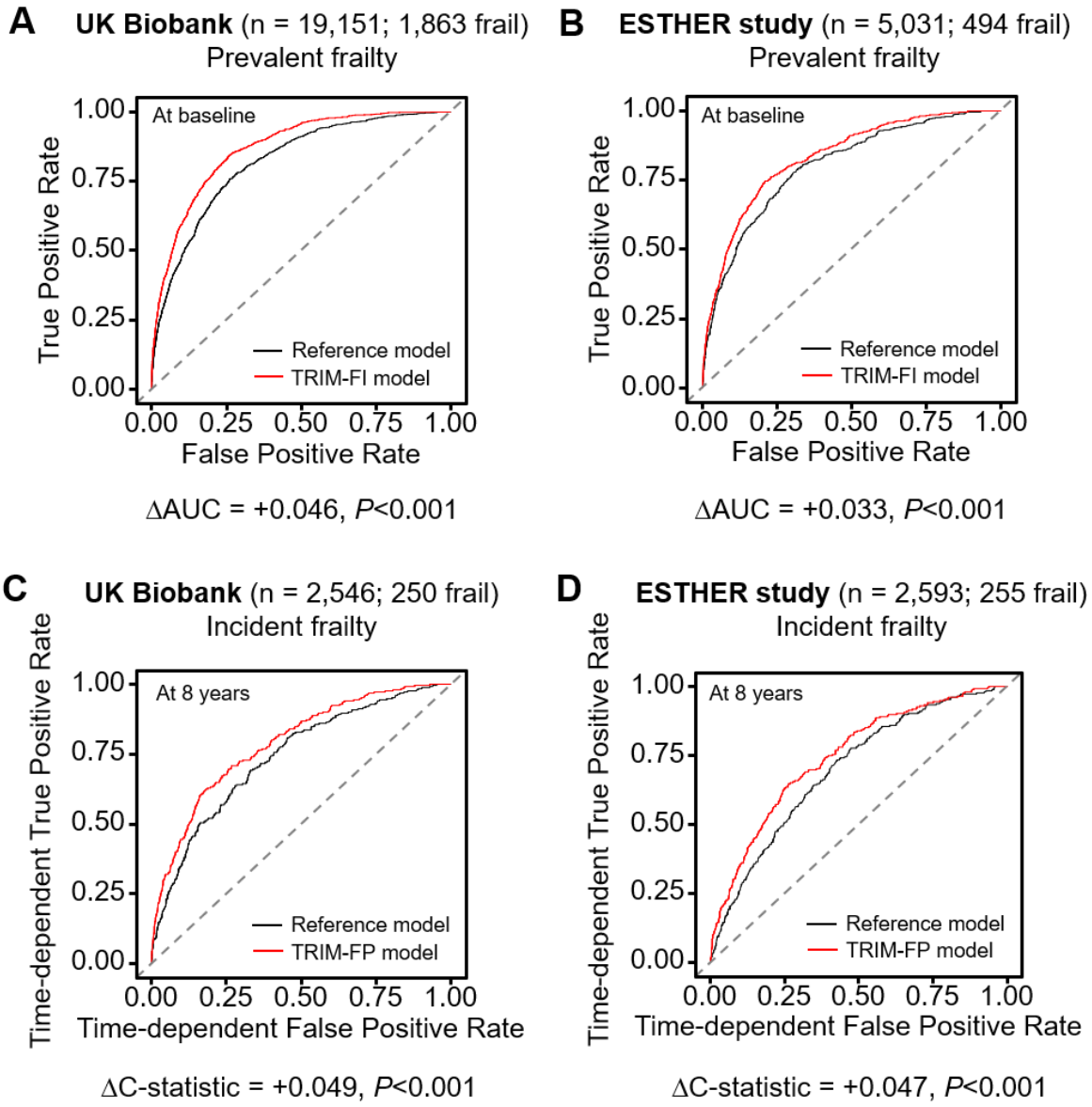



**Figure 13. ROC curves for discrimination of the TRIM-FI and TRIM-FP models for prevalent and incident frailty, respectively**

**(A)**, **(B)**, ROC curves and AUCs for prevalent frailty at baseline in the UK Biobank derivation cohort and the ESTHER external validation cohort, respectively. **(C)**, **(D)**, Time-dependent ROC curves and Harrell's C-statistics for incident frailty in the UK Biobank and ESTHER study, respectively. The reference models included conventional risk factors (chronological age, sex, education, smoking, physical activity, alcohol intake, BMI, dyslipidemia, hypertension, diabetes, cardiovascular disease, and history of cancer). The TRIM models additionally included the LASSO-selected proteomic and metabolomic biomarkers.

Abbreviations: AUC, Area Under the Curve; ROC, Receiver Operating Characteristic; TRIM-FI, Traditional Risk factor, Inflammation, and Metabolism based Frailty Identification model; TRIM-FP, Traditional Risk factor, Inflammation, and Metabolism based Frailty Prediction model.

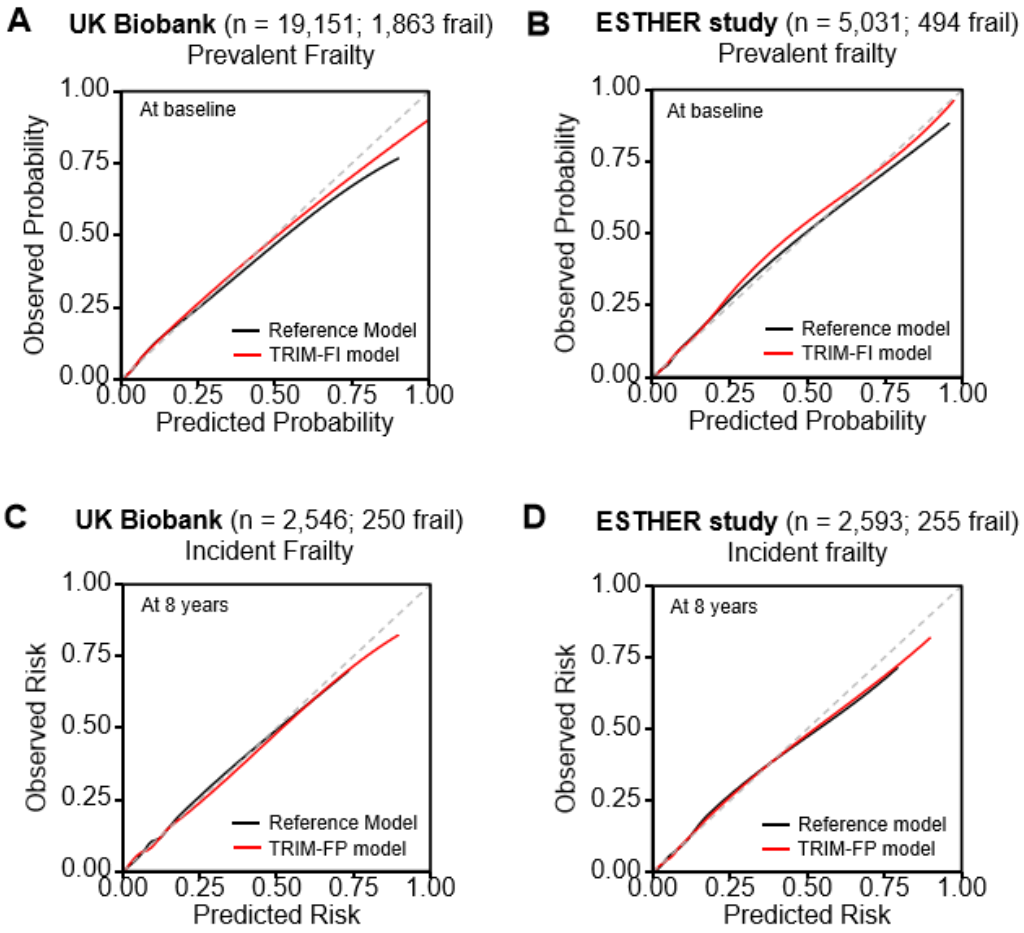



**Figure 14. Calibration of the TRIM models for prevalent and incident frailty**

**(A)**, **(B)**, Calibration curves for prevalent frailty at baseline in the UK Biobank (derivation cohort) and the ESTHER study (external validation cohort), respectively. **(C)**, **(D)**, Calibration curves for incident frailty in the UK Biobank and ESTHER study, respectively.

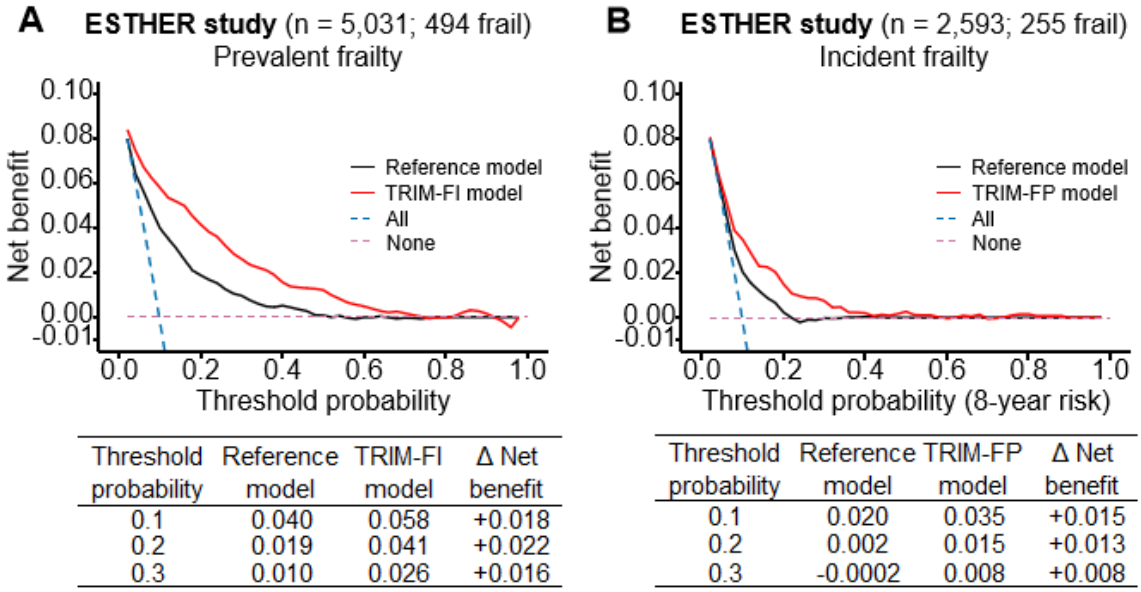
The reference model includes conventional risk factors (chronological age, sex, education, smoking, physical activity, alcohol intake, BMI, dyslipidemia, hypertension, diabetes, cardiovascular disease, and history of cancer). The TRIM models included the aforementioned conventional risk factors, 17 proteins, and 9 metabolites (**Appendix, Table A6**).

Abbreviations: BMI, body mass index; TRIM-FI, Traditional Risk factor, Inflammation, and Metabolism based Frailty Identification model; TRIM-FP, Traditional Risk factor, Inflammation, and Metabolism based Frailty Prediction model; the abbreviations for proteins and metabolites are listed in **Appendix, Table A5**.

### 3.2.4 Clinical utility of the TRIM models for prevalent and incident frailty in community-based populations

Decision curve analysis showed sustained net benefit of the TRIM models across threshold probabilities between 0.10 and 0.30 for both prevalent and incident frailty in the external validation cohort (**Figure 15**). Across this clinically relevant range, absolute differences in net benefit ranged from 0.016 to 0.022 for prevalent frailty and from 0.008 to 0.015 for incident frailty. The incremental net benefit diminished beyond threshold probabilities of 0.30.

To facilitate clinical and research use, an interactive web-based calculator for the TRIM models was developed (https://trim-riskclassifier.shinyapps.io/trimcalculator/), allowing for streamlined individualized frailty assessment and 8-year risk prediction.




| Threshold probability | Reference model | TRIM-FI model | Δ Net benefit |
|---|---|---|---|
| 0.1 | 0.040 | 0.058 | +0.018 |
| 0.2 | 0.019 | 0.041 | +0.022 |
| 0.3 | 0.010 | 0.026 | +0.016 |

| Threshold probability | Reference model | TRIM-FP model | Δ Net benefit |
|---|---|---|---|
| 0.1 | 0.020 | 0.035 | +0.015 |
| 0.2 | 0.002 | 0.015 | +0.013 |
| 0.3 | -0.0002 | 0.008 | +0.008 |

**Figure 15. Decision curve analysis in the external validation cohort (ESTHER study)**

(**A**), the TRIM-FI model for prevalent frailty. (**B**), the TRIM-FP model for incident frailty. Net benefit is plotted across a range of threshold probabilities to quantify the clinical utility of the reference models and the TRIM models. Abbreviations: TRIM-FI, Traditional Risk factor, Inflammation, and Metabolism based Frailty Identification model; TRIM-FP, Traditional Risk factor, Inflammation, and Metabolism based Frailty Prediction model.

### 3.2.5 Sensitivity analyses using incident frailty-derived omics biomarkers

To assess the robustness of biomarker selection, the same bootstrap-enhanced LASSO feature-selection procedure was repeated in UKB using incident frailty as the outcome. The resulting biomarker panel, comprising 9 proteomic and 5 metabolomic biomarkers, partially overlapped with that derived from prevalent frailty, including 6 shared proteins (*FGF23*, *HGF*, *IL6*, *TGFA*, *TNFRSF9*, and *VEGFA*) and 4 shared metabolites (albumin, glucose, GlycA, and lactate) (**Figure 16**).

In prediction analyses, the incident frailty-derived biomarker panel yielded consistent improvements in model performance, although the effect sizes were approximately half those of the prevalent frailty-derived biomarker panel (**Table 8**). In ESTHER, the time-dependent C-statistic increased from 0.699 to 0.726 (ΔC=0.027, *P*=0.006), accompanied by improvements in continuous NRI (13.9%), categorical NRI (6.5%), and IDI (1.7%).

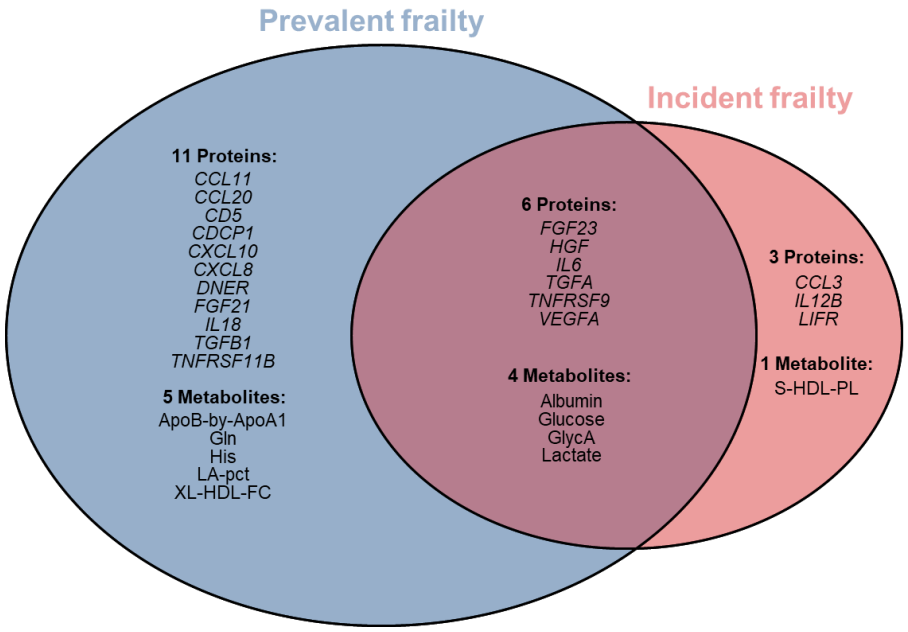



**Figure 16. Overlap of proteomic and metabolomic biomarkers selected at baseline and follow-up in the UK Biobank**

Venn diagram illustrating the overlap between biomarker panels identified using the same bootstrap-enhanced LASSO procedure in the UK Biobank, with prevalent frailty at baseline and incident frailty during an 8-year follow-up as feature-selection outcomes. Among the selected omics biomarkers, 6 proteins and 4 metabolites were shared between the approaches, suggesting partially consistent molecular signatures underlying both prevalent and incident frailty.

**Table 8. Sensitivity analysis: predictive value of proteomic and metabolomic biomarkers selected only based on incident frailty in the UK Biobank and validated in the ESTHER study**

| Outcome / Cohort / Model | C-statistic (95% CI) [f] | *Δ*C-statistics [f] | *Δ* Continuous NRI (95% CI) | *Δ* Categorical NRI (95% CI) [g] | *Δ* IDI (95% CI) |
|---|---|---|---|---|---|
| **Incident frailty [a]** | | | | | |
| **UK Biobank (derivation set)** | | | | | |
| **n=2,546; 250 (9.8%) frail** | | | | | |
| Reference model [b] | 0.739 (0.706–0.771) | Reference | — | — | — |
| + Proteomics [c] | 0.760 (0.728–0.792) | +0.021, *P*=0.010 | 13.4% (2.1%–24.7%), *P*=0.020 | 5.8% (0.6%–11.0%), *P*=0.028 | 1.7% (0.4%–3.0%), *P*=0.011 |
| + Metabolomics [d] | 0.754 (0.722–0.786) | +0.015, *P*=0.041 | 10.2% (0.8%–19.6%), *P*=0.033 | 4.1% (0.1%–8.1%), *P*=0.046 | 1.2% (0.2%–2.2%), *P*=0.018 |
| Combined model [e] | 0.769 (0.738–0.800) | +0.030, *P*=0.002 | 17.8% (6.5%–29.1%), *P*=0.001 | 8.6% (3.1%–14.1%), *P*=0.002 | 2.3% (0.9%–3.7%), *P*=0.001 |
| **ESTHER Study (validation set)** | | | | | |
| **n=2,593; 255 (9.8%) frail** | | | | | |
| Reference model [b] | 0.699 (0.667–0.731) | Reference | — | — | — |
| + Proteomics [c] | 0.714 (0.682–0.746) | +0.015, *P*=0.048 | 9.6% (0.7%–18.5%), *P*=0.032 | 3.9% (0.2%–7.6%), *P*=0.039 | 1.1% (0.1%–2.1%), *P*=0.026 |
| + Metabolomics [d] | 0.709 (0.677–0.741) | +0.010, *P*=0.092 | 7.4% (-0.3%–15.1%), *P*=0.059 | 2.8% (-0.2%–5.8%), *P*=0.067 | 0.8% (0.0%–1.6%), *P*=0.050 |
| Combined model [e] | 0.726 (0.695–0.757) | +0.027, *P*=0.006 | 13.9% (3.8%–24.0%), *P*=0.007 | 6.5% (1.2%–11.8%), *P*=0.016 | 1.7% (0.5%–2.9%), *P*=0.005 |

Abbreviations: BMI, body mass index; C-statistic, concordance statistic; CI, confidence interval; IDI, integrated discrimination improvement; NRI, net reclassification improvement.

Notes: Differences in time-dependent C-statistics were assessed using bootstrap resampling, and NRI and IDI were tested using likelihood ratio-based inference.

[a] Incident frailty analysis during a follow-up period of 8 years (median of 7.8 years in the UK Biobank and 7.7 years in the ESTHER study).

[b] The reference model included the following 12 conventional risk factors: chronological age, sex, education, smoking status, physical activity, alcohol consumption, BMI, dyslipidemia, hypertension, diabetes, cardiovascular disease, and history of cancer.

[c] The proteomics model included the variables of the reference model plus 9 proteins: *CCL3*, *FGF23*, *HGF*, *IL12B*, *IL6*, *LIFR*, *TGFA*, *TNFRSF9*, and *VEGFA* (see **Appendix, Table A5** for abbreviations).

[d] The metabolomics model included the variables of the reference model plus 5 metabolites: Albumin, Glucose, GlycA, Lactate, and S-HDL-PL (see **Appendix, Table A5** for abbreviations).

[e] The combined model included the reference model plus the above 9 proteins and 5 metabolites.

[f] AUC for baseline frailty or Harrell's C-statistic for incident frailty.

[g] Categorical NRI was calculated based on predefined risk thresholds of <5%, 5%–10%, and >10% (Ersbøll et al. 2013; Peng et al. 2026).

## 3.3 Study 3: Metabolomics-based biological ageing clocks for mortality

### 3.3.1 Distribution of traditional risk factors and their associations with 10-year mortality in UKB and ESTHER

The baseline characteristics of the 6,820 participants in ESTHER were broadly comparable to the 209,144 participants in UKB (**Table 9**), with similar mean age (60.4 ± 5.5 years *vs.* 60.0 ± 5.4 years) and proportion of women (55.0% *vs.* 53.9%). The distribution of education levels, physical activity levels, body mass index (BMI), and the prevalence of hypertension, diabetes, cardiovascular disease, cancer history, and treated dyslipidemia were also consistent across cohorts. Differences were observed in smoking and alcohol consumption, with the proportion of current smokers being higher in ESTHER than in UKB (17.6% *vs.* 9.5%) and the proportion of participants with moderate or high alcohol consumption being higher in UKB than in ESTHER (29.4% *vs.* 8.1%).

During a median follow-up of 10 years (mean duration: 9.75 years in UKB and 9.56 years in ESTHER), a total of 12,347 deaths occurred in UKB and 804 deaths in ESTHER (**Figure 5**).

In both cohorts, older age, male sex, current smoking, low physical activity, low weight (BMI<21.5 kg/m²), high weight (BMI≥35 kg/m²), hypertension, diabetes, cardiovascular disease, and cancer history were associated with an increased risk of 10-year all-cause mortality (**Table 10**). The HRs for these risk and preventive factors were highly comparable between the two cohorts. Treated dyslipidemia was not significantly associated with mortality in either cohort. Notable differences were observed for medium or high alcohol consumption (associated with increased mortality in UKB only) and higher education (associated with lower mortality in UKB only).

**Table 9. Baseline characteristics of study participants in the UK Biobank and ESTHER study**

| Characteristics | UK Biobank | | | ESTHER Study |
|---|---|---|---|---|
| | Derivation Set (n=146,400) | Internal Test Set (n=62,744) | Overall (n=209,144) | External Validation Set (n=6,820) |
| **Age, years** | 60.0 (5.4) | 60.0 (5.4) | 60.0 (5.4) | 60.4 (5.5) |
| **Women, n (%)** | 78,856 (53.9) | 33,976 (54.2) | 112,832 (53.9) | 3,753 (55.0) |
| **High education level, n (%)** [a, g] | 41,319 (28.2) | 17,655 (28.1) | 58,974 (28.2) | 1,744 (25.6) |
| **Smoking status, n (%)** [g] | / | / | / | / |
| Never smoker | 77,236 (52.8) | 33,109 (52.8) | 110,345 (52.8) | 3,420 (50.1) |
| Former smoker | 55,227 (37.7) | 23,686 (37.7) | 78,913 (37.7) | 2,198 (32.3) |
| Current smoker | 13,937 (9.5) | 5,949 (9.5) | 19,886 (9.5) | 1,202 (17.6) |
| **Physical activity, n (%)** [g] | / | / | / | / |
| Low | 30,595 (20.9) | 13,277 (21.2) | 43,872 (21.0) | 1,347 (19.8) |
| Moderate | 60,424 (41.3) | 25,693 (40.9) | 86,117 (41.2) | 3,089 (45.2) |
| High | 55,381 (37.8) | 23,774 (37.9) | 79,155 (37.8) | 2,384 (35.0) |
| **Alcohol consumption, n (%)** [b, g] | / | / | / | / |
| Abstainer | 44,669 (30.5) | 19,280 (30.7) | 63,949 (30.6) | 2,242 (32.9) |
| Low | 58,589 (40.0) | 25,074 (40.0) | 83,663 (40.0) | 4,028 (59.0) |
| Medium or high | 43,142 (29.5) | 18,390 (29.3) | 61,532 (29.4) | 550 (8.1) |
| **Body mass index, n (%)** [c] | / | / | / | / |
| $< 21.5$ kg/m$^2$ | 9,560 (6.4) | 4,062 (6.5) | 13,622 (6.5) | 352 (5.2) |
| $\geq 21.5 - < 25$ kg/m$^2$ | 36,513 (25.0) | 15,736 (25.1) | 52,249 (25.0) | 1,539 (22.6) |
| $\geq 25 - < 30$ kg/m$^2$ | 63,552 (43.4) | 27,126 (43.2) | 90,678 (43.3) | 3,189 (46.7) |
| $\geq 30 - < 35$ kg/m$^2$ | 26,709 (18.2) | 11,345 (18.1) | 38,054 (18.2) | 1,346 (19.7) |
| $\geq 35$ kg/m$^2$ | 10,066 (7.0) | 4,475 (7.1) | 14,541 (7.0) | 394 (5.8) |
| **Medical conditions, n (%)** | / | / | / | / |
| Treated dyslipidemia [g] | 31,879 (21.8) | 13,566 (21.6) | 45,445 (21.7) | 712 (10.4) |
| Hypertension [d, g] | 85,968 (58.7) | 36,819 (58.7) | 122,787 (58.7) | 4,146 (60.8) |
| Diabetes [e] | 15,075 (10.3) | 6,503 (10.4) | 21,578 (10.3) | 964 (14.1) |
| Cardiovascular disease [f, g] | 10,791 (7.4) | 4,567 (7.3) | 15,358 (7.3) | 503 (7.4) |
| History of cancer [g] | 12,825 (8.8) | 5,541 (8.8) | 18,366 (8.8) | 471 (6.9) |

Notes: Values are means (SDs) for continuous variables and percentages for categorical variables.

[a] In the UK Biobank, participants who reported "None of the above," "Certificate of Secondary Education (CSEs) or equivalent," "Ordinary Levels/General Certificate of Secondary Education (O levels/GCSEs) or equivalent," "Other professional qualifications (e.g., nursing, teaching)," or "Advanced Levels/Advanced Subsidiary Levels (A levels/AS levels) or equivalent" were categorized as having a low or medium education level. Those who reported "National Vocational Qualification (NVQ), Higher National Diploma (HND), Higher National Certificate (HNC) or equivalent," or a

"College or University degree" were categorized as having a high education level. In the German ESTHER study, participants who reported "No education or 'Hauptschule'" or "'Realschule'/'Mittlere Reife'" were categorized as having a low education level, whereas those who reported "'Fachhochschulreife' or 'Abitur'" were categorized as having a high education level.

[b] Alcohol consumption was categorized as low for women consuming 0–19.99 g/day and men consuming 0–39.99 g/day. It was categorized as medium or high for women consuming 20–39.99 g/day or ≥40 g/day, and for men consuming 40–59.99 g/day or ≥60 g/day.

[c] To enhance comparability between the two cohorts, body mass index (BMI) categories were defined according to the WHO classification with slight modifications to the cut-off values. Participants with a BMI less than 21.5 kg/m², including both underweight individuals and those at the lower end of the normal range, were grouped together to account for population-specific differences and to ensure sufficient sample sizes across categories.

[d] Hypertension was defined as self-reported hypertension, the use of antihypertensive medication, a systolic blood pressure ≥140 mm Hg, or a diastolic blood pressure ≥90 mm Hg.

[e] In this study, diabetes was defined as self-reported diabetes, the use of glucose-lowering medication, or an $HbA_{1c}$ level of ≥6.5%.

[f] Cardiovascular disease in this study was defined as having myocardial infarction or stroke.

[g] Missing values for education level (1.2%), smoking status (0.1%), physical activity (9.6%), hypertension (0.02%), diabetes (0.02%), lipid-lowering medication (0.02%), and cancer (0.4%) in the UK Biobank, as well as, education level (2.0%), smoking status (2.4%), physical activity (0.2%), alcohol consumption (8.0%), diabetes (1.7%), and cardiovascular disease (2.6%) in the ESTHER study, were imputed using random forest imputation with the R package *missRanger* (version 2.1.5).

**Table 10. Longitudinal association between traditional risk factors and 10-year all-cause mortality in the UK Biobank and ESTHER study**

| Baseline Characteristics | UK Biobank (209,144 participants) | | | ESTHER Study (6,820 participants) | | |
|---|---|---|---|---|---|---|
| | $N_{death}$ (%) | HR (95% CI) [a] | *P*-value | $N_{death}$ (%) | HR (95% CI) [a] | *P*-value |
| **Age** | | | | | | |
| 50–59 years old | 3,056 (3.4) | Ref. | | 215 (5.3) | Ref. | |
| 60–69 years old | 9,291 (7.9) | 2.31 (2.22–2.40) | <0.001 | 589 (14.5) | 2.09 (1.81–2.42) | <0.001 |
| **Sex** | | | | | | |
| Women | 4,852 (4.3) | Ref. | | 335 (8.9) | Ref. | |
| Men | 7,495 (7.8) | 1.71 (1.65–1.78) | <0.001 | 469 (15.3) | 1.72 (1.47–2.02) | <0.001 |
| **Education level** | | | | | | |
| Medium or low | 8,979 (6.0) | Ref. | | 622 (12.3) | Ref. | |
| High | 3,368 (5.7) | 0.95 (0.92–0.99) | 0.036 | 182 (10.4) | 0.99 (0.85–1.18) | 0.849 |
| **Smoking status** | | | | | | |
| Never smoker | 4,625 (4.2) | Ref. | | 312 (9.1) | Ref. | |
| Former smoker | 5,402 (6.9) | 1.36 (1.30–1.42) | <0.001 | 284 (12.9) | 1.20 (1.01–1.43) | 0.041 |
| Current smoker | 2,320 (11.7) | 2.66 (2.53–2.80) | <0.001 | 208 (17.3) | 2.15 (1.78–2.59) | <0.001 |
| **Physical activity** | | | | | | |
| Low | 3,198 (7.3) | Ref. | | 213 (15.8) | Ref. | |
| Moderate | 4,849 (5.6) | 0.82 (0.79–0.86) | <0.001 | 378 (12.2) | 0.78 (0.66–0.92) | 0.004 |
| High | 4,300 (5.4) | 0.78 (0.74–0.81) | <0.001 | 213 (8.9) | 0.63 (0.52–0.77) | <0.001 |
| **Alcohol consumption** | | | | | | |
| Abstainer | 3,281 (5.1) | Ref. | | 242 (10.8) | Ref. | |
| Low | 4,493 (5.4) | 1.08 (1.01–1.17) | 0.011 | 496 (12.3) | 1.12 (1.02–1.23) | 0.037 |
| Medium or high | 4,573 (7.4) | 1.39 (1.33–1.46) | <0.001 | 66 (12.0) | 1.04 (0.77–1.35) | 0.919 |
| **Body mass index** | | | | | | |
| < 21.5 kg/m$^2$ | 947 (7.0) | 1.52 (1.41–1.64) | <0.001 | 50 (14.2) | 1.66 (1.20–2.30) | 0.002 |
| ≥ 21.5 – < 25 kg/m$^2$ | 2,578 (4.9) | Ref. | | 154 (10.0) | Ref. | |
| ≥ 25 – < 30 kg/m$^2$ | 5,085 (5.6) | 0.94 (0.89–0.98) | 0.007 | 350 (11.0) | 0.92 (0.76–1.11) | 0.378 |
| ≥ 30 – < 35 kg/m$^2$ | 2,475 (6.5) | 0.98 (0.93–1.04) | 0.576 | 179 (13.3) | 1.03 (0.82–1.29) | 0.801 |
| ≥ 35 kg/m$^2$ | 1,262 (8.7) | 1.30 (1.21–1.39) | <0.001 | 71 (18.0) | 1.51 (1.13–2.02) | 0.006 |
| **Treated dyslipidemia** | | | | | | |
| No | 8,072 (4.9) | Ref. | | 692 (11.3) | Ref. | |
| Yes | 4,275 (9.4) | 0.98 (0.93–1.02) | 0.319 | 112 (15.7) | 0.98 (0.80–1.21) | 0.858 |

| Baseline Characteristics | UK Biobank (209,144 participants) | | | ESTHER Study (6,820 participants) | | |
|---|---|---|---|---|---|---|
| | $N_{death}$ (%) | HR (95% CI) [a] | *P*-value | $N_{death}$ (%) | HR (95% CI) [a] | *P*-value |
| **Hypertension** | | | | | | |
| No | 3,515 (4.1) | Ref. | | 230 (8.6) | Ref. | |
| Yes | 8,832 (7.2) | 1.23 (1.18–1.29) | <0.001 | 574 (13.8) | 1.25 (1.06–1.47) | 0.009 |
| **Diabetes** | | | | | | |
| No | 9,844 (5.3) | Ref. | | 595 (10.2) | Ref. | |
| Yes | 2,503 (11.6) | 1.50 (1.43–1.57) | <0.001 | 209 (21.7) | 1.66 (1.41–1.96) | <0.001 |
| **Cardiovascular disease** | | | | | | |
| No | 10,130 (5.2) | Ref. | | 677 (10.7) | Ref. | |
| Yes | 2,217 (14.4) | 1.75 (1.66–1.84) | <0.001 | 127 (25.3) | 1.68 (1.37–2.06) | <0.001 |
| **Cancer** | | | | | | |
| No | 10,040 (5.3) | Ref. | | 698 (11.0) | Ref. | |
| Yes | 2,307 (12.5) | 2.31 (2.20–2.41) | <0.001 | 106 (22.5) | 2.12 (1.73–2.61) | <0.001 |

[a] Cox proportional hazards regression model included all covariates: age, sex, education level, smoking status, physical activity, alcohol consumption, body mass index, treated dyslipidemia, hypertension, diabetes, cardiovascular disease, and cancer. Proportional hazards assumption was evaluated using Schoenfeld residuals (Sjölander and Dickman 2024; Stensrud and Hernán 2020). CI, confidence interval; HR, hazard ratio.

### 3.3.2 Associations of metabolomic biomarkers with all-cause mortality

Among 209,144 UKB participants, 224 of the 249 metabolomic biomarkers were significantly associated with 10-year all-cause mortality after adjustment for conventional risk factors and multiple testing (FDR-adjusted $P$<0.05) (**Figure 17A**). Of these, 68 biomarkers were validated in the independent ESTHER (n=6,820), including 26 with positive and 42 with negative associations with mortality (FDR-adjusted $P$<0.05) (**Figure 17B**). Most of the validated biomarkers were related to lipid metabolism.

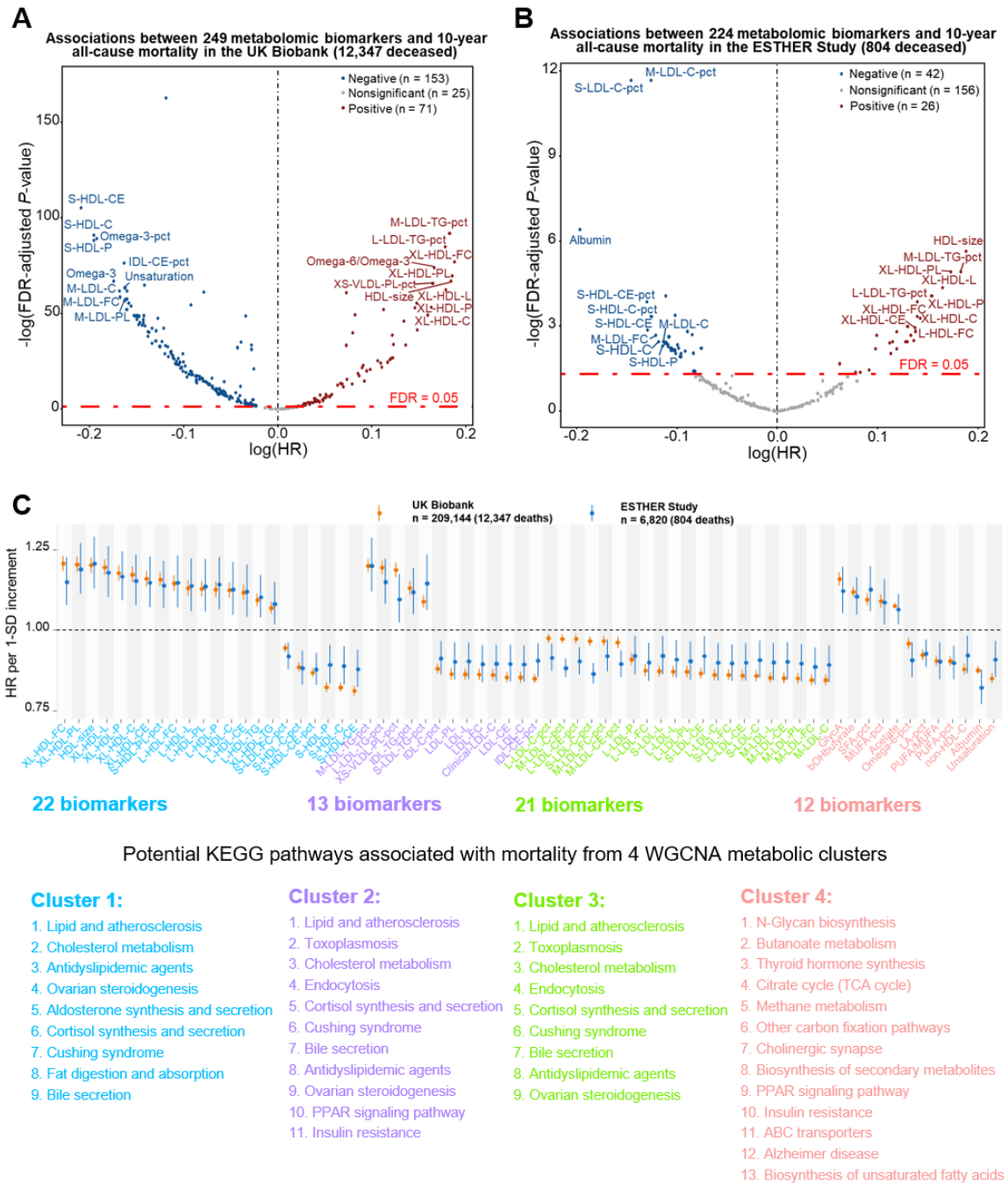



**Figure 17. Metabolome-wide discovery and validation of biomarkers associated with 10-year all-cause mortality in the UK Biobank and ESTHER study**

**(A)** In the UK Biobank (n = 209,144; 12,347 deaths), 224 of 249 measured metabolomic biomarkers showed significant associations with 10-year all-cause mortality in multivariable Cox regression analyses. The 10 biomarkers with the strongest positive and inverse HRs are highlighted by their abbreviations (the full names are detailed in **Appendix, Table A2**). **(B)** Of these, 68 biomarkers were validated in the ESTHER study (n = 6,820; 804 deaths). The corresponding volcano plot depicts their effect estimates. **(C)** Forest plots display HRs with 95% CIs per 1-SD increment for the validated metabolomic biomarkers in both cohorts. The 68 biomarkers were grouped into four co-expression clusters by WGCNA; KEGG pathway enrichment analyses for each cluster suggested biological processes of potential relevance, summarized below the plot. All models were adjusted for age, sex, education level, smoking status, physical activity, alcohol consumption, body-mass index, treated dyslipidemia, hypertension, diabetes, cardiovascular disease, and cancer.
Abbreviations: CI, confidence interval; FDR, false discovery rate; HR, hazard ratio; KEGG, Kyoto Encyclopedia of Genes and Genomes; WGCNA, Weighted Gene Co-expression Network Analysis.

### 3.3.3 Metabolic pathways associated with all-cause mortality

To take into account inter-correlations of the 68 validated metabolomic biomarkers, I conducted weighted gene co-expression network analysis (WGCNA) in UKB and identified 4 clusters (**Figure 18**). **Table 11** lists the Kyoto Encyclopedia of Genes and Genomes (KEGG)-based pathways enriched in these clusters. **Figure 17C** gives an overview on the strengths of the associations of the 68 validated metabolomic biomarkers, their clustering and potential KEGG pathways. Cluster 1 (22 biomarkers) primarily consists of high-density lipoprotein (HDL) particles and is enriched in pathways related to lipid metabolism, cardiovascular function, and endocrine regulation. Clusters 2 and 3 (13 and 21 biomarkers, respectively) are dominated by intermediate-density lipoprotein (IDL), low-density lipoprotein (LDL), and very-low-density lipoprotein (VLDL) particles and are associated with pathways involved in cardiovascular disease, lipid metabolism, and metabolic disorders. Cluster 4 (12 biomarkers) includes GlycA, *β*-hydroxybutyrate (bOHbutyrate), various fatty acid markers, acetate, non-HDL cholesterol, and albumin. This cluster is enriched in pathways related to carbohydrate metabolism, energy metabolism, endocrine signaling, and neurodegenerative diseases, including Alzheimer's disease.

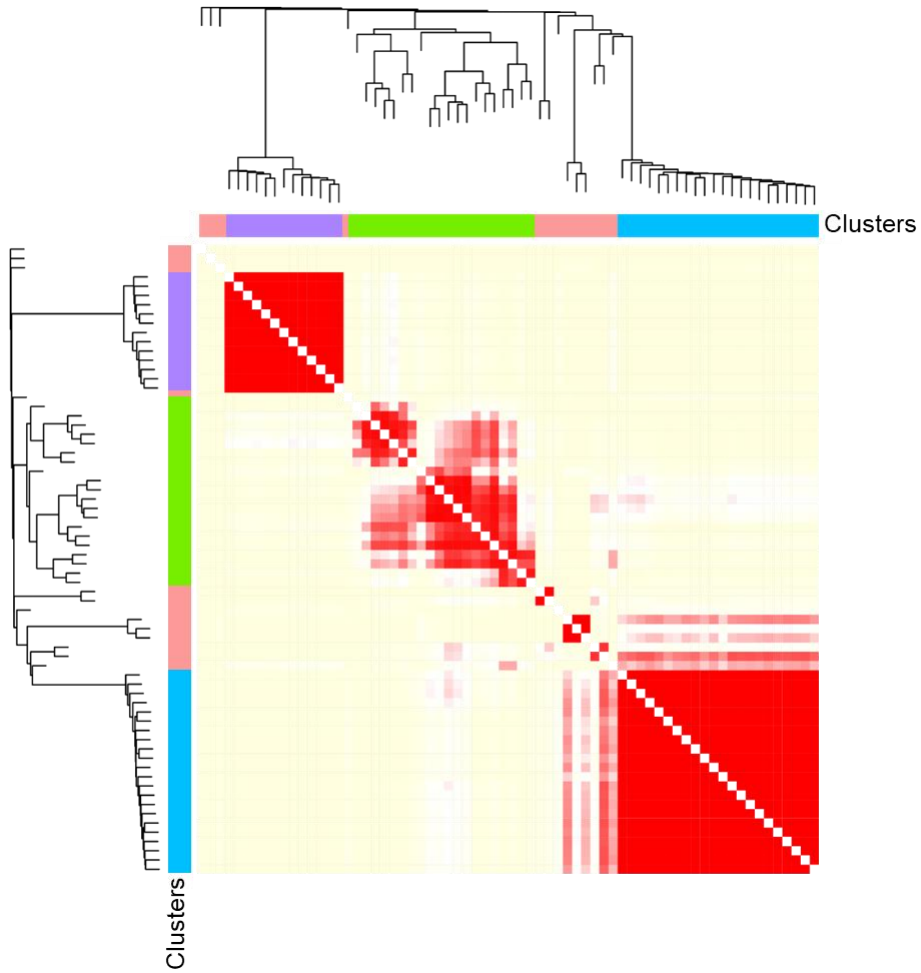



**Figure 18. Topological overlap map of 68 metabolomic biomarkers**

Metabolic biomarkers are arranged in rows and columns based on the clustering tree. Light yellow shades indicate low topological overlap (low similarity), while red shades indicate higher overlap and similarity (Langfelder and Horvath 2008). Metabolite clusters are represented by the squares along the diagonal.

**Table 11. Summary of potential KEGG pathways associated with 68 validated metabolic biomarkers identified from WGCNA-derived metabolite clusters**

| Metabolite clusters | Class of pathways [a] | Subclass of pathways | Pathways in KEGG database |
|---|---|---|---|
| Cluster 1 | | | |
| | Human Diseases | Cardiovascular disease | Lipid and atherosclerosis (map05417) |
| | Organismal Systems | Digestive system | Cholesterol metabolism (map04979) |
| | Drug Development | Chronology: Other drugs | Antidyslipidemic agents (map07052) |
| | Organismal Systems | Endocrine system | Ovarian steroidogenesis (map04913) |
| | Organismal Systems | Endocrine system | Aldosterone synthesis and secretion (map04925) |
| | Organismal Systems | Endocrine system | Cortisol synthesis and secretion (map04927) |
| | Human Diseases | Endocrine and metabolic disease | Cushing syndrome (map04934) |
| | Organismal Systems | Digestive system | Fat digestion and absorption (map04975) |
| | Organismal Systems | Digestive system | Bile secretion (map04976) |
| Cluster 2 | | | |
| | Human Diseases | Cardiovascular disease | Lipid and atherosclerosis (map05417) |
| | Human Diseases | Infectious disease: parasitic | Toxoplasmosis (map05145) |
| | Organismal Systems | Digestive system | Cholesterol metabolism (map04979) |
| | Cellular Processes | Transport and catabolism | Endocytosis (map04144) |
| | Organismal Systems | Endocrine system | Cortisol synthesis and secretion (map04927) |
| | Human Diseases | Endocrine and metabolic disease | Cushing syndrome (map04934) |
| | Organismal Systems | Digestive system | Bile secretion (map04976) |
| | Drug Development | Chronology: Other drugs | Antidyslipidemic agents (map07052) |
| | Organismal Systems | Endocrine system | Ovarian steroidogenesis (map04913) |
| | Organismal Systems | Endocrine system | PPAR signaling pathway (map03320) |
| | Human Diseases | Endocrine and metabolic disease | Insulin resistance (map04931) |
| Cluster 3 | | | |
| | Human Diseases | Cardiovascular disease | Lipid and atherosclerosis (map05417) |
| | Human Diseases | Infectious disease: parasitic | Toxoplasmosis (map05145) |
| | Organismal Systems | Digestive system | Cholesterol metabolism (map04979) |
| | Cellular Processes | Transport and catabolism | Endocytosis (map04144) |
| | Organismal Systems | Endocrine system | Cortisol synthesis and secretion (map04927) |
| | Human Diseases | Endocrine and metabolic disease | Cushing syndrome (map04934) |
| | Organismal Systems | Digestive system | Bile secretion (map04976) |
| | Drug Development | Chronology: Other drugs | Antidyslipidemic agents (map07052) |

| Metabolite clusters | Class of pathways [a] | Subclass of pathways | Pathways in KEGG database |
|---|---|---|---|
| | Organismal Systems | Endocrine system | Ovarian steroidogenesis (map04913) |
| Cluster 4 | | | |
| | Metabolism | Glycan biosynthesis and metabolism | N-Glycan biosynthesis (hsa00510) |
| | Metabolism | Carbohydrate metabolism | Butanoate metabolism (map00650) |
| | Organismal Systems | Endocrine system | Thyroid hormone synthesis (map04918) |
| | Metabolism | Carbohydrate metabolism | Citrate cycle (TCA cycle) (map00020) |
| | Metabolism | Energy metabolism | Methane metabolism (map00680) |
| | Metabolism | Energy metabolism | Other carbon fixation pathways (map00720) |
| | Organismal Systems | Nervous system | Cholinergic synapse (map04725) |
| | Metabolism | Global and overview maps | Biosynthesis of secondary metabolites (map01110) |
| | Organismal Systems | Endocrine system | PPAR signaling pathway (map03320) |
| | Human Diseases | Endocrine and metabolic disease | Insulin resistance (map04931) |
| | Environmental Information Processing | Membrane transport | ABC transporters (map02010) |
| | Human Diseases | Neurodegenerative disease | Alzheimer disease (map05010) |
| | Metabolism | Lipid metabolism | Biosynthesis of unsaturated fatty acids (map01040) |

[a] The KEGG PATHWAY database (Kanehisa et al. 2014) is a collection of manually curated pathway maps representing current knowledge of molecular interactions, reactions, and network relationships across various biological and biomedical areas, including: 1. Metabolism; 2. Genetic Information Processing; 3. Environmental Information Processing; 4. Cellular Processes; 5. Organismal Systems; 6. Human Diseases; and 7. Drug Development.

### 3.3.4 Selected metabolomic biomarkers for predicting 10-year all-cause mortality

To develop all-cause mortality prediction models, I applied least absolute shrinkage and selection operator (LASSO) regularized Cox regression to identify sex- and age-specific metabolomic biomarkers from a panel of 249 metabolites (**Figure 19**). A total of 20 biomarkers in younger men (**Figure 19A**), 18 in older men (**Figure 19B**), 12 in younger women (**Figure 19C**), and 13 in older women (**Figure 19D**) were selected for risk score development. In models adjusted for conventional risk factors, 8, 9, 6, and 7 biomarkers, respectively, were statistically significantly associated with 10-year all-cause mortality across the derivation (UKB 70%), internal validation (UKB 30%) and external validation (ESTHER) sets with consistent HRs per 1-SD increment. Among these validated biomarkers, the inflammatory biomarkers, GlycA and albumin, were statistically significantly associated with 10-year mortality in all age- and sex-specific strata. **Figure 20** ranks the standardized $\beta$-coefficients of the selected metabolomic biomarkers in the sex- and age-specific risk scores. GlycA and the Omega-6/Omega-3 ratio carried strong positive weights ($\beta$>0.15) in all 4 subgroups, whereas VLDL size (in all 4 subgroups), LA-pct (in 3 subgroups), valine, and albumin (each in 2 subgroups) showed large negative weights ($\beta$<-0.15) in more than one subgroup. These leading metabolites largely reflect pathways of systemic inflammation, energy metabolism, and lipid regulation.

Overall, 26 metabolomic biomarkers were selected and **Figure 19E** illustrates the overlap in selection of metabolites in the subgroups. Nine biomarkers were selected in all four subgroups, 12 biomarkers were uniquely detected in men, four biomarkers were identified exclusively in women, and one biomarker was selected specifically in the older men and older women subgroups. The subgroup-specific pathways highlighted distinct biological processes: in younger men, pathways related to insulin resistance, glucocorticoid activity, and cancer metabolism predominated; in older men, bile acid metabolism and viral infection pathways were enriched; in younger women, branched-chain amino acid metabolism was prominent; and in older women, pathways related to inflammation and host-microbial interactions were identified (**Figure 19F**).

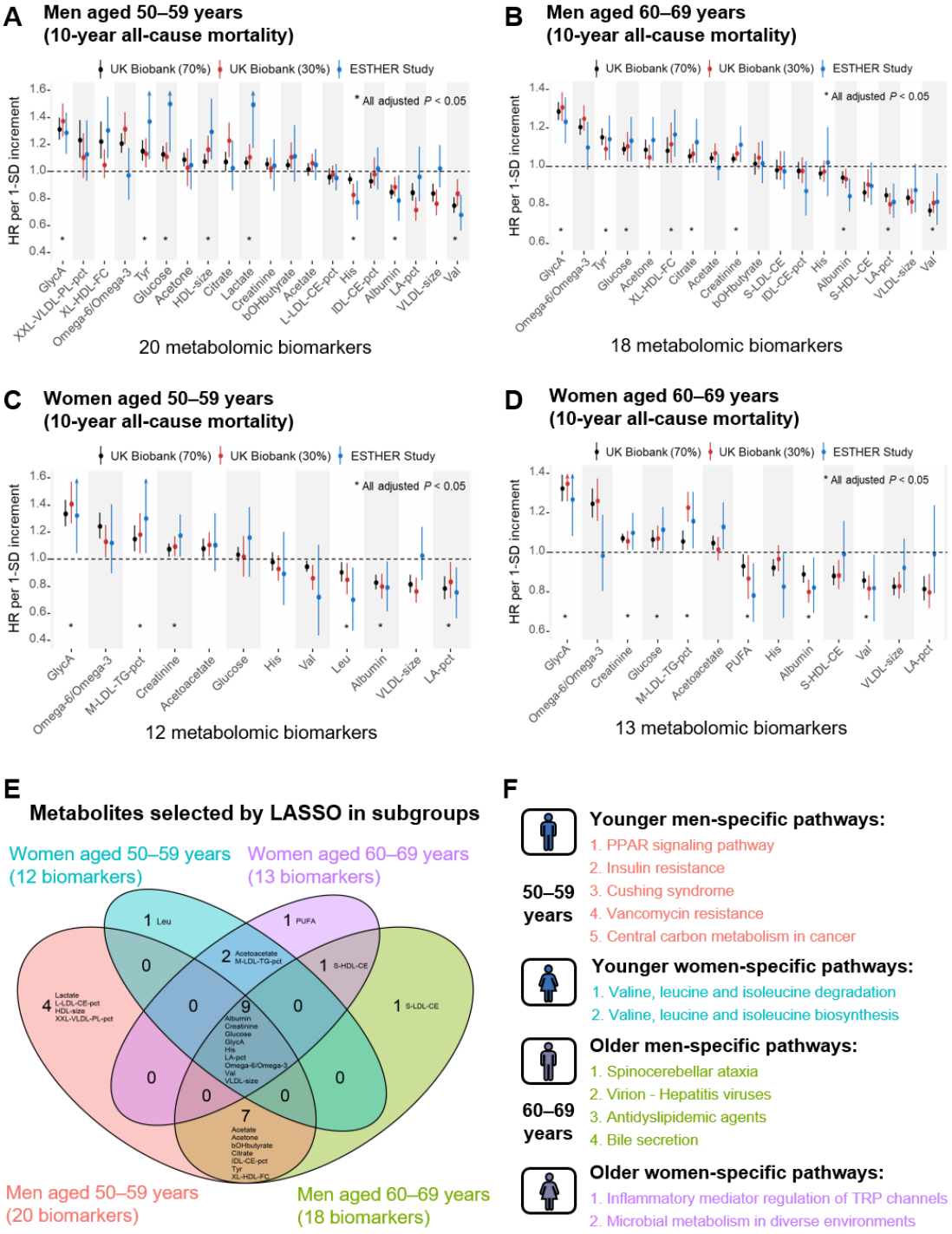



**Figure 19. Sex- and age-specific metabolomic biomarkers associated with 10-year all-cause mortality**

(**A**), Younger men. (**B**), Older men. (**C**), Younger women. (**D**), Older women. (**E**), Overlap of mortality-associated metabolomic biomarkers across age- and sex-specific subgroups. (**F**), KEGG pathway enrichment analysis of subgroup-specific metabolomic biomarkers. Associations between LASSO-selected metabolomic biomarkers and 10-year all-cause mortality (panels **A–D**) were assessed in multivariable Cox models adjusted for age, sex, education level, smoking status, physical activity, alcohol consumption, body mass index, treated dyslipidemia, hypertension, diabetes, cardiovascular disease, and cancer. HRs and 95% CIs per 1-SD increment are shown across the derivation set (black; 70% of the UK Biobank), the internal validation set (red; remaining 30% of the UK Biobank), and the external validation set (blue; ESTHER study). Asterisks (*) indicate biomarkers that were consistently and significantly associated with mortality across all three sets.

Abbreviations: HR, hazard ratio; CI, confidence interval; SD, standard deviation; LASSO, least absolute shrinkage and selection operator; KEGG, Kyoto Encyclopedia of Genes and Genomes.

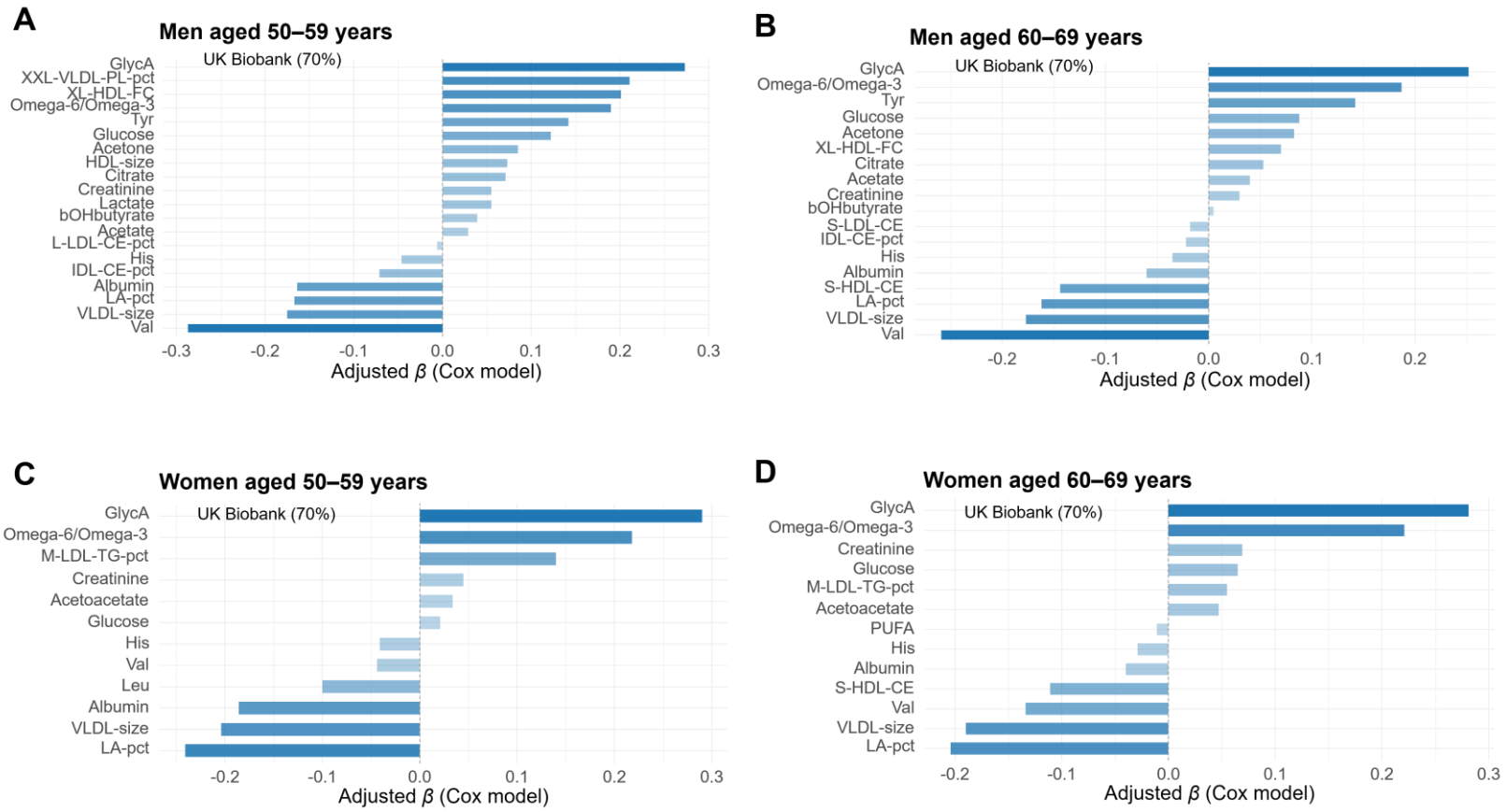



**Figure 20. Effect sizes of standardized β-coefficients of selected biomarkers used to derive mortality risk scores in the training set (70% of the UK Biobank)**

(**A**), Younger men. (**B**), Older men. (**C**), Younger women. (**D**), Older women. Biomarkers are ranked by their effect sizes. Biomarker names are abbreviated in the figure; full names are listed in **Appendix, Table A2**.

### 3.3.5 Discrimination and reclassification analyses of 10-year mortality prediction models including metabolomic biomarkers

Since the 26 selected metabolomic biomarkers were neither strongly correlated with the conventional risk factors (absolute Spearman correlation coefficients $|\rho| \leq 0.45$) nor with each other ($|\rho| \leq 0.67$) (**Figure 21**), I evaluated whether combined models of traditional risk factors and metabolomic biomarkers could enhance risk discrimination for 10-year all-cause mortality beyond conventional risk factors. Risk scores were derived in sex- and age-specific groups using Cox regression and validated in both internal and external validation cohorts. The variables' $\beta$-coefficients, obtained in the derivation set and used in the two validation sets, are provided in **Appendix, Table A7**.

Across all analyses, models based solely on metabolomic biomarkers yielded lower C-statistics compared to traditional risk factor models (**Table 12**). However, the C-statistics of the metabolomics models were significantly higher than 0.5 in the range of 0.639–0.684 across the subgroups for 10-year all-cause mortality prediction. The C-statistics for 5-year all-cause mortality prediction in the total study population were slightly higher than for 10-year all-cause mortality prediction in the derivation, internal validation, and external validation set (0.701, 0.685, and 0.673 versus 0.690, 0.679 and 0.662, respectively).

Combining metabolomic biomarkers with traditional risk factors significantly improved risk discrimination, as reflected by increases in the C-statistic compared to models with conventional risk factors alone (**Table 12; Figure 22**). Improvements in 10-year mortality prediction were observed in all the derivation (70% UKB), internal validation (30% UKB), and external validation (ESTHER) in younger men (+0.063, +0.055, and +0.084, respectively), older men (+0.053, +0.048, and +0.055, respectively), younger women (+0.054, +0.050, and +0.036, respectively), and older women **(**+0.047, +0.042, and +0.055, respectively). In the total study population, the C-statistic increased from 0.757 to 0.833 (+0.076), from 0.723 to 0.796 (+0.073), and from 0.720 to 0.788 (+0.068) in the derivation, internal validation, and external validation set, respectively. The improvement in the total population was even more pronounced for 5-year mortality risk prediction (C-statistic increase: +0.095, +0.087, +0.078, respectively).

Moreover, reclassification analyses demonstrated statistically significant improvements in net reclassification improvement (NRI) and integrated discrimination improvement (IDI) for 10-year all-cause mortality across all sex- and age-specific subgroups in both the internal and external validation cohorts, except for the group of younger women in ESTHER (**Table 13**). With the exception of this group, the ranges of the categorical NRIs, continuous NRIs, and IDIs were ranged

from 8.2%–18.9%, 24.4%–41.3%, and 4.6–12.4%, respectively across the strata and validation sets for 10-year mortality prediction. The ranges of categorical NRI, continuous NRI, and IDI were even higher for 5-year mortality prediction in the total study population: 19.1%–24.7%, 33.4%–42.2%, and 9.8%–12.3% respectively.

Calibration curves showed that models based on conventional risk factors were well-calibrated in both studies and all subgroups except for younger women in ESTHER (**Figure 23**). Adding metabolomic biomarkers further improved model calibration by bringing predicted and observed mortality rates closer to the ideal calibration line, which represents perfect agreement between predicted and observed cases. The model was also well calibrated in the group of younger women from ESTHER after inclusion of the metabolomic biomarkers.

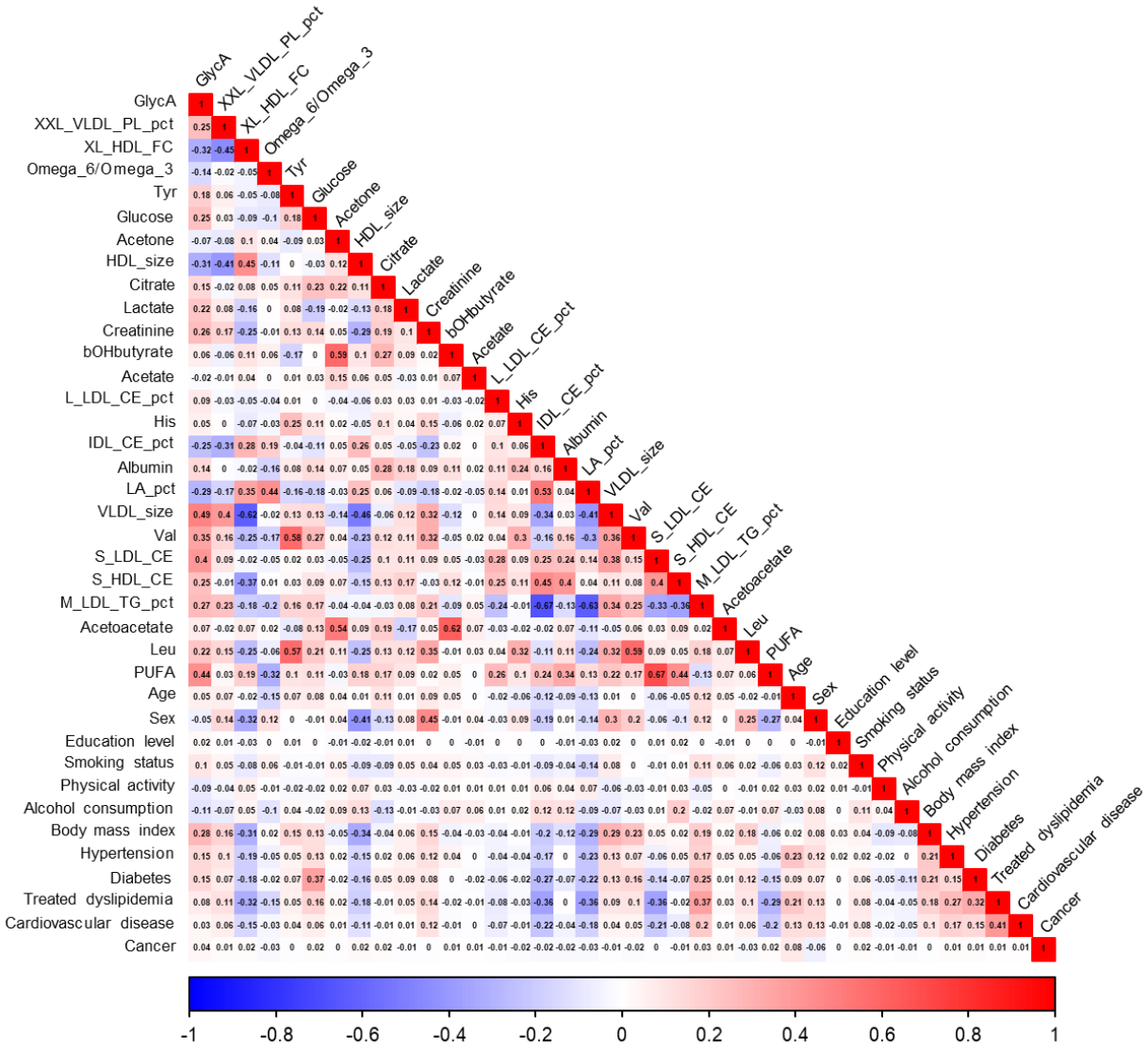



**Figure 21. Spearman correlations between the conventional risk factors and 26 metabolomic biomarkers in the derivation set (70% UK Biobank)**

**Table 12. C-statistics of prediction models for 10- and 5-year all-cause mortality in the UK Biobank and ESTHER cohorts**

| | C-statistic (*P*) [h] | | |
|---|---|---|---|
| | **Derivation set UK Biobank (70%)** | **Internal validation set UK Biobank (30%)** | **External validation set ESTHER** |
| **10-year all-cause mortality in men aged 50–59 years** | **n=28,158 (1,233 deaths)** | **n=12,043 (522 deaths)** | **n=1,557 (120 deaths)** |
| Traditional risk factor model [a] | 0.781 (0.766, 0.795) | 0.776 (0.754, 0.799) | 0.732 (0.688, 0.776) |
| Metabolomics model [b] | 0.684 (0.669, 0.698) | 0.672 (0.649, 0.696) | 0.658 (0.612, 0.706) |
| Combined model | 0.844 (0.830, 0.858) | 0.831 (0.811, 0.852) | 0.816 (0.779, 0.852) |
| | *ΔC*=0.063; *P*<0.001 | *ΔC*=0.055; *P*<0.001 | *ΔC*=0.084; *P*<0.001 |
| **10-year all-cause mortality in men aged 60–69 years** | **n=39,386 (4,041 deaths)** | **n=16,725 (1,699 deaths)** | **n=2,196 (349 deaths)** |
| Traditional risk factor model [a] | 0.742 (0.734, 0.750) | 0.733 (0.724, 0.741) | 0.724 (0.695, 0.753) |
| Metabolomics model [c] | 0.681 (0.672, 0.691) | 0.667 (0.656, 0.677) | 0.650 (0.621, 0.679) |
| Combined model | 0.795 (0.787, 0.803) | 0.781 (0.771, 0.790) | 0.779 (0.751, 0.808) |
| | *ΔC*=0.053; *P*<0.001 | *ΔC*=0.048; *P*<0.001 | *ΔC*=0.055; *P*<0.001 |
| **10-year all-cause mortality in women aged 50–59 years** | **n=35,553 (890 deaths)** | **n=15,254 (411 deaths)** | **n=1,201 (95 deaths)** |
| Traditional risk factor model [a] | 0.746 (0.728, 0.764) | 0.737 (0.709, 0.765) | 0.725 (0.676, 0.774) |
| Metabolomics model [d] | 0.678 (0.660, 0.697) | 0.665 (0.637, 0.693) | 0.647 (0.594, 0.700) |
| Combined model | 0.800 (0.783, 0.817) | 0.787 (0.762, 0.811) | 0.761 (0.712, 0.810) |
| | *ΔC*=0.054; *P*<0.001 | *ΔC*=0.050; *P*<0.001 | *ΔC*=0.036; *P*=0.065 |
| **10-year all-cause mortality in women aged 60–69 years** | **n=43,303 (2,467 deaths)** | **n=18,722 (1,084 deaths)** | **n=1,866 (240 deaths)** |
| Traditional risk factor model [a] | 0.750 (0.739, 0.761) | 0.735 (0.718, 0.753) | 0.714 (0.679, 0.749) |
| Metabolomics model [e] | 0.670 (0.658, 0.681) | 0.654 (0.637, 0.672) | 0.639 (0.605, 0.674) |
| Combined model | 0.797 (0.786, 0.809) | 0.777 (0.762, 0.793) | 0.769 (0.738, 0.799) |
| | *ΔC*=0.047; *P*<0.001 | *ΔC*=0.042; *P*<0.001 | *ΔC*=0.055; *P*<0.001 |
| **10-year all-cause mortality in total study population** | **n=146,400 (8,631 deaths)** | **n=62,744 (3,716 deaths)** | **n=6,820 (804 deaths)** |
| Traditional risk factor model [a] | 0.757 (0.751, 0.763) | 0.723 (0.715, 0.732) | 0.720 (0.700, 0.741) |
| Metabolomics model [f] | 0.690 (0.684, 0.696) | 0.679 (0.670, 0.688) | 0.662 (0.645, 0.679) |
| Combined model | 0.833 (0.827, 0.839) | 0.796 (0.788, 0.804) | 0.788 (0.770, 0.808) |
| | *ΔC*=0.076; *P*<0.001 | *ΔC*=0.073; *P*<0.001 | *ΔC*=0.068; *P*<0.001 |
| **5-year all-cause mortality in total study population** | **n=146,400 (2,944 deaths)** | **n=62,744 (1,317 deaths)** | **n=6,820 (312 deaths)** |
| Traditional risk factor model [a] | 0.761 (0.749, 0.773) | 0.731 (0.717, 0.745) | 0.734 (0.705, 0.762) |
| Metabolomics model [g] | 0.701 (0.690, 0.712) | 0.685 (0.671, 0.699) | 0.673 (0.646, 0.700) |
| Combined model | 0.856 (0.847, 0.865) | 0.818 (0.805, 0.831) | 0.812 (0.788, 0.836) |
| | *ΔC*=0.095; *P*<0.001 | *ΔC*=0.087; *P*<0.001 | *ΔC*=0.078; *P*<0.001 |

[a] Variables of the traditional risk factor models are: age, sex, education level, smoking status, physical activity, alcohol consumption, body mass index, treated dyslipidemia, hypertension, diabetes, cardiovascular disease, and cancer.

[b] Metabolites included in this model were (see **Appendix, Table A2** for abbreviations): GlycA, XXL-VLDL-PL-pct, XL-HDL-FC, Omega-6/Omega-3, Tyr, Glucose, Acetone, HDL-size, Citrate, Lactate, Creatinine, bOHbutyrate, Acetate, L-LDL-CE-pct, His, IDL-CE-pct, Albumin, LA-pct, VLDL-size, Val.

[c] Metabolites included in this model were (see **Appendix, Table A2** for abbreviations): GlycA, Omega-6/Omega-3, Tyr, Glucose, Acetone, XL-HDL-FC, Citrate, Acetate, Creatinine, bOHbutyrate, S-LDL-CE, IDL-CE-pct, His, Albumin, S-HDL-CE, LA-pct, VLDL-size, Val.

[d] Metabolites included in this model were (see **Appendix, Table A2** for abbreviations): GlycA, Omega-6/Omega-3, M-LDL-TG-pct, Creatinine, Acetoacetate, Glucose, His, Val, Leu, Albumin, VLDL-size, LA-pct.

[e] Metabolites included in this model were (see **Appendix, Table A2** for abbreviations): GlycA, Omega-6/Omega-3, Creatinine, Glucose, M-LDL-TG-pct, Acetoacetate, PUFA, His, Albumin, S-HDL-CE, Val, VLDL-size, LA-pct.

[f/g] Metabolites included in this model were (see **Appendix, Table A2** for abbreviations): GlycA, XXL-VLDL-PL-pct, XL-HDL-FC, Omega-6/Omega-3, Tyr, Glucose, Acetone, HDL-size, Citrate, Lactate, Creatinine, bOHbutyrate, Acetate, L-LDL-CE-pct, S-LDL-CE, IDL-CE-pct, M-LDL-TG-pct, Acetoacetate, PUFA, His, Val, Leu, Albumin, S-HDL-CE, LA-pct, VLDL-size.
[h] The C-statistic of the combined model was tested against the C-statistic of the traditional risk factor model with the DeLong test.

**Table 13. Reclassification improvements in 10- and 5-year mortality prediction by adding metabolomic biomarkers to conventional risk models: evidence from the UK Biobank and ESTHER study**

| | Categorial NRI [f] | Continuous NRI | IDI |
|---|---|---|---|
| **10-year all-cause mortality in men aged 50–59 years [a]** | | | |
| UK Biobank (70%) | 17.5±5.3%, *P*<0.001 | 37.5±10.9%, *P*<0.001 | 10.7±2.9%, *P*<0.001 |
| UK Biobank (30%) | 12.6±4.3%, *P*<0.001 | 32.0±9.1%, *P*<0.001 | 9.9±2.7%, *P*<0.001 |
| ESTHER Study | 18.9±5.9%, *P*<0.001 | 41.3±11.1%, *P*<0.001 | 12.4±3.2%, *P*<0.001 |
| **10-year all-cause mortality in men aged 60–69 years [b]** | | | |
| UK Biobank (70%) | 15.0±4.7%, *P*<0.001 | 30.1±7.5%, *P*<0.001 | 6.1±2.0%, *P*<0.001 |
| UK Biobank (30%) | 9.4±3.4%, *P*<0.001 | 26.5±6.8%, *P*<0.001 | 5.0±1.6%, *P*<0.001 |
| ESTHER Study | 8.2±2.5%, *P*<0.001 | 25.8±5.0%, *P*<0.001 | 4.6±1.0%, *P*<0.001 |
| **10-year all-cause mortality in women aged 50–59 years [c]** | | | |
| UK Biobank (70%) | 16.3±5.1%, *P*<0.001 | 31.4±8.1%, *P*<0.001 | 6.4±2.3%, *P*<0.001 |
| UK Biobank (30%) | 10.7±2.5%, *P*<0.001 | 29.7±7.9%, *P*<0.001 | 5.8±1.7%, *P*<0.001 |
| ESTHER Study | -0.2±0.4%, *P*=0.146 | 1.2±0.8%, *P*=0.045 | -0.05±0.08%, *P*=0.086 |
| **10-year all-cause mortality in women aged 60–69 years [d]** | | | |
| UK Biobank (70%) | 11.5±3.9%, *P*<0.001 | 26.7±7.4%, *P*<0.001 | 5.9±1.7%, *P*<0.001 |
| UK Biobank (30%) | 9.1±3.0%, *P*<0.001 | 24.4±5.1%, *P*<0.001 | 4.8±1.2%, *P*<0.001 |
| ESTHER Study | 14.5±4.3%, *P*<0.001 | 31.7±8.3%, *P*<0.001 | 6.2±1.8%, *P*<0.001 |
| **10-year all-cause mortality in total study population [e]** | | | |
| UK Biobank (70%) | 18.3±5.5%, *P*<0.001 | 41.6±11.8%, *P*<0.001 | 11.7±3.4%, *P*<0.001 |
| UK Biobank (30%) | 16.1±4.2%, *P*<0.001 | 38.2±8.4%, *P*<0.001 | 9.5±2.5%, *P*<0.001 |
| ESTHER Study | 14.7±3.8%, *P*<0.001 | 30.4±7.7%, *P*<0.001 | 8.3±1.6%, *P*<0.001 |
| **5-year all-cause mortality in total study population [e]** | | | |
| UK Biobank (70%) | 24.7±6.8%, *P*<0.001 | 42.2±12.4%, *P*<0.001 | 12.3±3.6%, *P*<0.001 |
| UK Biobank (30%) | 20.3±5.5%, *P*<0.001 | 39.7±10.1%, *P*<0.001 | 10.7±3.1%, *P*<0.001 |
| ESTHER Study | 19.1±4.7%, *P*<0.001 | 33.4±9.2%, *P*<0.001 | 9.8±2.2%, *P*<0.001 |

Notes: Data are expressed as point estimates with 95% confidence intervals, presented as estimate ± half the confidence interval width. Both the conventional risk factor model and the extended model incorporating metabolomic biomarkers were developed in the derivation cohort (70% of UK Biobank) and validated in the internal validation cohort (30% of UK Biobank) and the external ESTHER study. The conventional risk factor model, included age, sex, education level, smoking status, physical activity, alcohol consumption, body mass index, treated dyslipidemia, hypertension, diabetes, cardiovascular disease, and cancer.
[a] 20 metabolic biomarkers, [b] 18 metabolic biomarkers, [c] 12 metabolic biomarkers, [d] 13 metabolic biomarkers were used in the younger men, older men, younger women and older women groups, respectively.
[e] The overlap of 26 unique biomarkers across four subgroups were used in the total study populations.
[f] Categorical reclassification was assessed for three risk categories (≤ 5%, >5%–10%, >10%) based on the conventional risk factors and the biomarker score. IDI, integrated discrimination improvement; NRI, net reclassification improvement.

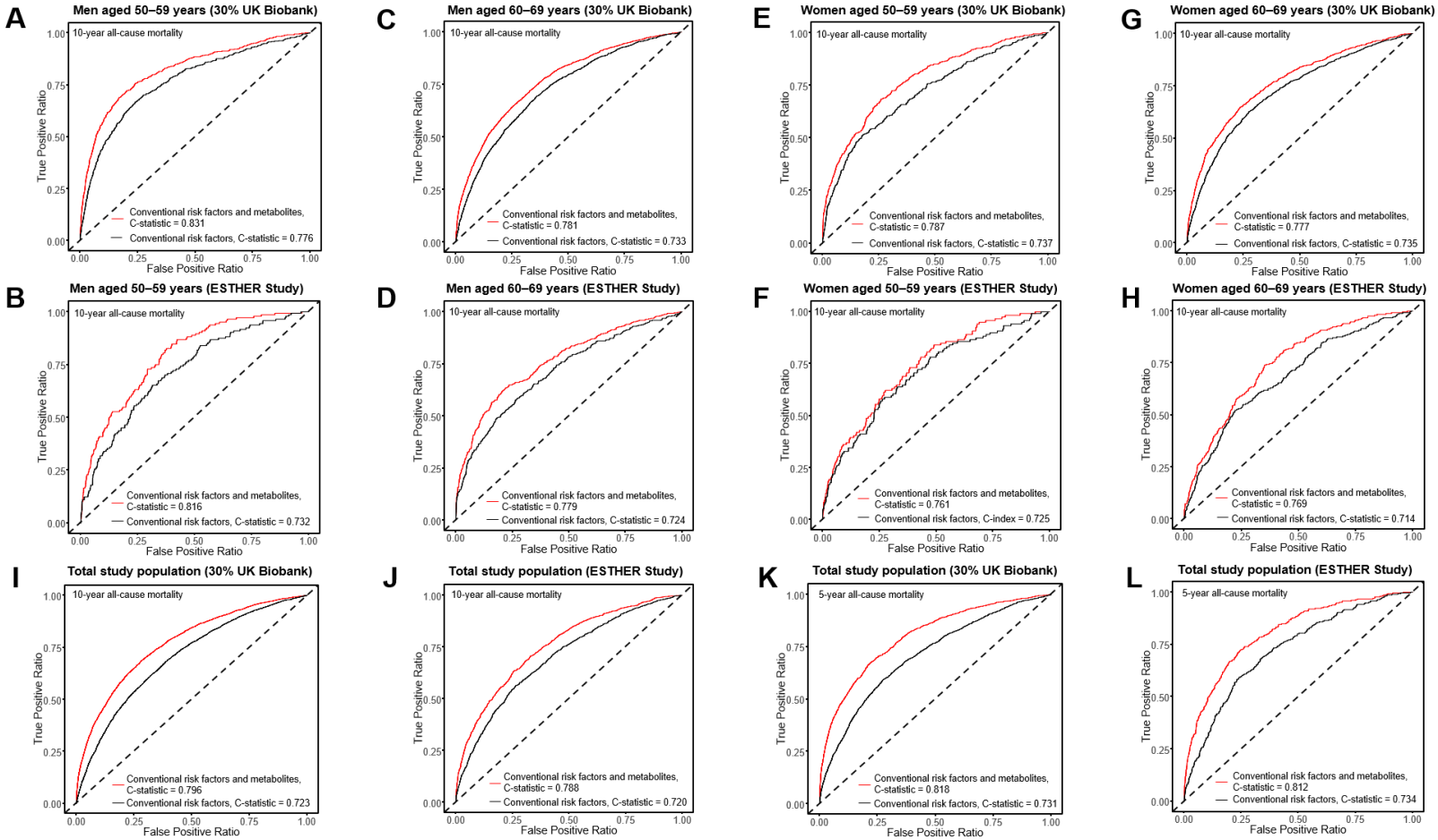



**Figure 22. Improved mortality prediction accuracy with identified metabolomic biomarkers in sex- and age-specific groups, as well as the total study population**

For 10-year all-cause mortality, ROC curves were assessed in: (**A**) younger men (30% UK Biobank), (**B**) younger men (ESTHER), (**C**) older men (30% UK Biobank), (**D**) older men (ESTHER), (**E**) younger women (30% UK Biobank), (**F**) younger women (ESTHER), (**G**) older women (30% UK Biobank), (**H**) older women (ESTHER), (**I**) the total study population (30% UK Biobank), and (**J**) the total study population (ESTHER). For 5-year all-cause mortality, ROC curves were assessed in (**K**) the total study population from 30% of UK Biobank and (**L**) the total study population from ESTHER. The curves are based on the predictions from the conventional risk factors (black) and the metabolic biomarkers (red).

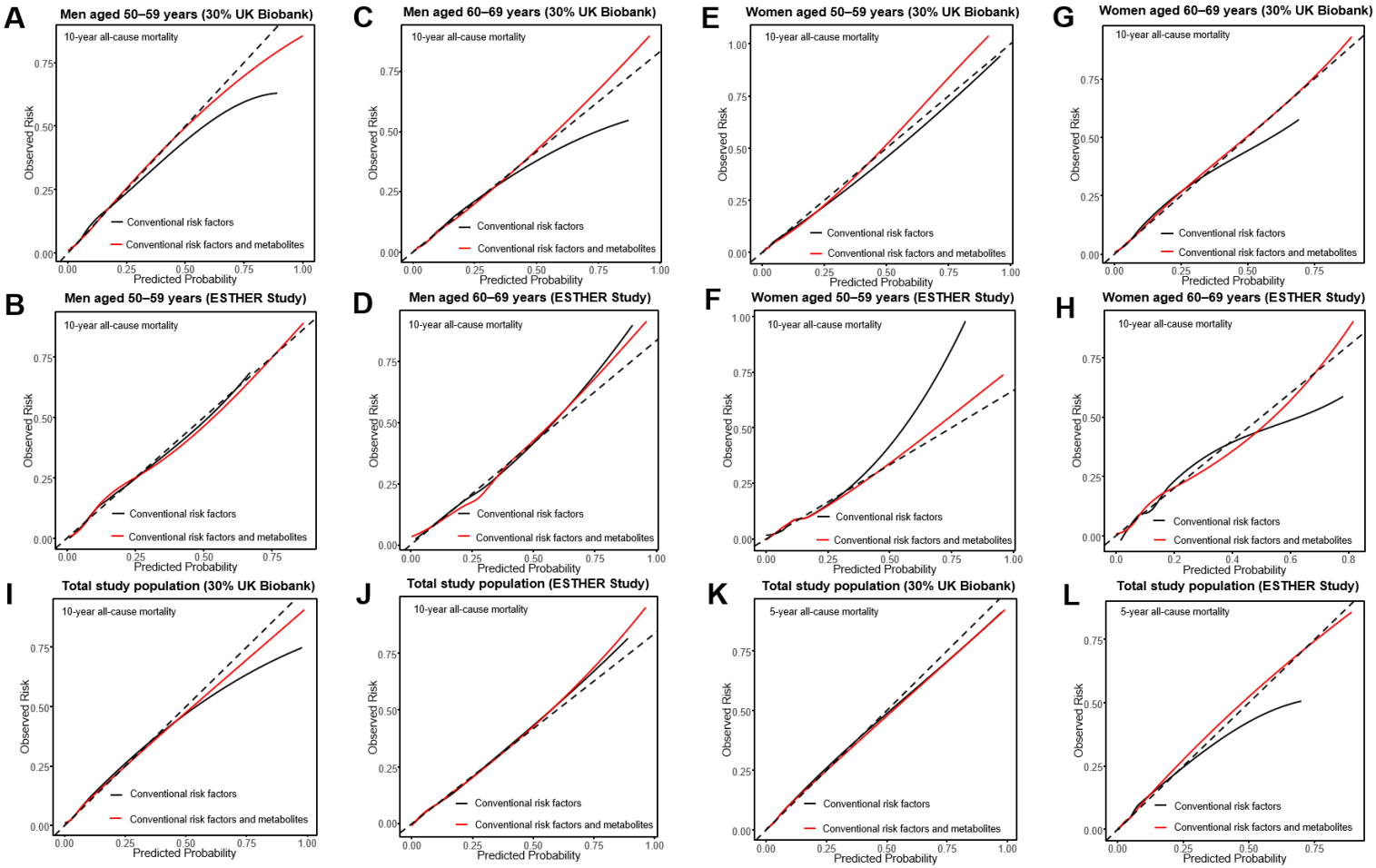



**Figure 23. Calibration curves of the models only including conventional risk factors and their combinations with the metabolites for risk prediction in the internal and external validation cohorts**

For 10-year all-cause mortality, calibration was assessed in: (**A**) younger men (30% UK Biobank), (**B**) younger men (ESTHER), (**C**) older men (30% UK Biobank), (**D**) older men (ESTHER), (**E**) younger women (30% UK Biobank), (**F**) younger women (ESTHER), (**G**) older women (30% UK Biobank), (**H**) older women (ESTHER), (**I**) the total study population (30% UK Biobank), and (**J**) the total study population (ESTHER). For 5-year all-cause mortality, calibration was assessed in (**K**) the total study population from 30% of UK Biobank and (**L**) the total study population from ESTHER.

### 3.3.6 Development and testing of metabolomic ageing clocks

The equations used to calculate the MetaboMR clock1 and MetaboMR clock2 are provided in **Appendix, Table A8**. The distributions of chronological age and the biological age according to the two clocks in the training, internal validation, and external validation datasets are illustrated in **Figure 24A–F**. **Figure 25** shows strong correlations between the MetaboMR clock1, which was derived from the selected metabolomic biomarkers only, and chronological age across all sub-groups and in the overall study population (training set: $r = 0.83–0.85$; internal validation set: $r = 0.82–0.85$; external validation set: $r = 0.80–0.83$). The corresponding Pearson correlation coefficients for the MetabMR clock2, which was derived from the biomarkers and traditional mortality risk factors, were slightly higher (**Figure 26**; training set: $r = 0.85–0.86$; internal validation set: $r = 0.84–0.86$; external validation set: $r = 0.80–0.84$). The C-statistics for 10-year mortality prediction using the MetaboMR clocks were substantially higher than the C-statistics of chronological age in the total population of both studies and in all age and sex-specific subgroups, and MetaboMR clock2 outperformed MetaboMR clock1 (**Figure 27**). For instance, in the external validation cohort (ESTHER), the MetaboMR clock2 achieved a C-statistic of 0.781 (0.766–0.796) in the total population, followed by a C-statistic of 0.682 (0.665–0.698) of MetaboMR clock1, whereas the C-statistic of the chronological age was 0.616 (0.601–0.631).

The distributions of metabolomic age acceleration derived from these clocks are shown in **Figure 28**. The metabolomic age acceleration derived from the MetaboMR clock1 (MetAA1) and MetaboMR clock2 (MetAA2) were consistently, significantly associated with 10-year all-cause mortality across all sub-groups and the total population in both the internal and external validation sets with stronger associations among men than women (**Table 14**). Overall, the results for MetAA1 and MetAA2 were comparable. In the total internal validation set (UKB) and total external validation set (ESTHER), each year increase of MetAA1 was significantly associated with a 10-year all-cause mortality increase in the adjusted model of 9% (HR [95%CI], 1.09 [1.08–1.10]) and 8% (HR [95%CI], 1.08 [1.06–1.09]), respectively. Excluding subjects on lipid-lowering therapy led to almost the same results (**Table 15**).

MetAAs were slightly stronger associated with cardiovascular disease (CVD) mortality than with cancer mortality in the total UKB cohort (**Table 16**). Each 1-year increase in MetAA1 and MetAA2 was associated with 13% and 14% higher risks of CVD mortality, respectively, whereas MetAA1 and MetAA2 were associated with a 9% and 10% higher risk of cancer mortality, respectively. Cause-specific mortality results were not shown for ESTHER due to low case numbers.

The algorithms for the MetaboMR clocks have been made publicly available on GitHub (https://github.com/leipeng0825/metabolomics_10y_all_cause_mortality) to facilitate independent validation, and an interactive web calculator has been developed to enable clinical implementation (https://metabomrclock.shinyapps.io/metabomrclock/).

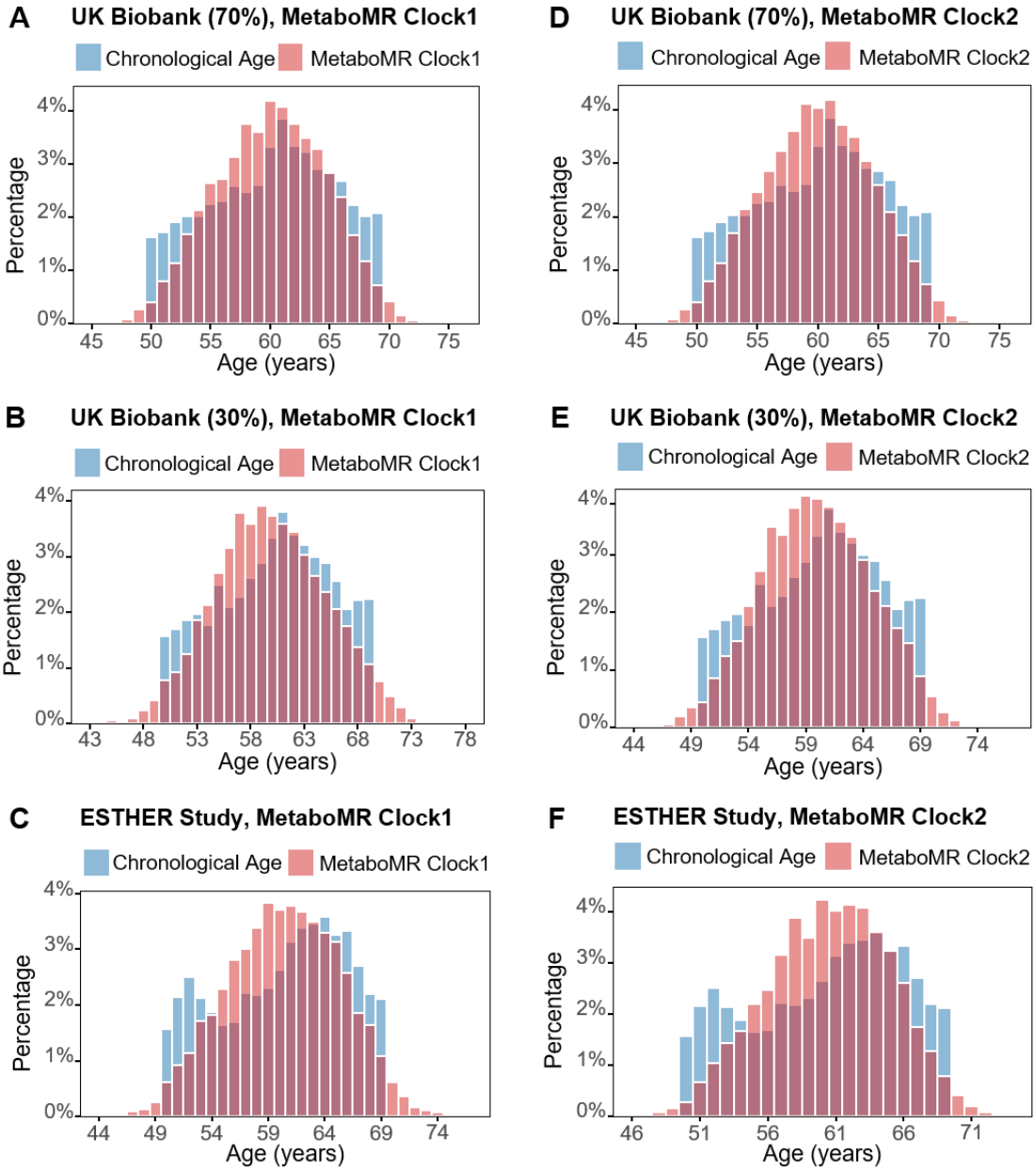



**Figure 24. Distributions of chronological age and metabolomics-based mortality-related ageing clocks in the UK Biobank and ESTHER cohorts**

(**A–C**) MetaboMR clock1 (based on LASSO-selected metabolomic biomarkers only); (**D–F**) MetaboMR clock2 (based on LASSO-selected metabolomic biomarkers + conventional risk factors).
Abbreviations: LASSO, least absolute shrinkage and selection operator; MetaboMR, metabolomics-based mortality risk.

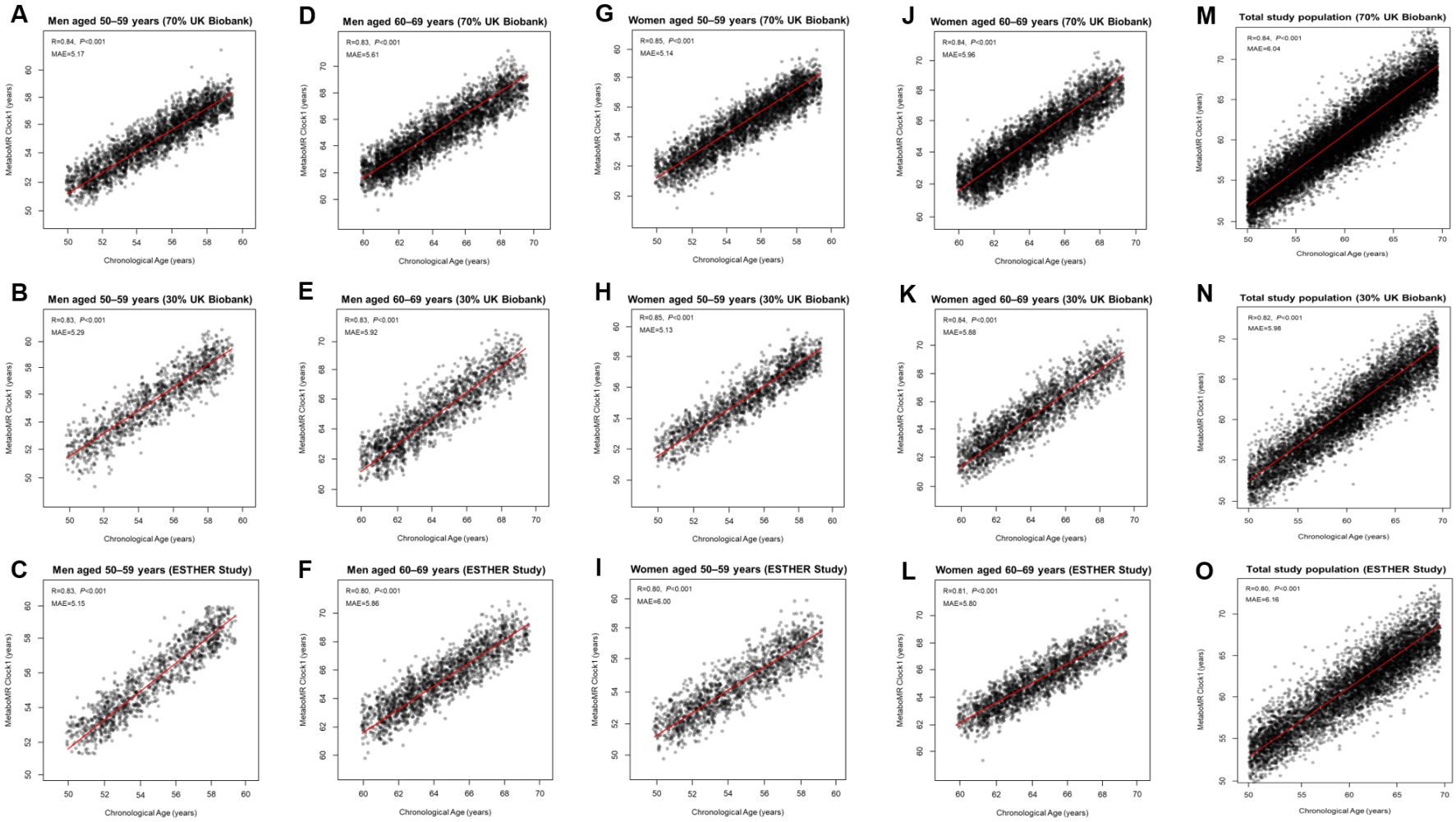



**Figure 25. Linear correlation between chronological age and the metabolomic age based on 26 selected metabolomic biomarkers only (metabolomics-based mortality risk clock 1 (MetaboMR clock1))**

(**A–C**) younger men; (**D–F**) older men; (**G–I**) younger women; (**J–L**) older women; (**M–O**) total population. (**A**, **D**, **G**, **J**, **M**) 70% UK Biobank; (**B**, **E**, **H**, **K**, **N**) 30% UK Biobank; (**C**, **F**, **I**, **L**, **O**) ESTHER. The Pearson correlation coefficient (*r*) is shown in the top left corner of the panels of the figure. In addition, the mean absolute error (MAE) is stated, which reflects the average absolute deviation between predicted metabolomic age and chronological age. Lower MAE values indicate greater model prediction accuracy.

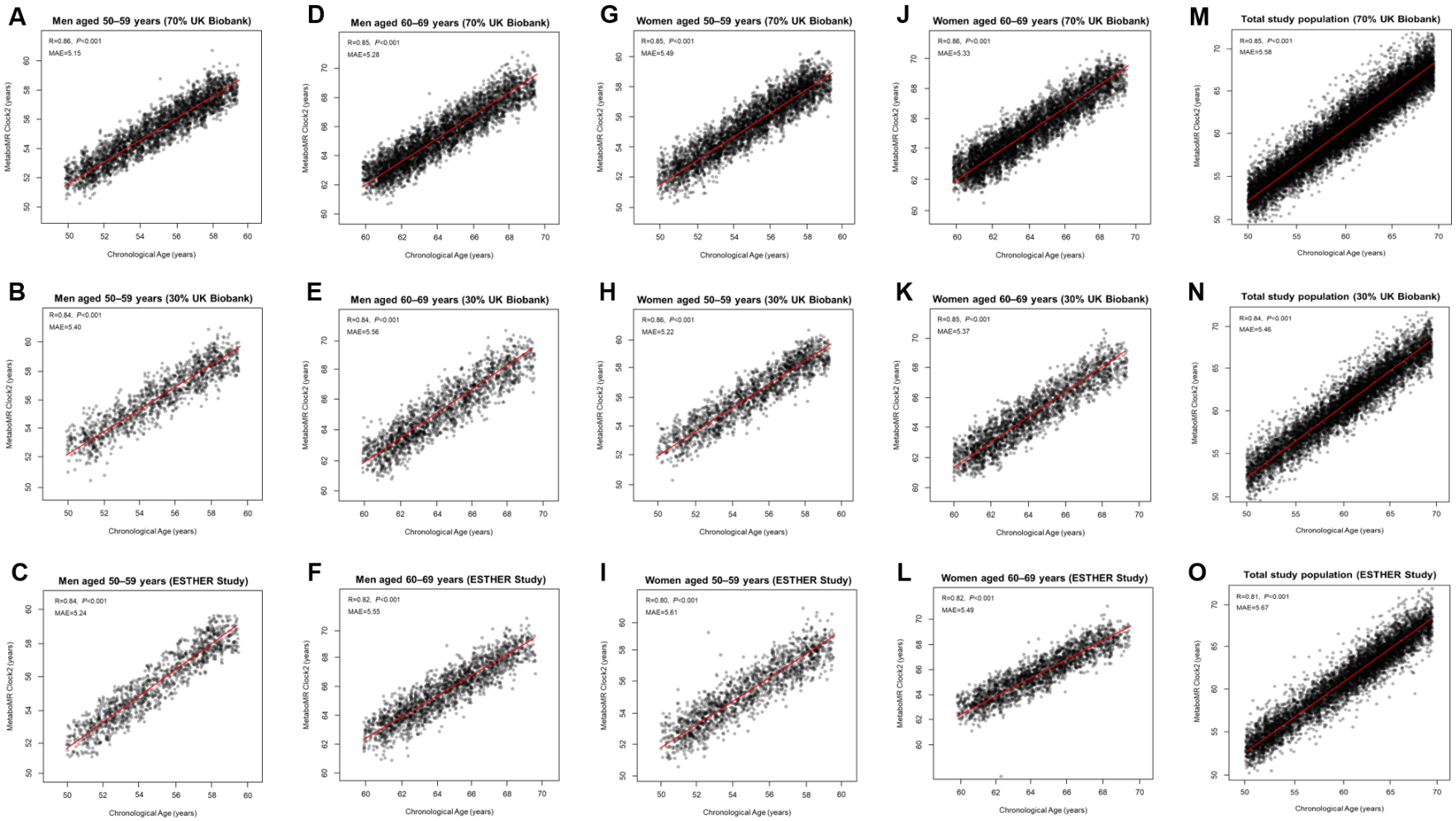



**Figure 26. Linear correlation between chronological age and metabolomic age based on 26 selected metabolomic biomarkers and traditional mortality risk factors (metabolomics-based mortality risk clock 2 (MetaboMR clock2))**

(**A–C**) younger men; (**D–F**) older men; (**G–I**) younger women; (**J–L**) older women; (**M–O**) total population. (**A**, **D**, **G**, **J**, **M**) 70% UK Biobank; (**B**, **E**, **H**, **K**, **N**) 30% UK Biobank; (**C**, **F**, **I**, **L**, **O**) ESTHER. The Pearson correlation coefficient (*r*) is shown in the top left corner of the panels of the figure. In addition, the mean absolute error (MAE) is stated, which reflects the average absolute deviation between predicted metabolomic age and chronological age. Lower MAE values indicate greater model prediction accuracy.

**Table 14. Association of metabolomic age acceleration based on metabolomic biomarkers only (MetAA1) and based on both biomarkers and traditional risk factors (MetAA2) with 10-year all-cause mortality**

| MetAA / Study population | Training set UK Biobank (70%) | | | Internal validation set UK Biobank (30%) | | | External validation set ESTHER | | |
|---|---|---|---|---|---|---|---|---|---|
| | Mean MetAA ± SD (years) | Crude HR per 1-year MetAA increase [a] | Adjusted HR per 1-year MetAA increase [b] | Mean MetAA ± SD (years) | Crude HR per 1-year MetAA increase [a] | Adjusted HR per 1-year MetAA increase [b] | Mean MetAA ± SD (years) | Crude HR per 1-year MetAA increase [a] | Adjusted HR per 1-year MetAA increase [b] |
| **MetAA1** | | | | | | | | | |
| **Men aged 50–59 years** | 0.00±5.71 | **1.16 (1.12–1.20)** | **1.13 (1.09–1.17)** | 0.27±5.94 | **1.20 (1.14–1.26)** | **1.17 (1.11–1.23)** | 0.40±5.76 | **1.19 (1.08–1.32)** | **1.16 (1.04–1.29)** |
| **Men aged 60–69 years** | 0.00±5.82 | **1.20 (1.18–1.22)** | **1.17 (1.15–1.19)** | 0.30±6.09 | **1.19 (1.17–1.23)** | **1.17 (1.14–1.20)** | 0.51±5.95 | **1.20 (1.14–1.26)** | **1.19 (1.13–1.26)** |
| **Women aged 50–59 years** | 0.00±5.73 | **1.13 (1.09–1.17)** | **1.10 (1.05–1.14)** | 0.17±5.66 | **1.12 (1.06–1.19)** | **1.10 (1.04–1.16)** | 0.47±6.14 | **1.11 (1.01–1.23)** | **1.11 (1.00–1.23)** |
| **Women aged 60–69 years** | 0.00±5.77 | **1.18 (1.16–1.21)** | **1.16 (1.14–1.19)** | 0.22±5.88 | **1.19 (1.16–1.23)** | **1.17 (1.13–1.21)** | 0.53±5.77 | **1.13 (1.05–1.20)** | **1.10 (1.03–1.18)** |
| **Total study population** | 0.00±5.81 | **1.12 (1.11–1.13)** | **1.10 (1.09–1.11)** | 0.24±5.62 | **1.11 (1.10–1.12)** | **1.09 (1.08–1.10)** | 0.48±6.08 | **1.10 (1.09–1.12)** | **1.08 (1.06–1.09)** |
| **MetAA2** | | | | | | | | | |
| **Men aged 50–59 years** | 0.00±5.42 | **1.14 (1.10–1.18)** | N.A.[c] | 0.16±5.73 | **1.17 (1.11–1.23)** | N.A.[c] | 0.34±5.62 | **1.15 (1.04–1.28)** | N.A.[c] |
| **Men aged 60–69 years** | 0.00±5.58 | **1.19 (1.17–1.21)** | N.A.[c] | 0.23±5.84 | **1.18 (1.16–1.22)** | N.A.[c] | 0.46±5.81 | **1.17 (1.11–1.23)** | N.A.[c] |
| **Women aged 50–59 years** | 0.00±5.65 | **1.12 (1.08–1.16)** | N.A.[c] | 0.12±5.49 | **1.11 (1.05–1.18)** | N.A.[c] | 0.39±6.13 | **1.10 (1.01–1.22)** | N.A.[c] |
| **Women aged 60–69 years** | 0.00±5.70 | **1.15 (1.13–1.18)** | N.A.[c] | 0.20±5.76 | **1.16 (1.13–1.20)** | N.A.[c] | 0.42±5.97 | **1.12 (1.04–1.19)** | N.A.[c] |
| **Total study population** | 0.00±5.57 | **1.11 (1.10–1.12)** | N.A.[c] | 0.18±5.61 | **1.10 (1.09–1.11)** | N.A.[c] | 0.37±5.79 | **1.09 (1.08–1.11)** | N.A.[c] |

Abbreviations: CI, confidence interval; HR, hazard ratio; MetAA1, metabolomic age acceleration.

Notes: Printed in bold: Statistically significant ($P$<0.05).

[a] HRs for mortality were estimated per 1-year increase in metabolomic age acceleration (MetAA1), reflecting the relative increase in mortality risk for each year that a person's metabolomic age exceeds their chronological age.

[b] The adjusted Cox proportional hazards regression model included the following covariates: chronological age, sex, education level, smoking status, physical activity, alcohol consumption, body mass index, treated dyslipidemia, hypertension, diabetes, cardiovascular disease, and cancer.

[c] As the MetaboMR clock2 already includes the traditional risk factors, it cannot be further adjusted for them.

**Table 15. Association of metabolomic age acceleration based on metabolomic biomarkers only (MetAA1) and based on both biomarkers and traditional risk factors (MetAA2) with 10-year all-cause mortality, excluding participants on lipid-lowering therapy**

| MetAA / Study population | Training set UK Biobank (70%) | | | Internal validation set UK Biobank (30%) | | | External validation set ESTHER | | |
|---|---|---|---|---|---|---|---|---|---|
| | Mean MetAA ± SD (years) | Crude HR per 1-year MetAA increase [a] | Adjusted HR per 1-year MetAA increase [b] | Mean MetAA ± SD (years) | Crude HR per 1-year MetAA increase [a] | Adjusted HR per 1-year MetAA increase [b] | Mean MetAA ± SD (years) | Crude HR per 1-year MetAA increase [a] | Adjusted HR per 1-year MetAA increase [b] |
| **MetAA1** | | | | | | | | | |
| **Men aged 50–59 years** | 0.00±5.68 | **1.17 (1.13–1.21)** | **1.14 (1.10–1.18)** | 0.23±5.92 | **1.21 (1.14–1.28)** | **1.18 (1.12–1.24)** | 0.37±5.78 | **1.20 (1.09–1.33)** | **1.17 (1.05–1.30)** |
| **Men aged 60–69 years** | 0.00±5.79 | **1.23 (1.20–1.26)** | **1.20 (1.18–1.22)** | 0.28±6.05 | **1.20 (1.17–1.23)** | **1.19 (1.16–1.22)** | 0.48±5.91 | **1.21 (1.15–1.28)** | **1.20 (1.14–1.27)** |
| **Women aged 50–59 years** | 0.00±5.71 | **1.16 (1.12–1.20)** | **1.13 (1.08–1.17)** | 0.14±5.64 | **1.15 (1.09–1.22)** | **1.11 (1.07–1.15)** | 0.44±6.12 | **1.14 (1.03–1.26)** | **1.12 (1.01–1.24)** |
| **Women aged 60–69 years** | 0.00±5.73 | **1.19 (1.16–1.22)** | **1.17 (1.14–1.20)** | 0.19±5.84 | **1.20 (1.16–1.24)** | **1.17 (1.13–1.21)** | 0.50±5.75 | **1.16 (1.08–1.23)** | **1.11 (1.04–1.19)** |
| **Total study population** | 0.00±5.78 | **1.14 (1.13–1.15)** | **1.12 (1.11–1.13)** | 0.20±5.61 | **1.13 (1.12–1.14)** | **1.11 (1.10–1.12)** | 0.45±6.06 | **1.12 (1.10–1.14)** | **1.10 (1.08–1.12)** |
| **MetAA2** | | | | | | | | | |
| **Men aged 50–59 years** | 0.00±5.39 | **1.15 (1.11–1.19)** | N.A.[c] | 0.14±5.70 | **1.18 (1.13–1.23)** | N.A.[c] | 0.31±5.60 | **1.16 (1.05–1.29)** | N.A.[c] |
| **Men aged 60–69 years** | 0.00±5.55 | **1.22 (1.20–1.24)** | N.A.[c] | 0.21±5.81 | **1.18 (1.15–1.21)** | N.A.[c] | 0.44±5.79 | **1.18 (1.12–1.24)** | N.A.[c] |
| **Women aged 50–59 years** | 0.00±5.63 | **1.15 (1.10–1.20)** | N.A.[c] | 0.10±5.47 | **1.12 (1.08–1.16)** | N.A.[c] | 0.36±6.11 | **1.13 (1.04–1.25)** | N.A.[c] |
| **Women aged 60–69 years** | 0.00±5.68 | **1.18 (1.15–1.21)** | N.A.[c] | 0.18±5.74 | **1.16 (1.13–1.19)** | N.A.[c] | 0.41±5.95 | **1.13 (1.06–1.20)** | N.A.[c] |
| **Total study population** | 0.00±5.54 | **1.13 (1.12–1.14)** | N.A.[c] | 0.16±5.59 | **1.12 (1.11–1.13)** | N.A.[c] | 0.35±5.76 | **1.10 (1.08–1.12)** | N.A.[c] |

Abbreviations: CI, confidence interval; HR, hazard ratio; MetAA1, metabolomic age acceleration.

Notes: Printed in bold: Statistically significant ($P$<0.05).

[a] HRs for mortality were estimated per 1-year increase in metabolomic age acceleration (MetAA1), reflecting the relative increase in mortality risk for each year that a person's metabolomic age exceeds their chronological age.

[b] The adjusted Cox proportional hazards regression model included the following covariates: chronological age, sex, education level, smoking status, physical activity, alcohol consumption, body mass index, hypertension, diabetes, cardiovascular disease, and cancer.

[c] As the MetaboMR clock2 already includes the traditional risk factors, it cannot be further adjusted for them.

**Table 16. Association of metabolomic age acceleration based on metabolomic biomarkers only (MetAA1) and based on both biomarkers and traditional risk factors (MetAA2) with 10-year CVD and cancer mortality in the total UK Biobank**

| MetAA / Study population | CVD mortality | | | | Cancer mortality | | | |
|---|---|---|---|---|---|---|---|---|
| | $N_{total}$ | $N_{cases}$ | Crude HR per 1-year MetAA increase [a] | Adjusted HR per 1-year MetAA increase [b] | $N_{total}$ | $N_{cases}$ | Crude HR per 1-year MetAA increase [a] | Adjusted HR per 1-year MetAA increase [b] |
| **MetAA1** | | | | | | | | |
| **Men aged 50–59 years** | 40,201 | 431 | **1.30 (1.27–1.34)** | **1.13 (1.09–1.16)** | 40,201 | 870 | **1.19 (1.17–1.21)** | **1.11 (1.09–1.13)** |
| **Men aged 60–69 years** | 56,111 | 1,426 | **1.37 (1.35–1.40)** | **1.18 (1.16–1.20)** | 56,111 | 2,831 | **1.22 (1.20–1.24)** | **1.13 (1.12–1.14)** |
| **Women aged 50–59 years** | 50,807 | 146 | **1.34 (1.29–1.39)** | **1.15 (1.11–1.20)** | 50,807 | 878 | **1.17 (1.14–1.20)** | **1.09 (1.07–1.11)** |
| **Women aged 60–69 years** | 62,025 | 578 | **1.45 (1.41–1.49)** | **1.21 (1.18–1.25)** | 62,025 | 2,135 | **1.21 (1.19–1.23)** | **1.12 (1.10–1.14)** |
| **Total study population** | 209,144 | 2,581 | **1.30 (1.29–1.31)** | **1.13 (1.12–1.14)** | 209,144 | 6,714 | **1.17 (1.15–1.19)** | **1.09 (1.08–1.10)** |
| **MetAA2** | | | | | | | | |
| **Men aged 50–59 years** | 40,201 | 431 | **1.15 (1.11–1.18)** | N.A.[c] | 40,201 | 870 | **1.12 (1.10–1.14)** | N.A.[c] |
| **Men aged 60–69 years** | 56,111 | 1,426 | **1.20 (1.18–1.22)** | N.A.[c] | 56,111 | 2,831 | **1.15 (1.14–1.17)** | N.A.[c] |
| **Women aged 50–59 years** | 50,807 | 146 | **1.16 (1.12–1.21)** | N.A.[c] | 50,807 | 878 | **1.10 (1.08–1.12)** | N.A.[c] |
| **Women aged 60–69 years** | 62,025 | 578 | **1.23 (1.20–1.27)** | N.A.[c] | 62,025 | 2,135 | **1.13 (1.11–1.15)** | N.A.[c] |
| **Total study population** | 209,144 | 2,581 | **1.14 (1.13–1.15)** | N.A.[c] | 209,144 | 6,714 | **1.10 (1.09–1.11)** | N.A.[c] |

Abbreviations: CI, confidence interval; CVD, cardiovascular disease; HR, hazard ratio; MetAA1, metabolomic age acceleration.

Notes: Printed in bold: Statistically significant ($P$<0.05).

[a] HRs for mortality were estimated per 1-year increase in metabolomic age acceleration (MetAA1), reflecting the relative increase in mortality risk for each year that a person's metabolomic age exceeds their chronological age.

[b] The adjusted Cox proportional hazards regression model included the following covariates: chronological age, sex, education level, smoking status, physical activity, alcohol consumption, body mass index, treated dyslipidemia, hypertension, diabetes, cardiovascular disease, and cancer.

[c] As the MetaboMR clock2 already includes the traditional risk factors, it cannot be further adjusted for them.

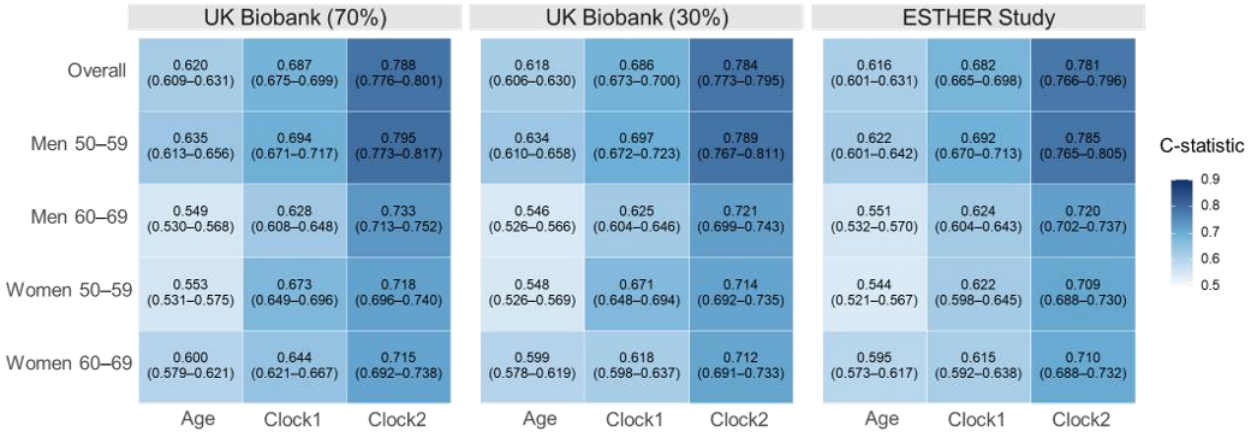



**Figure 27. Comparison of the C-statistics of chronological age and the MetaboMR clocks for 10-year mortality prediction in the UK Biobank and ESTHER study**

The figure shows C-statistics with 95% confidence intervals.

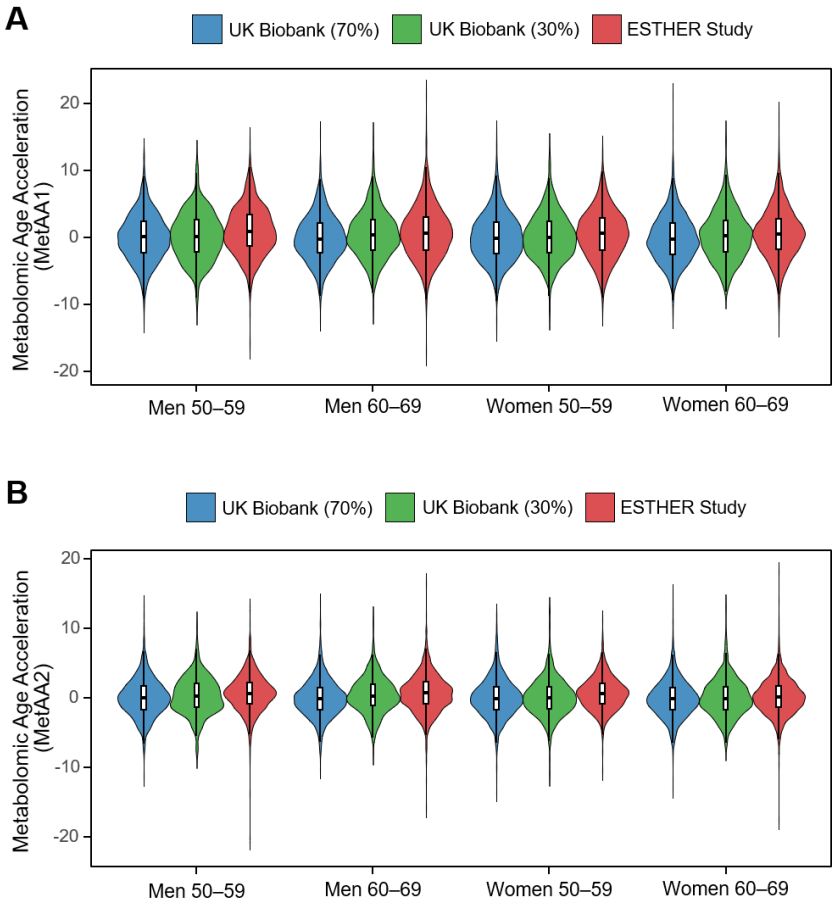



**Figure 28. Distribution of metabolomic age acceleration (MetAA) derived from MetaboMR clocks in the UK Biobank and ESTHER cohorts**

(**A**), MetAA derived from MetaboMR clock1. (**B**), MetAA derived from MetaboMR clock2. Violin plots show the distribution of MetAA in each cohort. The center line, box limits, and whiskers represent the median, interquartile range, and full range, respectively. Abbreviations: MetAA, metabolomic age acceleration; MetaboMR, metabolomics-based mortality risk.

## 3.4 Study 4: Inflammaging and proteomic ageing clocks for mortality

### 3.4.1 Baseline characteristics and conventional mortality risk factors

Baseline characteristics were broadly comparable between UKB and ESTHER (**Table 17**), although ESTHER participants were slightly older and had higher prevalences of current smoking, hypertension, diabetes, and cardiovascular disease. In contrast, ESTHER participants consumed less alcohol and were less frequently diagnosed with dyslipidemia.

Associations of conventional risk factors with 10-year all-cause mortality were largely consistent across cohorts (**Table 18**). Older age, current smoking, diabetes, cardiovascular disease, and a history of cancer were associated with substantially increased mortality risk, whereas higher levels of physical activity were associated with lower mortality risk. The direction and magnitude of these associations were broadly similar in UKB and ESTHER.

**Table 17. Baseline characteristics of study participants in the UK Biobank and ESTHER Study**

| | UK Biobank [h] | | | ESTHER Study [h] | |
|---|---|---|---|---|---|
| | Derivation Set (n=24,396) | Internal Test Set (n=10,456) | Overall (n=34,852) | External Validation Set 1 (n=8,082) | External Validation Set 2 (n=408) |
| **Age, years** | 60.1 (5.5) | 60.1 (5.4) | 60.1 (5.4) | 62.1 (6.5) | 62.3 (6.7) |
| **Women, n (%)** | 13,330 (54.6%) | 5,572 (53.3%) | 18,902 (54.2%) | 4,456 (55.1%) | 201 (49.3%) |
| **High education level, n (%)** [a, g] | 6,752 (27.7%) | 2,902 (27.8%) | 9,654 (27.7%) | 1,978 (24.5%) | 100 (24.5%) |
| **Smoking status, n (%)** [g] | / | / | / | / | / |
| Never smoker | 12,866 (52.8%) | 5,535 (52.9%) | 18,401 (52.8%) | 4,045 (50.0%) | 197 (48.2%) |
| Former smoker | 9,160 (37.5%) | 3,916 (37.5%) | 13,076 (37.5%) | 2,625 (32.5%) | 150 (36.8%) |
| Current smoker | 2,370 (9.7%) | 1,005 (9.6%) | 3,375 (9.7%) | 1,412 (17.5%) | 61 (15.0%) |
| **Physical activity, n (%)** [g] | / | / | / | / | / |
| Low | 4,811 (19.7%) | 2,019 (19.3%) | 6,830 (19.6%) | 1,730 (21.4%) | 87 (21.3%) |
| Moderate | 11,283 (46.2%) | 4,717 (45.1%) | 16,000 (45.9%) | 3,701 (45.8%) | 197 (48.3%) |
| High | 8,302 (34.1%) | 3,720 (35.6%) | 12,022 (34.5%) | 2,651 (32.8%) | 124 (30.4%) |
| **Alcohol consumption, n (%)** [b, g] | / | / | / | / | / |
| Abstainer | 7,404 (30.3%) | 3,109 (29.7%) | 10,513 (30.2%) | 2,542 (31.5%) | 135 (33.1%) |
| Low | 9,740 (40.0%) | 4,301 (41.1%) | 14,041 (40.3%) | 4,955 (61.3%) | 249 (61.0%) |
| Medium or high | 7,252 (29.7%) | 3,046 (29.2%) | 10,298 (29.5%) | 585 (7.2%) | 24 (5.9%) |
| **Body mass index, n (%)** [c, g] | / | / | / | / | / |
| $< 21.5$ kg/m$^2$ | 1,649 (6.8%) | 653 (6.2%) | 2,302 (6.6%) | 420 (5.2%) | 21 (5.1%) |
| $\geq 21.5 - < 25$ kg/m$^2$ | 6,042 (24.8%) | 2,612 (25.0%) | 8,654 (24.8%) | 1,818 (22.5%) | 96 (23.5%) |
| $\geq 25 - < 30$ kg/m$^2$ | 10,739 (44.0%) | 4,532 (43.4%) | 15,271 (43.8%) | 3,765 (46.6%) | 185 (45.4%) |
| $\geq 30 - < 35$ kg/m$^2$ | 4,295 (17.6%) | 1,938 (18.5%) | 6,233 (17.9%) | 1,609 (19.9%) | 71 (17.4%) |
| $\geq 35$ kg/m$^2$ | 1,671 (6.8%) | 721 (6.9%) | 2,392 (6.9%) | 470 (5.8%) | 35 (8.6%) |
| **Medical conditions, n (%)** | / | / | / | / | / |
| Dyslipidemia [g] | 5,216 (21.4%) | 2,175 (20.8%) | 7,391 (21.2%) | 916 (11.3%) | 47 (11.5%) |
| Hypertension [d, g] | 14,244 (58.4%) | 6,067 (58.0%) | 20,311 (58.3%) | 5,140 (63.6%) | 267 (65.4%) |
| Diabetes [e, g] | 2,496 (10.2%) | 1,060 (10.1%) | 3,556 (10.2%) | 1,198 (14.8%) | 73 (17.9%) |
| Cardiovascular disease [f, g] | 1,854 (7.6%) | 773 (7.4%) | 2,627 (7.5%) | 700 (8.7%) | 50 (12.3%) |
| History of cancer [g] | 2,144 (8.8%) | 907 (8.7%) | 3,051 (8.8%) | 623 (7.7%) | 26 (6.4%) |

Notes: Values are means (SDs) for continuous variables and percentages for categorical variables.

[a] In the UK Biobank, participants who reported "None of the above," "Certificate of Secondary Education (CSEs) or equivalent," "Ordinary Levels/General Certificate of Secondary Education (O levels/GCSEs) or equivalent," "Other professional qualifications (e.g., nursing, teaching)," or "Advanced Levels/Advanced Subsidiary Levels (A levels/AS levels) or equivalent" were categorized as having a low or medium education level. Those who reported "National Vocational Qualification (NVQ), Higher National Diploma (HND), Higher National Certificate (HNC) or equivalent," or a "College or University degree" were categorized as having a high education level. In the German ESTHER study, participants who reported "No education or 'Hauptschule'" or "'Realschule'/'Mittlere Reife'" were categorized as having a low education level, whereas those who reported "'Fachhochschulreife' or 'Abitur'" were categorized as having a high education level.

[b] Alcohol consumption was categorized as low for women consuming 0–19.99 g/day and men consuming 0–39.99 g/day. It was categorized as medium or high for women consuming 20–39.99 g/day or ≥40 g/day, and for men consuming 40–59.99 g/day or ≥60 g/day.

[c] To enhance comparability between the two cohorts, body mass index (BMI) categories were defined according to the WHO classification with slight modifications to the cut-off values. Participants with a BMI less than 21.5 kg/m², including both underweight individuals and those at the lower end of the normal range, were grouped together to account for population-specific differences and to ensure sufficient sample sizes across categories.

[d] Hypertension was defined as self-reported hypertension, the use of antihypertensive medication, a systolic blood pressure ≥140 mm Hg, or a diastolic blood pressure ≥90 mm Hg.

[e] In this study, diabetes was defined as self-reported diabetes, the use of glucose-lowering medication, or an $HbA_{1c}$ level of ≥6.5%.

[f] Cardiovascular disease in this study was defined as having myocardial infarction or stroke.

[g] Missing data were present in the UK Biobank for education level (1.7%), smoking status (0.1%), physical activity (8.4%), alcohol consumption (2.0%), BMI (1.2%), dyslipidemia (0.02%), hypertension (0.02%), diabetes (0.01%), cardiovascular disease (0.4%), and cancer (0.2%). Corresponding proportions for the same variables in the ESTHER study were 2.6%, 2.8%, 0.4%, 8.8%, 1.3%, 1.4%, 1.9%, 1.3%, 3.0%, and 2.7% for external validation set 1, and 2.4%, 2.5%, 0.5%, 7.2%, 1.0%, 1.3%, 1.5%, 1.1%, 1.9%, and 1.2% for external validation set 2. All missing values were imputed using random forest methods as implemented in the R package *missRanger* (version 2.1.5).

[h] Median follow-up times were 9.76, 9.50, and 9.49 years in the UK Biobank, external validation subset 1, and external validation subset 2, respectively. Numbers of deaths during 10-year follow-up were 1,983, 958, and 100.

**Table 18. Longitudinal association between traditional risk factors and 10-year all-cause mortality in the UK Biobank and ESTHER study**

| Baseline Characteristics | UK Biobank (34,852 participants) | | | Validation set 1 (8,082 participants) [c] | | | Validation set 2 (408 participants) [c] | | |
|---|---|---|---|---|---|---|---|---|---|
| | $N_{death}$ (%) | HR (95% CI) [a] | FDR [b] | $N_{death}$ (%) | HR (95% CI) [a] | FDR [b] | $N_{death}$ (%) | HR (95% CI) [a] | FDR [b] |
| **Age (per 10 years)** | 1,983 (100) | 2.36 (2.14–2.59) | <0.001 | 958 (100) | 2.07 (1.86–2.31) | <0.001 | 100 (100) | 2.02 (1.42–2.88) | 0.002 |
| **Sex** | | | | | | | | | |
| Women | 766 (4.1) | Ref. | | 396 (8.9) | Ref. | | 36 (17.9) | Ref. | |
| Men | 1,217 (7.6) | 1.72 (1.56–1.90) | <0.001 | 562 (15.5) | 1.81 (1.56–2.10) | <0.001 | 64 (31.1) | 2.27 (1.41–3.66) | 0.007 |
| **Education** | | | | | | | | | |
| Medium or low | 1,460 (5.8) | Ref. | | 758 (12.6) | Ref. | | 76 (24.7) | Ref. | |
| High | 523 (5.4) | 0.93 (0.84–1.03) | 0.214 | 200 (9.8) | 0.90 (0.85–1.18) | 0.232 | 24 (24.2) | 0.82 (0.51–1.33) | 0.518 |
| **Smoking status** | | | | | | | | | |
| Never smoker | 713 (3.9) | Ref. | | 373 (9.1) | Ref. | | 42 (21.3) | Ref. | |
| Former smoker | 865 (6.6) | 1.44 (1.30–1.59) | <0.001 | 341 (13.1) | 1.15 (0.98–1.35) | 0.127 | 36 (24.2) | 1.07 (0.73–1.63) | 0.654 |
| Current smoker | 405 (12.0) | 3.08 (2.72–3.49) | <0.001 | 244 (17.5) | 2.18 (1.84–2.59) | <0.001 | 22 (26.1) | 1.60 (0.89–2.88) | 0.360 |
| **Physical activity** | | | | | | | | | |
| Low | 435 (6.4) | Ref. | | 286 (15.6) | Ref. | | 25 (28.7) | Ref. | |
| Moderate | 701 (5.8) | 0.96 (0.85–1.09) | 0.599 | 456 (12.3) | 0.74 (0.64–0.87) | <0.001 | 52 (26.5) | 0.93 (0.56–1.54) | 0.830 |
| High | 847 (5.3) | 0.87 (0.78–0.98) | 0.037 | 216 (8.2) | 0.58 (0.48–0.69) | <0.001 | 23 (18.6) | 0.72 (0.39–1.33) | 0.442 |
| **Alcohol consumption** | | | | | | | | | |
| Abstainer | 753 (5.4) | Ref. | | 566 (11.5) | Ref. | | 36 (26.7) | Ref. | |
| Low | 542 (5.3) | 1.01 (0.91–1.13) | 0.674 | 53 (10.6) | 1.03 (0.78–1.36) | 0.765 | 58 (23.4) | 1.06 (0.77–1.54) | 0.442 |
| Medium or high | 688 (6.5) | 1.25 (1.11–1.41) | <0.001 | 339 (12.8) | 1.09 (0.87–1.38) | 0.468 | 6 (25.0) | 1.46 (0.57–3.77) | 0.518 |
| **BMI** | | | | | | | | | |
| < 21.5 kg/m$^2$ | 149 (6.5) | 1.53 (1.27–1.85) | <0.001 | 59 (14.9) | 1.50 (1.12–2.02) | 0.012 | 5 (23.8) | 1.11 (0.42–2.97) | 0.835 |
| ≥ 21.5 – < 25 kg/m$^2$ | 405 (4.7) | Ref. | | 199 (11.0) | Ref. | | 28 (29.2) | Ref. | |
| ≥ 25 – < 30 kg/m$^2$ | 833 (5.5) | 0.95 (0.84–1.07) | 0.468 | 417 (11.0) | 0.81 (0.69–0.97) | 0.029 | 45 (24.5) | 0.76 (0.46–1.23) | 0.442 |
| ≥ 30 – < 35 kg/m$^2$ | 400 (6.4) | 0.99 (0.86–1.15) | 0.916 | 205 (12.7) | 0.88 (0.72–1.07) | 0.244 | 12 (16.9) | 0.43 (0.21–0.88) | 0.094 |
| ≥ 35 kg/m$^2$ | 196 (8.2) | 1.29 (1.08–1.55) | 0.009 | 78 (16.6) | 1.23 (0.94–1.62) | 0.183 | 10 (28.6) | 1.27 (0.95–1.98) | 0.139 |
| **Dyslipidemia** | | | | | | | | | |
| No | 1,269 (4.6) | Ref. | | 814 (11.4) | Ref. | | 86 (23.9) | Ref. | |
| Yes | 714 (9.7) | 1.07 (0.96–1.20) | 0.297 | 144 (15.8) | 1.05 (0.89–1.24) | 0.578 | 14 (29.8) | 1.01 (0.68–1.64) | 0.442 |
| **Hypertension** | | | | | | | | | |
| No | 551 (3.8) | Ref. | | 217 (7.4) | Ref. | | 25 (17.7) | Ref. | |
| Yes | 1,432 (7.1) | 1.28 (1.15–1.42) | <0.001 | 741 (14.4) | 1.47 (1.25–1.73) | <0.001 | 75 (28.2) | 1.25 (0.76–2.06) | 0.518 |
| **Diabetes** | | | | | | | | | |
| No | 1,573 (5.0) | Ref. | | 682 (9.9) | Ref. | | 70 (20.9) | Ref. | |
| Yes | 410 (11.5) | 1.54 (1.36–1.73) | <0.001 | 276 (22.4) | 1.71 (1.48–1.98) | <0.001 | 30 (41.7) | 1.78 (1.06–2.99) | 0.104 |
| **Cardiovascular disease** | | | | | | | | | |
| No | 1,642 (5.1) | Ref. | | 786 (10.6) | Ref. | | 79 (22.1) | Ref. | |
| Yes | 341 (14.1) | 1.63 (1.43–1.86) | <0.001 | 172 (26.2) | 1.75 (1.47–2.09) | <0.001 | 21 (42.0) | 1.44 (0.83–2.49) | 0.439 |
| **History of cancer** | | | | | | | | | |
| No | 1,600 (5.0) | Ref. | | 812 (10.9) | Ref. | | 92 (23.5) | Ref. | |
| Yes | 383 (12.6) | 2.44 (2.18–2.73) | <0.001 | 146 (23.4) | 2.09 (1.75–2.49) | <0.001 | 8 (50.0) | 3.00 (1.37–6.59) | 0.037 |

Abbreviations: BMI, body mass index; CI, confidence interval; HR, hazard ratio; Ref., reference.

[a] Cox proportional hazards regression model included all covariates: age, sex, education level, smoking status, physical activity, alcohol consumption, body mass index, dyslipidemia, hypertension, diabetes, cardiovascular disease, and cancer. Proportional hazards assumption was evaluated using Schoenfeld residuals (Grambsch and Therneau 1994; Sjölander and Dickman 2024; Stensrud and Hernán 2020).

[b] *P* values were adjusted for multiple comparisons using the Benjamini-Hochberg false discovery rate (FDR) method (Benjamini and Hochberg 1995).

[c] Validation sets 1 and 2 were derived from random subsets of the ESTHER Study. Serum proteins were measured using the Olink Target Inflammation Panel in validation set 1 and the Olink Explore HT Panel in validation set 2.

### 3.4.2 Sex-specific inflammatory and proteomic biomarkers associated with 10-year all-cause mortality risk

During 10 years of follow-up, 1,405 and 578 deaths occurred in the derivation set and internal validation set of the UKB, respectively. In the external validation sets 1 and 2 of the ESTHER study, 958 and 100 deaths were recorded in the first 10 years of follow-up, respectively. The breakdown of the case numbers in the sex-specific strata is shown in **Figure 6**.

From the Olink Target Inflammation panel, 11 and 9 proteins were selected for 10-year mortality prediction in the UKB's derivation sets for men and women, respectively, of which 7 were retrieved for both sexes (**Appendix, Table A9** and **Figure 29A**). **Figure 30A** shows the HRs and 95% confidence intervals for these proteins. Overall, 8 and 6 proteins were significantly associated with 10-year all-cause mortality after FDR correction in both internal and external validation datasets among men and women, respectively. More proteins (19 in men and 15 in women) were selected from the broader Olink Explore panels in the derivation set for 10-year mortality prediction, of which 7 were retrieved for both sexes (**Appendix, Table A9** and **Figure 29B**). Of these, 8 and 4 proteins were significantly associated with 10-year all-cause mortality after FDR correction in both internal and external validation datasets among men and women, respectively (**Figure 30B**). Several additional proteins showed directionally consistent associations in the internal and external validation datasets but were not statistically significant in the latter, likely due to the substantially lower number of deaths (n=100) compared with the external validation dataset used for the inflammatory proteins (n=958).

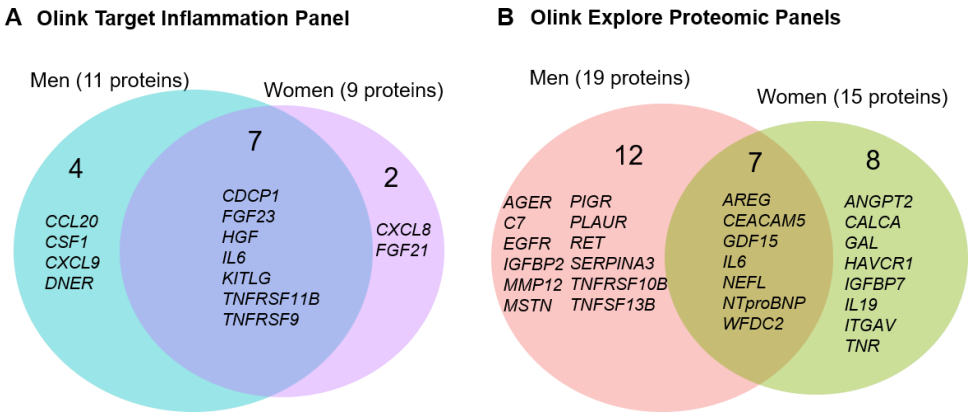



**Figure 29. Sex-specific overlap of inflammatory and proteomic biomarkers associated with 10-year all-cause mortality**

(**A**, **B**) Venn diagrams showing the overlap between LASSO-selected proteins in men and women from the Olink Target Inflammation Panel (**A**) and Olink Explore Proteomic Panels (**B**).

Abbreviations: LASSO, least absolute shrinkage and selection operator.

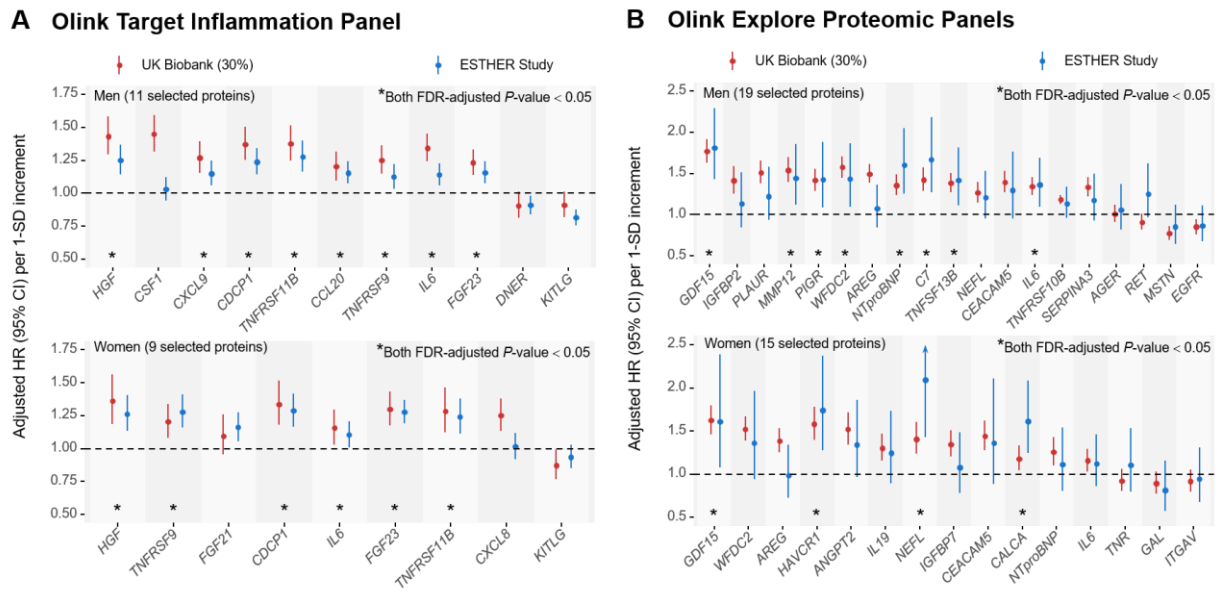



**Figure 30. Sex-specific inflammatory and proteomic biomarkers associated with 10-year all-cause mortality**

(**A**, **B**) Adjusted HRs and 95% CIs per 1-SD increment for LASSO-selected proteins in men and women from the Olink Target Inflammation Panel (**A**) and Olink Explore Proteomic Panels (**B**). Asterisks indicate proteins that remained significant after FDR correction in the UK Biobank internal validation and corresponding ESTHER external validation datasets.

Abbreviations: CI, confidence interval; FDR, false discovery rate; HR, hazard ratio; LASSO, least absolute shrinkage and selection operator; SD, standard deviation.

### 3.4.3 Biological pathways represented by mortality-associated protein signatures

Protein-protein interaction network analyses among the proteins selected from the Olink Target Inflammation panel exhibited broadly similar *interleukin-6* (*IL6)*-centered cytokine networks in men and women (**Figure 31A–B**). A comparable *IL6*-centered network was observed for women among the proteins selected from the broader Olink Explore panels. In contrast, the broader proteomic panels showed greater network divergence among men, characterized by a dual *epidermal growth factor receptor* (*EGFR)–IL6* hub structure.

Pathway enrichment analyses revealed that the protein selections from both Olink platforms are involved in cytokine-cytokine receptor interaction, pathways in cancer, viral protein interaction with cytokine and cytokine receptor, and PI3K-Akt signaling (**Figure 31C–D**). However, the broader proteomic panel selected more proteins from the PI3K-Akt signaling pathway and was additionally enriched in proteoglycan pathways in cancer.

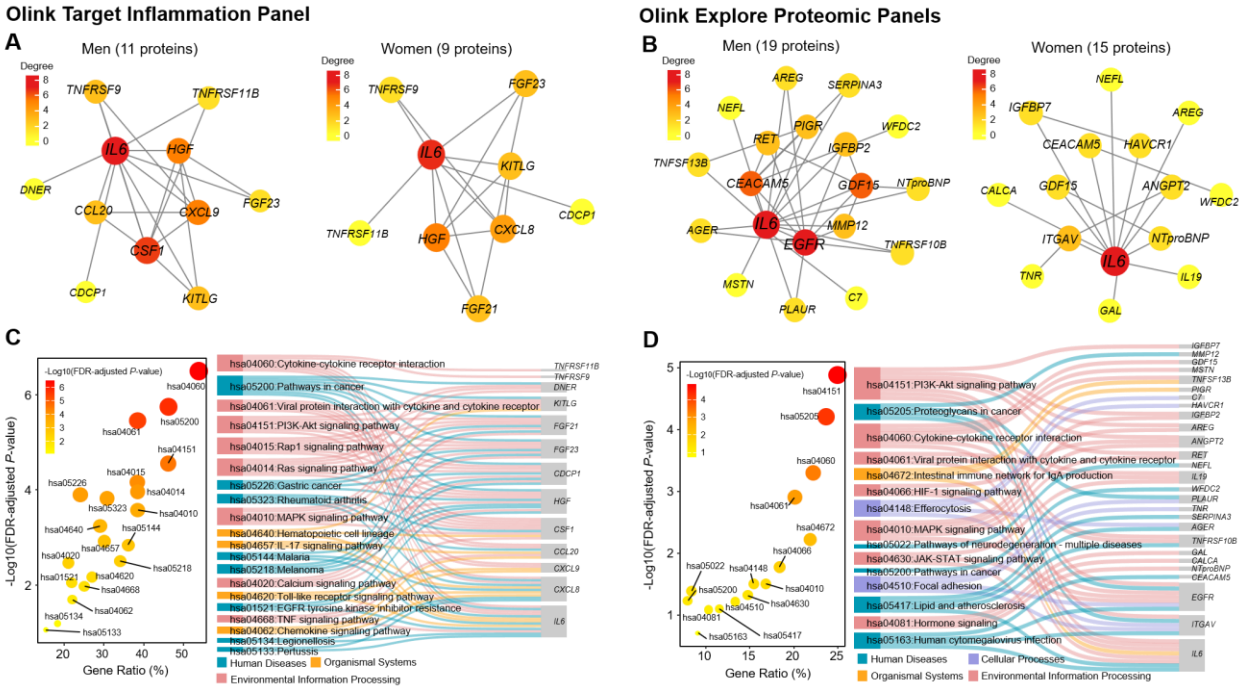



**Figure 31. Biological characteristics of sex-specific inflammatory and proteomic biomarkers associated with 10-year all-cause mortality**

(**A**, **B**) Protein-protein interaction networks of LASSO-selected proteins in men and women from the Olink Target Inflammation Panel (**A**) and Olink Explore Proteomic Panels (**B**). Nodes represent proteins and edges represent protein-protein interactions. (**C**, **D**) KEGG pathway enrichment analyses based on the union of LASSO-selected proteins in men and women from the Olink Target Inflammation Panel (**C**) and Olink Explore Proteomic Panels (**D**). Bubble plots (left) show significantly enriched pathways, and Sankey diagrams (right) illustrate relationships between enriched pathways and contributing proteins.

Abbreviations: KEGG, Kyoto Encyclopedia of Genes and Genomes; LASSO, least absolute shrinkage and selection operator.

### 3.4.4 Inflammatory and proteomic biomarkers improve prediction of mortality risk

The selected proteins showed only modest intercorrelations and correlations with conventional risk factors, with only a few high Spearman correlation coefficients between 0.40 and 0.57 (**Figure 32A–B**). Sex-specific risk scores were derived using LASSO Cox regression in the UKB derivation set and subsequently combined for the total population. The obtained $\beta$-coefficients of the final models are presented in **Appendix, Table A10**. The models were evaluated without re-estimation in the internal and external validation sets.

In all analyses, the model based solely on inflammatory biomarkers showed lower discrimination than the conventional risk factor model; the model based solely on the broader proteomic panel achieved higher discrimination than the conventional risk factor model, and the combined models achieved the highest C-statistic for 10-year mortality prediction (**Table 19**). Furthermore, predictive performance was always higher when proteins selected from the broader proteomic panel were added than when proteins from the Olink Target Inflammation panel were added. The highest C-

statistic was observed in the external validation set in the model combining proteins selected from the broader proteomic panel and traditional risk factors (C-statistic, 0.844), which was statistically significantly higher by 0.113 index points ($P$<0.001) than for the model of traditional risk factors alone (C-statistic, 0.731) and was accompanied by high, statistically significant (all $P$<0.001) NRI (65.2% and 42.6% for continuous and categorical NRI) and IDI values (9.1%). The model with only proteomics data achieved an almost as high C-statistic (0.827) as the model combining proteins and traditional risk factors (0.844).

**Table 20** shows the discrimination statistics for 5-year mortality prediction for these models. The patterns were the same, but generally higher C-statistics were obtained than for 10-year mortality prediction. The model combining proteins from the broader proteomics panel and traditional risk factors achieved a C-statistic of 0.891.

The patterns of model differences were also the same in sex-specific analyses. **Tables 21 and 22** show the results for 10-year mortality prediction for men and women, respectively. The best discriminating model combining proteins from the broader proteomic panel and traditional risk factors achieved C-statistics of 0.893 and 0.873 in men and women, respectively, and statistically significantly improved the performance of the reference model (C-statistic + 0.171, $P$<0.001 in men and C-statistic +0.107, $P$=0.006, respectively).

Individual protein analyses demonstrated that predictive improvements reflected the combined contribution of almost all proteins rather than dependence on a few dominant biomarkers (**Figure 33**). Moreover, the C-statistic gains by each biomarker were strikingly similar in the internal and external validation sets, underlining the robustness of the selected proteins.

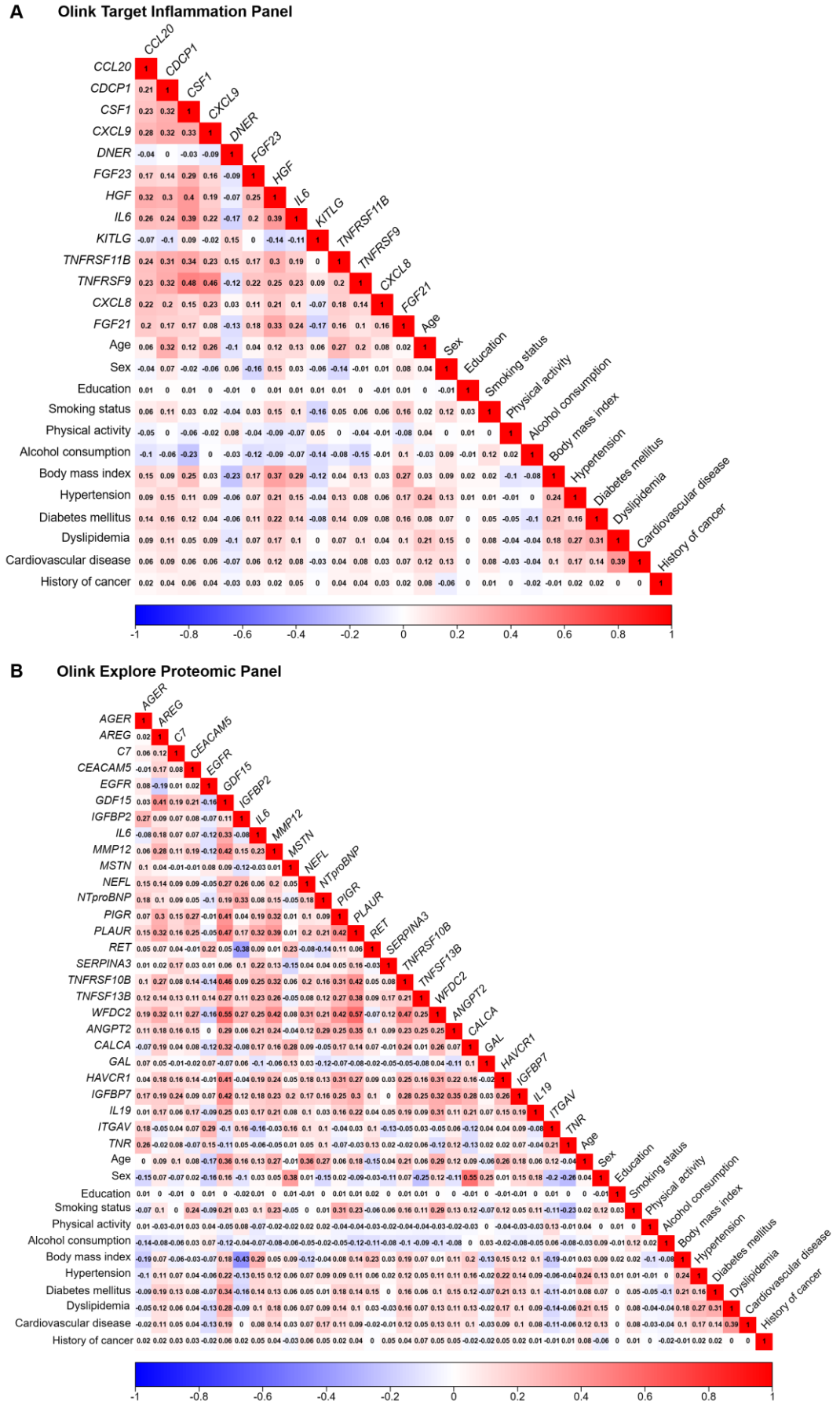



**Figure 32. Spearman correlations between conventional risk factors and selected inflammatory and proteomic biomarkers in the derivation set (70% UKB)**

(**A**), Olink Target Inflammation Panel. (**B**), Olink Explore Proteomic Panels. Red indicates positive correlations and blue indicates negative correlations, with color intensity proportional to correlation magnitude.

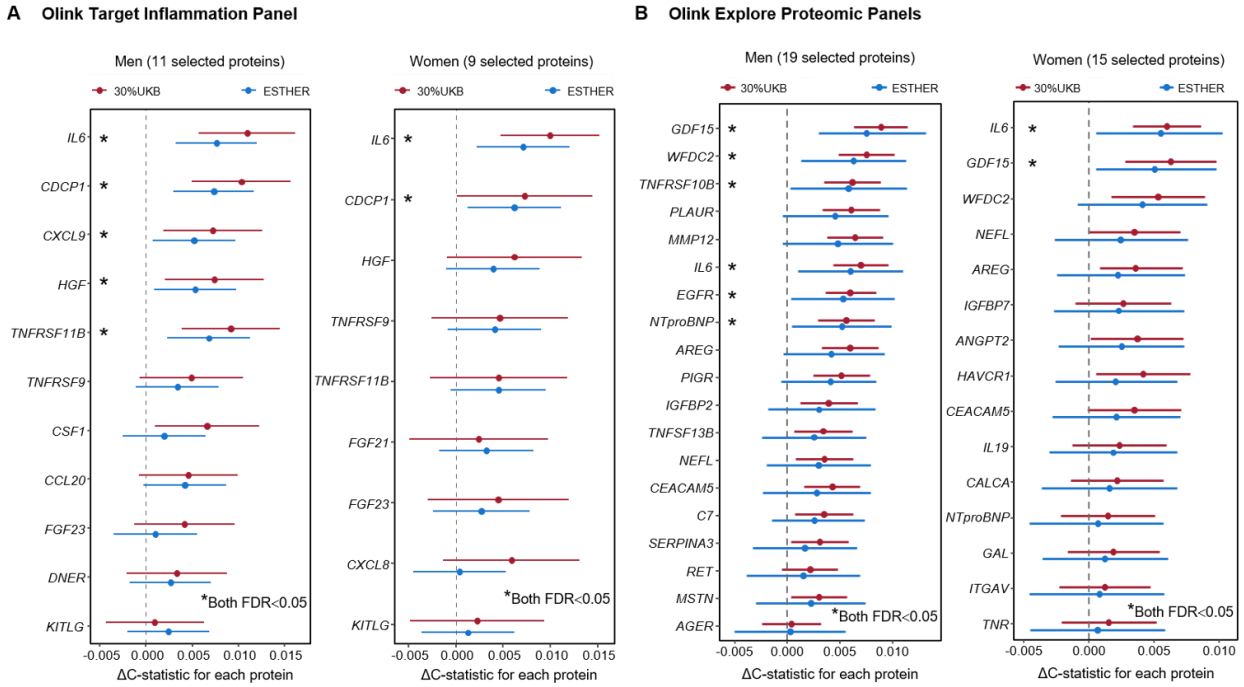



**Figure 33. Incremental improvements in discrimination (ΔC-statistics) for 10-year all-cause mortality after addition of individual inflammatory and proteomic biomarkers in internal and external validation sets**

(**A**) Olink Target Inflammation Panel in the UKB internal validation set (30% UKB) and ESTHER external validation subset 1. (**B**) Olink Explore Proteomic Panels in the UKB internal validation set (30% UKB) and ESTHER external validation subset 2. Forest plots show changes in time-dependent C-statistics (ΔC-statistics) after adding individual proteins to the conventional risk score for 10-year all-cause mortality prediction. The conventional risk score included established risk factors: age, sex, education, smoking status, physical activity, alcohol consumption, body mass index, dyslipidemia, hypertension, diabetes, cardiovascular disease, and history of cancer. Red and blue dots correspond to the UKB internal validation and ESTHER external validation sets, respectively. Error bars indicate 95% confidence intervals. Asterisks denote biomarkers with false discovery rate (FDR)-adjusted *P* values <0.05 in both validation sets. Protein annotations and corresponding UniProt identifiers are provided in **Supplementary Table 2**.

**Table 19. Inflammatory and proteomic panels differentially improve prediction of 10-year all-cause mortality in the overall population of UK Biobank and ESTHER cohorts**

| Outcome / Cohort / Model | C-statistic (95% CI) | *Δ* C-statistic | *Δ* Continuous NRI (95% CI) | *Δ* Categorical NRI (95% CI) [d] | *Δ* IDI (95% CI) |
|---|---|---|---|---|---|
| **Olink Target Inflammation Panel (10-Year)** | | | | | |
| **30% UK Biobank (Internal validation set)** | | | | | |
| Overall: n=10,456; 578 (5.5%) deaths | | | | | |
| Reference model [a] | 0.725 (0.704–0.747) | Ref. | Ref. | Ref. | Ref. |
| Inflammation model [b] | 0.716 (0.693–0.739) | - | - | - | - |
| Combined model | 0.766 (0.746–0.785) | +0.041, *P*<0.001 | 14.5% (13.4% – 15.6%), *P*<0.001 | 7.3% (3.6% – 11.3%), *P*<0.001 | 2.6% (2.0% – 3.1%), *P*<0.001 |
| **ESTHER Study (External validation set 1)** | | | | | |
| Overall: n=8,082; 958 (11.9%) deaths | | | | | |
| Reference model [a] | 0.720 (0.703–0.736) | Ref. | Ref. | Ref. | Ref. |
| Inflammation model [b] | 0.712 (0.690–0.733) | - | - | - | - |
| Combined model | 0.760 (0.738–0.781) | +0.040, *P*<0.001 | 14.3% (12.8% – 15.8%), *P*<0.001 | 6.9% (4.1% – 9.8%), *P*<0.001 | 2.2% (1.5% – 2.8%), *P*<0.001 |
| **Olink Explore Proteomic Panels (10-Year)** | | | | | |
| **30% UK Biobank (Internal validation set)** | | | | | |
| Overall: n=10,456; 578 (5.5%) deaths | | | | | |
| Reference model [a] | 0.725 (0.704–0.747) | Ref. | Ref. | Ref. | Ref. |
| Proteomics model [c] | 0.767 (0.746–0.788) | - | - | - | - |
| Combined model | 0.782 (0.762–0.802) | +0.057, *P*<0.001 | 34.4% (31.9% – 36.9%), *P*<0.001 | 16.4% (11.8% – 20.9%), *P*<0.001 | 4.0% (2.8% – 5.2%), *P*<0.001 |
| **ESTHER Study (External validation set 2)** | | | | | |
| Overall: n=408; 100 (24.5%) deaths | | | | | |
| Reference model [a] | 0.731 (0.673–0.789) | Ref. | Ref. | Ref. | Ref. |
| Proteomics model [c] | 0.827 (0.775–0.879) | - | - | - | - |
| Combined model | 0.844 (0.795–0.894) | +0.113, *P*<0.001 | 65.2% (58.5% – 71.9%), *P*<0.001 | 42.6% (29.4% – 55.7%), *P*<0.001 | 9.1% (4.5% – 14.3%), *P*<0.001 |

Abbreviations: BMI, body mass index; C-statistic, concordance statistic; CI, confidence interval; IDI, integrated discrimination improvement; NRI, net reclassification improvement; t-, time-dependent.
Notes: All *P* values for model improvements were <0.05. Differences in time-dependent C-statistics were assessed using bootstrap resampling, and NRI and IDI were tested using likelihood ratio-based inference.
[a] The reference model included the following conventional risk covariates: chronological age, sex, education, smoking status, physical activity, alcohol consumption, BMI, dyslipidemia, hypertension, diabetes, cardiovascular disease, and history of cancer.
[b] The inflammation model included 13 proteins from the Olink Target Inflammation Panel (see **Appendix, Table A9** for detailed information): *CCL20*, *CSF1*, *CXCL9*, *DNER*, *CDCP1*, *FGF23*, *HGF*, *IL6*, *KITLG*, *TNFRSF11B*, *TNFRSF9*, *CXCL8*, and *FGF21*.
[c] The proteomics model included 27 proteins from the Olink Explore Proteomic Panels (see **Appendix, Table A9** for detailed information): *AGER*, *C7*, *EGFR*, *IGFBP2*, *MMP12*, *MSTN*, *PIGR*, *PLAUR*, *RET*, *SERPINA3*, *TNFRSF10B*, *TNFSF13B*, *AREG*, *CEACAM5*, *GDF15*, *IL6*, *NEFL*, *NTproBNP*, *WFDC2*, *ANGPT2*, *CALCA*, *GAL*, *HAVCR1*, *IGFBP7*, *IL19*, *ITGAV*, and *TNR*.
[d] Categorical NRI was calculated based on predefined risk thresholds of <5%, 5%–10%, and >10%.

**Table 20. Inflammatory and proteomic panels differentially improve prediction of 5-year all-cause mortality in the overall population of UK Biobank and ESTHER cohorts**

| Outcome / Cohort / Model | C-statistic (95% CI) | *Δ* C-statistic | *Δ* Continuous NRI (95% CI) | *Δ* Categorical NRI (95% CI) [d] | *Δ* IDI (95% CI) |
|---|---|---|---|---|---|
| **Olink Target Inflammation Panel (5-Year)** | | | | | |
| **30% UK Biobank (Internal validation set)** | | | | | |
| Overall: n=10,456; 209 (2.0%) deaths | | | | | |
| Reference model [a] | 0.742 (0.707–0.778) | Ref. | Ref. | Ref. | Ref. |
| Inflammation model [b] | 0.740 (0.705–0.774) | - | - | - | - |
| Combined model | 0.796 (0.765–0.828) | +0.054, *P*<0.001 | 22.4% (18.8%–25.9%), *P*<0.001 | 11.5% (3.9%–18.5%), *P*<0.001 | 3.6% (3.0%–4.1%), *P*<0.001 |
| **ESTHER Study (External validation set 1)** | | | | | |
| Overall: n=8,082; 370 (4.6%) deaths | | | | | |
| Reference model [a] | 0.730 (0.696–0.763) | Ref. | Ref. | Ref. | Ref. |
| Inflammation model [b] | 0.725 (0.692–0.758) | - | - | - | - |
| Combined model | 0.783 (0.754–0.812) | +0.053, *P*<0.001 | 29.2% (26.5%–31.9%), *P*<0.001 | 13.6% (8.4%–18.6%), *P*<0.001 | 3.2% (2.5%–3.8%), *P*<0.001 |
| **Olink Explore Proteomic Panels (5-Year)** | | | | | |
| **30% UK Biobank (Internal validation set)** | | | | | |
| Overall: n=10,456; 209 (2.0%) deaths | | | | | |
| Reference model [a] | 0.742 (0.707–0.778) | Ref. | Ref. | Ref. | Ref. |
| Proteomics model [c] | 0.789 (0.766–0.811) | - | - | - | - |
| Combined model | 0.827 (0.798–0.857) | +0.085, *P*<0.001 | 46.7% (43.2%–50.2%), *P*<0.001 | 30.3% (22.5%–37.5%), *P*<0.001 | 5.6% (5.0%–6.2%), *P*<0.001 |
| **ESTHER Study (External validation set 2)** | | | | | |
| Overall: n=408; 34 (8.3%) deaths | | | | | |
| Reference model [a] | 0.722 (0.626–0.819) | Ref. | Ref. | Ref. | Ref. |
| Proteomics model [c] | 0.840 (0.753–0.927) | - | - | - | - |
| Combined model | 0.891 (0.826–0.957) | +0.169, *P*<0.001 | 78.7% (67.4%–90.1%), *P*<0.001 | 53.7% (31.7%–74.9%), *P*<0.001 | 11.8% (10.7%–12.9%), *P*<0.001 |

Abbreviations: BMI, body mass index; C-statistic, concordance statistic; CI, confidence interval; IDI, integrated discrimination improvement; NRI, net reclassification improvement; t-, time-dependent.
Notes: All *P* values for model improvements were <0.05. Differences in time-dependent C-statistics were assessed using bootstrap resampling, and NRI and IDI were tested using likelihood ratio-based inference.
[a] The reference model included the following conventional risk covariates: chronological age, sex, education, smoking status, physical activity, alcohol consumption, BMI, dyslipidemia, hypertension, diabetes, cardiovascular disease, and history of cancer.
[b] The inflammation model included 13 proteins from the Olink Target Inflammation Panel (see **Appendix, Table A9** for detailed information): *CCL20*, *CSF1*, *CXCL9*, *DNER*, *CDCP1*, *FGF23*, *HGF*, *IL6*, *KITLG*, *TNFRSF11B*, *TNFRSF9*, *CXCL8*, and *FGF21*.
[c] The proteomics model included 27 proteins from the Olink Explore Proteomic Panels (see **Appendix, Table A9** for detailed information): *AGER*, *C7*, *EGFR*, *IGFBP2*, *MMP12*, *MSTN*, *PIGR*, *PLAUR*, *RET*, *SERPINA3*, *TNFRSF10B*, *TNFSF13B*, *AREG*, *CEACAM5*, *GDF15*, *IL6*, *NEFL*, *NTproBNP*, *WFDC2*, *ANGPT2*, *CALCA*, *GAL*, *HAVCR1*, *IGFBP7*, *IL19*, *ITGAV*, and *TNR*.
[d] Categorical NRI was calculated based on predefined risk thresholds of <5%, 5%–10%, and >10%.

**Table 21. Inflammatory and proteomic panels differentially improve prediction of 10-year all-cause mortality in men from UK Biobank and ESTHER cohorts**

| Outcome / Cohort / Model | C-statistic (95% CI) | *Δ* C-statistic | *Δ* Continuous NRI (95% CI) | *Δ* Categorical NRI (95% CI) [d] | *Δ* IDI (95% CI) |
|---|---|---|---|---|---|
| **Olink Target Inflammation Panel (10-Year)** | | | | | |
| **30% UK Biobank (Internal validation set)** | | | | | |
| Men: n=4,884; 365 (7.5%) deaths | | | | | |
| Reference model [a] | 0.716 (0.678–0.753) | Ref. | Ref. | Ref. | Ref. |
| Inflammation model [b] | 0.673 (0.634–0.712) | - | - | - | - |
| Combined model | 0.745 (0.708–0.781) | +0.029, *P*<0.001 | 11.6% (9.1%–14.1%), *P*<0.001 | 5.9% (0.9%–11.1%), *P*=0.015 | 2.9% (2.0%–4.0%), *P*<0.001 |
| **ESTHER Study (External validation set 1)** | | | | | |
| Men: n=3,626; 562 (15.5%) deaths | | | | | |
| Reference model [a] | 0.740 (0.718–0.762) | Ref. | Ref. | Ref. | Ref. |
| Inflammation model [b] | 0.710 (0.687–0.733) | - | - | - | - |
| Combined model | 0.778 (0.757–0.799) | +0.038, *P*<0.001 | 19.2% (16.5%–21.9%), *P*<0.001 | 10.9% (6.9%–15.0%), *P*<0.001 | 4.8% (3.6%–6.0%), *P*<0.001 |
| **Olink Explore Proteomic Panels (10-Year)** | | | | | |
| **30% UK Biobank (Internal validation set)** | | | | | |
| Men: n=4,884; 365 (7.5%) deaths | | | | | |
| Reference model [a] | 0.716 (0.678–0.753) | Ref. | Ref. | Ref. | Ref. |
| Proteomics model [c] | 0.730 (0.694–0.766) | - | - | - | - |
| Combined model | 0.786 (0.752–0.819) | +0.070, *P*<0.001 | 39.6% (36.8%–42.4%), *P*<0.001 | 19.9% (14.5%–25.8%), *P*<0.001 | 6.6% (5.0%–8.5%), *P*<0.001 |
| **ESTHER Study (External validation set 2)** | | | | | |
| Men: n=207; 64 (30.9%) deaths | | | | | |
| Reference model [a] | 0.722 (0.646–0.799) | Ref. | Ref. | Ref. | Ref. |
| Proteomics model [c] | 0.865 (0.810–0.920) | - | - | - | - |
| Combined model | 0.893 (0.847–0.940) | +0.171, *P*<0.001 | 53.6% (43.2%–64.0%), *P*<0.001 | 35.7% (14.8%–55.0%), *P*<0.001 | 9.7% (3.7%–15.7%), *P*<0.001 |

Abbreviations: BMI, body mass index; C-statistic, concordance statistic; CI, confidence interval; IDI, integrated discrimination improvement; NRI, net reclassification improvement; t-, time-dependent.
Notes: All *P* values for model improvements were <0.05. Differences in time-dependent C-statistics were assessed using bootstrap resampling, and NRI and IDI were tested using likelihood ratio-based inference.
[a] The reference model included the following conventional risk covariates: chronological age, sex, education, smoking status, physical activity, alcohol consumption, BMI, dyslipidemia, hypertension, diabetes, cardiovascular disease, and history of cancer.
[b] The inflammation model included 11 proteins from the Olink Target Inflammation Panel (see **Appendix, Table A9** for detailed information): *CCL20*, *CSF1*, *CXCL9*, *DNER*, *CDCP1*, *FGF23*, *HGF*, *IL6*, *KITLG*, *TNFRSF11B*, and *TNFRSF9*.
[c] The proteomics model included 19 proteins from the Olink Explore Proteomic Panels (see **Appendix, Table A9** for detailed information): *AGER*, *C7*, *EGFR*, *IGFBP2*, *MMP12*, *MSTN*, *PIGR*, *PLAUR*, *RET*, *SERPINA3*, *TNFRSF10B*, *TNFSF13B*, *AREG*, *CEACAM5*, *GDF15*, *IL6*, *NEFL*, *NTproBNP*, and *WFDC2*.
[d] Categorical NRI was calculated based on predefined risk thresholds of <5%, 5%–10%, and >10%.

**Table 22. Inflammatory and proteomic panels differentially improve prediction of 10-year all-cause mortality in women from UKB and ESTHER cohorts**

| Outcome / Cohort / Model | C-statistic (95% CI) | Δ C-statistic | Δ Continuous NRI (95% CI) | Δ Categorical NRI (95% CI) [d] | Δ IDI (95% CI) |
|---|---|---|---|---|---|
| **Olink Target Inflammation Panel (10-Year)** | | | | | |
| **30% UK Biobank (Internal validation set)** | | | | | |
| Women: n=5,572; 213 (3.8%) deaths | | | | | |
| Reference model [a] | 0.727 (0.700–0.754) | Ref. | Ref. | Ref. | Ref. |
| Inflammation model [b] | 0.696 (0.667–0.725) | - | - | - | - |
| Combined model | 0.759 (0.733–0.786) | +0.032, *P*<0.001 | 10.7% (8.3%–13.1%), *P*<0.001 | 4.4% (0.6%–10.0%), *P*=0.029 | 1.4% (0.6%–2.3%), *P*=0.001 |
| **ESTHER Study (External validation set 1)** | | | | | |
| Women: n=4,456; 396 (8.9%) deaths | | | | | |
| Reference model [a] | 0.713 (0.687–0.739) | Ref. | Ref. | Ref. | Ref. |
| Inflammation model [b] | 0.636 (0.606–0.665) | - | - | - | - |
| Combined model | 0.748 (0.722–0.775) | +0.035, *P*<0.001 | 13.7% (12.1%–15.3%), *P*<0.001 | 6.5% (2.3%–10.7%), *P*<0.001 | 2.1% (0.5%–3.8%), *P*<0.001 |
| **Olink Explore Proteomic Panels (10-Year)** | | | | | |
| **30% UK Biobank (Internal validation set)** | | | | | |
| Women: n=5,572; 213 (3.8%) deaths | | | | | |
| Reference model [a] | 0.727 (0.700–0.754) | Ref. | Ref. | Ref. | Ref. |
| Proteomics model [c] | 0.775 (0.750–0.800) | - | - | - | - |
| Combined model | 0.784 (0.760–0.809) | +0.057, *P*<0.001 | 28.8% (27.1%–30.5%), *P*<0.001 | 14.6% (13.1%–16.1%), *P*<0.001 | 4.1% (2.6%–5.6%), *P*<0.001 |
| **ESTHER Study (External validation set 2)** | | | | | |
| Women: n=201; 36 (17.9%) deaths | | | | | |
| Reference model [a] | 0.766 (0.677–0.856) | Ref. | Ref. | Ref. | Ref. |
| Proteomics model [c] | 0.838 (0.764–0.913) | - | - | - | - |
| Combined model | 0.873 (0.808–0.937) | +0.107, *P*=0.006 | 56.8% (46.3%–68.3%), *P*<0.001 | 36.9% (15.8%–58.9%), *P*<0.001 | 9.9% (8.8%–11.7%), *P*<0.001 |

Abbreviations: BMI, body mass index; C-statistic, concordance statistic; CI, confidence interval; IDI, integrated discrimination improvement; NRI, net reclassification improvement; t-, time-dependent.
Notes: All *P* values for model improvements were <0.05. Differences in time-dependent C-statistics were assessed using bootstrap resampling, and NRI and IDI were tested using likelihood ratio-based inference.
[a] The reference model included the following conventional risk covariates: chronological age, sex, education, smoking status, physical activity, alcohol consumption, BMI, dyslipidemia, hypertension, diabetes, cardiovascular disease, and history of cancer.
[b] The inflammation model included 9 proteins from the Olink Target Inflammation Panel (see **Appendix, Table A9** for detailed information): *CXCL8*, *FGF21*, *CDCP1*, *FGF23*, *HGF*, *IL6*, *KITLG*, *TNFRSF11B*, and *TNFRSF9*.
[c] The proteomics model included 15 proteins from the Olink Explore Proteomic Panels (see **Appendix, Table A9** for detailed information): *ANGPT2*, *CALCA*, *GAL*, *HAVCR1*, *IGFBP7*, *IL19*, *ITGAV*, *TNR*, *AREG*, *CEACAM5*, *GDF15*, *IL6*, *NEFL*, *NTproBNP*, and *WFDC2*.
[d] Categorical NRI was calculated based on predefined risk thresholds of <5%, 5%–10%, and >10%.

### 3.4.5 Development and validation of inflammation- and proteomics-based ageing clocks

LASSO-selected proteins associated with all-cause mortality were used to develop sex-specific inflammaging and proteomic ageing clocks by elastic net regression against chronological age, with or without traditional risk factors. The equations for the inflammaging clocks (InflamAge1 with traditional risk factors and InflamAge2 without) and proteomic ageing clocks (ProtAge1 with traditional risk factors and ProtAge2 without) are provided in **Appendix, Table A11**. Their distributions and strong correlations with chronological age in the internal and external validation sets are shown in **Figures 34–36**.

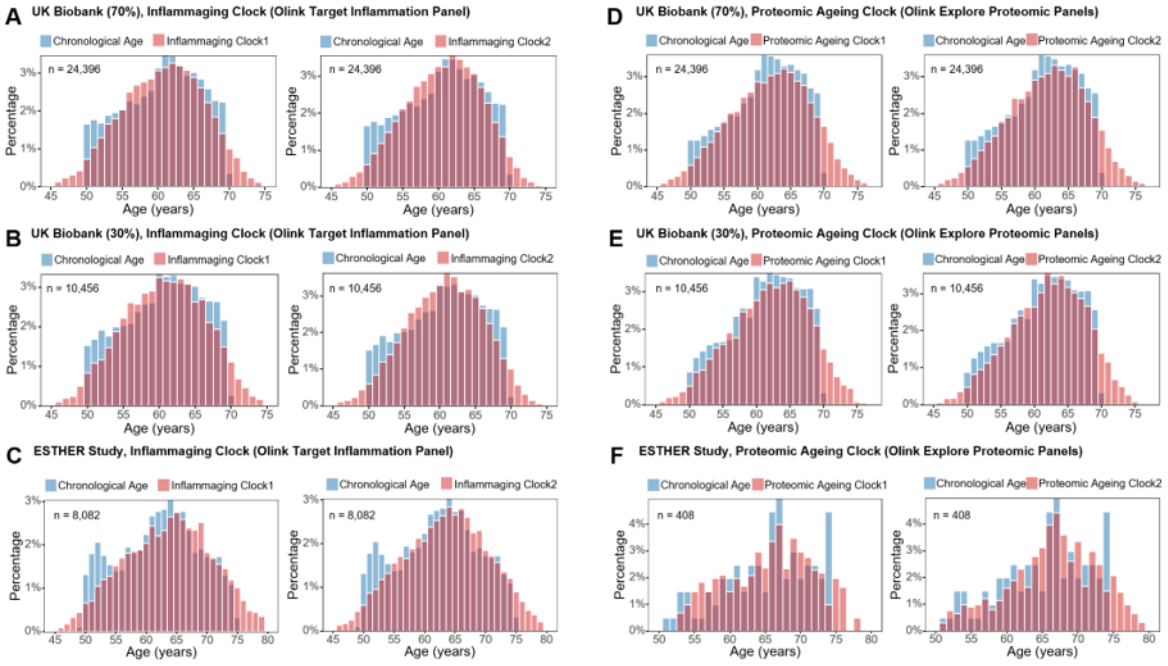



**Figure 34. Distributions of chronological age and mortality risk ageing clocks in the UK Biobank and ESTHER cohorts**

(**A–C**) Inflammaging Clocks in the UK Biobank derivation set (**A**), UK Biobank internal validation set (**B**), and ESTHER external validation subset 1 (**C**); (**D–F**) Proteomic Ageing Clocks in the UK Biobank derivation set (**D**), UK Biobank internal validation set (**E**), and ESTHER external validation subset 2 (**F**).

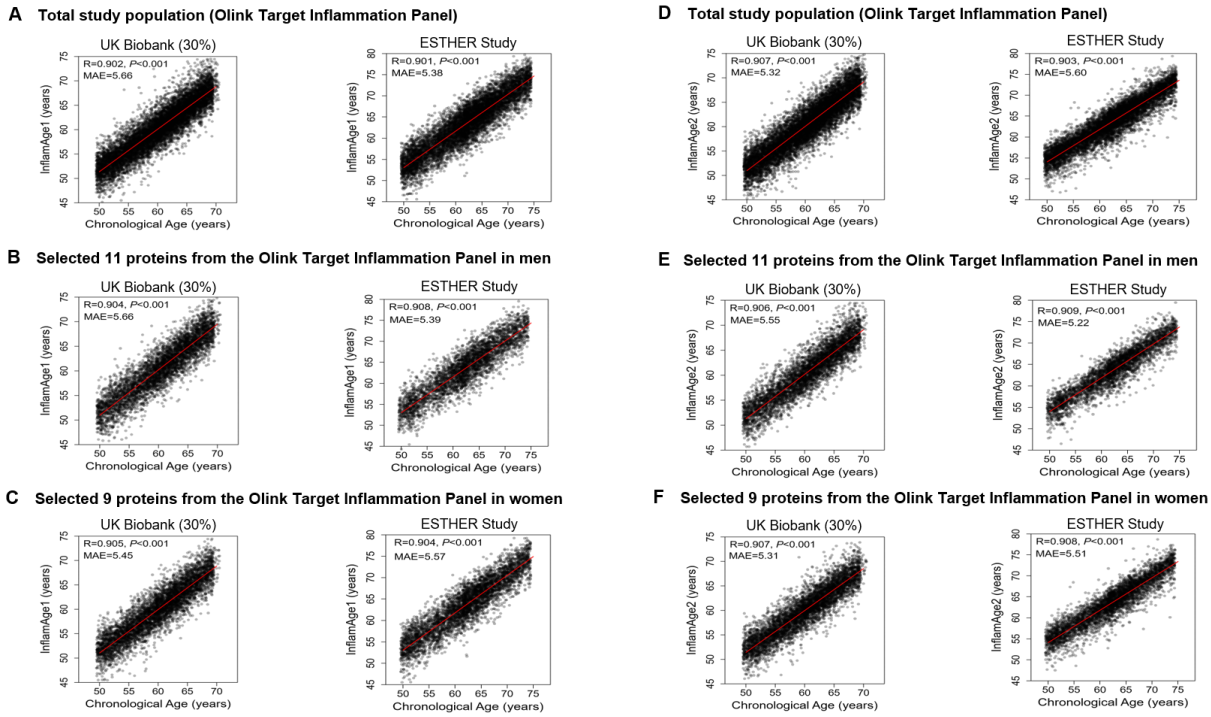



**Figure 35. Linear correlations between chronological age and inflammatory age estimated by the inflammaging clocks derived from the Olink Target Inflammation Panel**

(**A–C**) Linear correlations between chronological age and inflammatory age estimated by InflamAge1, derived from LASSO-selected proteomic proteins only, for the overall study population (**A**), men (**B**), and women (**C**) in the UK Biobank internal validation set (30% UKB) and ESTHER external validation subset 1. (**D–F**) Corresponding correlations for InflamAge2, derived from LASSO-selected proteomic proteins combined with conventional risk factors, for the overall study population (**D**), men (**E**), and women (**F**) in the UK Biobank internal validation set (30% UKB) and ESTHER external validation subset 1. Pearson correlation coefficients (*R*) are shown in the upper left corner of each panel. Mean absolute error (MAE), reflecting the average absolute deviation between estimated inflammatory age and chronological age, is also provided; lower MAE values indicate greater prediction accuracy.

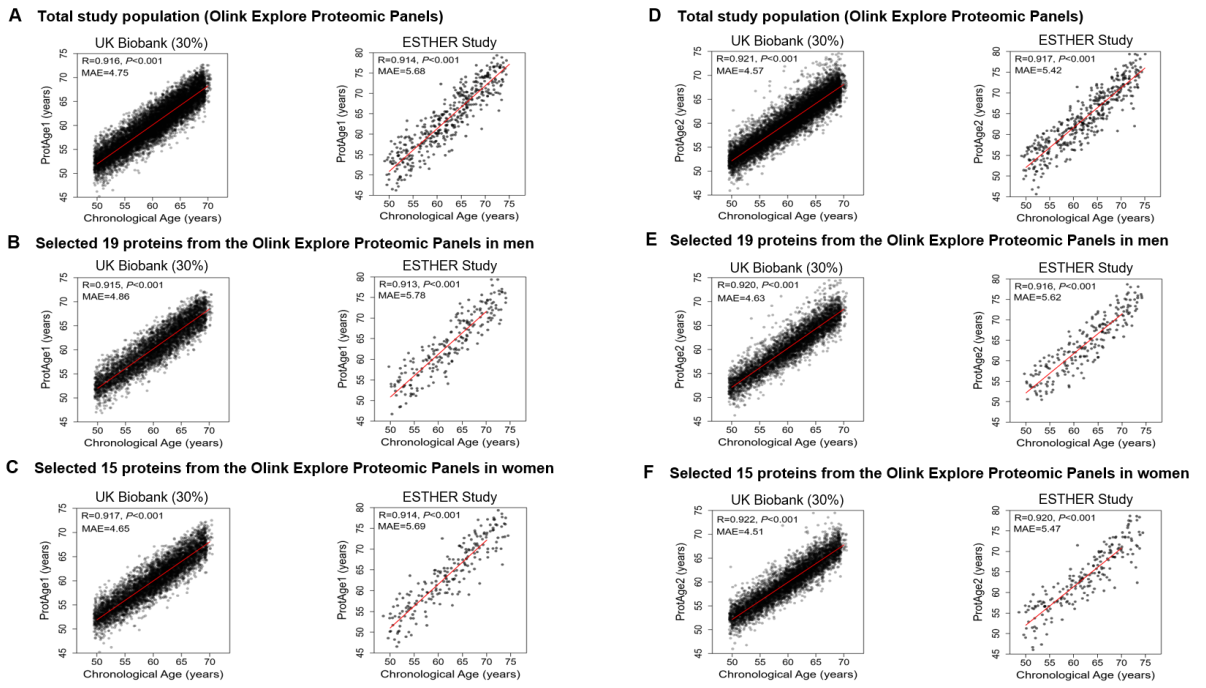


**Figure 36. Linear correlations between chronological age and proteomic age estimated by the proteomic ageing clocks derived from the Olink Explore Panels**

(**A–C**) Linear correlations between chronological age and proteomic age estimated by ProtAge1, derived from LASSO-selected proteomic proteins only, for the overall study population (**A**), men (**B**), and women (**C**) in the UK Biobank internal validation set (30% UKB) and ESTHER external validation subset 2. (**D–F**) Corresponding correlations for ProtAge2, derived from LASSO-selected proteomic proteins combined with conventional risk factors, for the overall study population (**D**), men (**E**), and women (**F**) in the UK Biobank internal validation set (30% UKB) and ESTHER external validation subset 2. Pearson correlation coefficients (*R*) are shown in the upper left corner of each panel. Mean absolute error (MAE), reflecting the average absolute deviation between estimated proteomic age and chronological age, is also provided; lower MAE values indicate greater prediction accuracy.

### 3.4.6 Inflammaging and proteomic age accelerations are associated with all-cause and cause-specific mortality

In internal and external validation cohorts, as well as across sexes, age acceleration, measured by the inflammation-based and proteomic ageing clocks, was consistently statistically significantly associated with higher 10-year all-cause mortality (**Table 23**). Results did not differ much by sex and incorporation of traditional risk factors in the clocks. However, the proteomic ageing clocks had a stronger association with 10-year mortality than the inflammaging clocks. Age acceleration based on ProtAge2 demonstrated the strongest associations overall. In the total external validation cohort, each one-year increase was associated with a 22% higher all-cause mortality (HR=1.22, 95% CI 1.15–1.31). The corresponding HR (95% CI) for age acceleration based on InflamAge2 was 1.14 (1.11–1.17).

The results for CVD and cancer mortality in the total UKB set are shown in **Table 24** (case numbers in the validation sets were too low for cause-specific mortality). Differences by clocks were similar to those seen for all-cause mortality, but associations were generally stronger with CVD mortality than with cancer mortality. For example, each one-year increase in age acceleration based on ProtAge2 was associated with a 60% higher CVD mortality (HR=1.60, 95% CI 1.54–1.66) and a 33% higher cancer mortality (HR=1.33, 95% CI 1.30–1.36).

**Table 23. Association of inflammaging/proteomic age acceleration based on proteins only and based on both proteins and traditional risk factors with 10-year all-cause mortality**

| Ageing Clock / Study population | Internal validation UK Biobank (30%) [a] | | External validation ESTHER [b] | |
|---|---|---|---|---|
| | Mean AA ± SD (years) | Adjusted HR per 1-year AA increase [c] | Mean AA ± SD (years) | Adjusted HR per 1-year AA increase [c] |
| **Inflammaging Clocks** | | | | |
| **Proteins only (InflamAge1)** | | | | |
| Men | 0.56±5.57 | **1.24 (1.18–1.30)** | 0.72±6.84 | **1.10 (1.04–1.16)** |
| Women | 0.48±5.29 | **1.23 (1.16–1.32)** | 0.54±6.37 | **1.09 (1.05–1.15)** |
| Overall | 0.54±5.44 | **1.24 (1.19–1.29)** | 0.62±6.61 | **1.10 (1.06–1.13)** |
| **Proteins and risk factors (InflamAge2) [c]** | | | | |
| Men | 0.41±5.19 | **1.25 (1.16–1.34)** | 0.55±5.76 | **1.14 (1.10–1.19)** |
| Women | 0.29±4.87 | **1.23 (1.17–1.30)** | 0.36±5.40 | **1.13 (1.09–1.17)** |
| Overall | 0.38±5.03 | **1.26 (1.21–1.31)** | 0.47±5.59 | **1.14 (1.11–1.17)** |
| **Proteomic Aging Clocks** | | | | |
| **Proteins only (ProtAge1)** | | | | |
| Men | 0.43±5.46 | **1.29 (1.22–1.38)** | 1.89±7.58 | **1.22 (1.07–1.38)** |
| Women | 0.35±4.83 | **1.23 (1.17–1.28)** | 1.75±6.97 | **1.19 (1.09–1.31)** |
| Overall | 0.40±5.21 | **1.25 (1.20–1.29)** | 1.80±7.24 | **1.20 (1.12–1.29)** |
| **Proteins and risk factors (ProtAge2) [c]** | | | | |
| Men | 0.34±5.05 | **1.33 (1.25–1.41)** | 1.74±6.93 | **1.23 (1.11–1.36)** |
| Women | 0.27±4.62 | **1.26 (1.21–1.32)** | 1.68±6.36 | **1.20 (1.11–1.30)** |
| Overall | 0.32±4.81 | **1.30 (1.25–1.35)** | 1.59±6.12 | **1.22 (1.15–1.31)** |

Abbreviations: AA, Age acceleration; BMI, body mass index; InflamAge, inflammaging age; ProtAge, proteomic age; CI, confidence interval; HR, hazard ratio; SD, standard deviation.
Notes: Printed in bold: Statistically significant ($P < 0.05$).
[a] n=10,456; 578 (5.5%) deaths.
[b] For inflammaging clocks: n = 8,082 including 958 deaths; for proteomic aging clocks: n = 408 including 100 deaths.
[c] Models for "Proteins only" clocks were adjusted for chronological age, sex, education, smoking status, physical activity, alcohol consumption, BMI, dyslipidemia, hypertension, diabetes, cardiovascular disease, and history of cancer. HRs were estimated per 1-year increase in inflammatory/proteomic age acceleration, defined as the residual from regressing the corresponding biological age estimate on chronological age. Models for "proteins and risk factors" already included these risk factors and were therefore not adjusted.

**Table 24.** Association of inflammaging/proteomic age acceleration based on proteomic biomarkers only and on both biomarkers and traditional risk factors with 10-year CVD and cancer mortality in the overall study population of UK Biobank

| AA / Study population | CVD mortality | | | Cancer mortality | | |
|---|---|---|---|---|---|---|
| | $N_{total}$ | $N_{cases}$ | Adjusted HR per 1-year MetAA increase [a] | $N_{total}$ | $N_{cases}$ | Adjusted HR per 1-year MetAA increase [a] |
| **InflamAA** | **UK Biobank (100%)** | | | **UK Biobank (100%)** | | |
| **Olink Target Inflammation Panel (InflamAA1)** | | | | | | |
| Men | 15,950 | 306 | **1.44 (1.22–1.67)** | 15,950 | 621 | **1.22 (1.15–1.30)** |
| Women | 18,902 | 103 | **1.32 (1.20–1.44)** | 18,902 | 506 | **1.15 (1.08–1.22)** |
| Overall | 34,852 | 409 | **1.45 (1.34–1.57)** | 34,852 | 1,127 | **1.22 (1.17–1.27)** |
| **Olink Target Inflammation Panel (InflamAA2) [b]** | | | | | | |
| Men | 15,950 | 306 | **1.47 (1.38–1.55)** | 15,950 | 621 | **1.25 (1.20–1.30)** |
| Women | 18,902 | 103 | **1.44 (1.31–1.58)** | 18,902 | 506 | **1.21 (1.16–1.26)** |
| Overall | 34,852 | 409 | **1.50 (1.43–1.57)** | 34,852 | 1,127 | **1.25 (1.22–1.29)** |
| **ProtAA** | **UK Biobank (100%)** | | | **UK Biobank (100%)** | | |
| **Olink Explore Proteomic Panels (ProtAA1)** | | | | | | |
| Men | 15,950 | 306 | **1.55 (1.40–1.71)** | 15,950 | 621 | **1.31 (1.26–1.36)** |
| Women | 18,902 | 103 | **1.49 (1.41–1.58)** | 18,902 | 506 | **1.21 (1.16–1.27)** |
| Overall | 34,852 | 409 | **1.56 (1.48–1.63)** | 34,852 | 1,127 | **1.27 (1.23–1.31)** |
| **Olink Explore Proteomic Panels (ProtAA2) [b]** | | | | | | |
| Men | 15,950 | 306 | **1.64 (1.52–1.76)** | 15,950 | 621 | **1.36 (1.32–1.40)** |
| Women | 18,902 | 103 | **1.52 (1.45–1.59)** | 18,902 | 506 | **1.24 (1.19–1.29)** |
| Overall | 34,852 | 409 | **1.60 (1.54–1.66)** | 34,852 | 1,127 | **1.33 (1.30–1.36)** |

Abbreviations: CI, confidence interval; CVD, cardiovascular disease; HR, hazard ratio; InflamAA, inflammaging age acceleration; ProtAA, proteomic age acceleration.

Notes: Printed in bold: Statistically significant ($P$<0.05).

[a] The adjusted Cox proportional hazards regression model included the following covariates: chronological age, sex, education, smoking status, physical activity, alcohol consumption, body mass index, dyslipidemia, hypertension, diabetes, cardiovascular disease, and history of cancer. HRs were estimated per 1-year increase in inflammatory/proteomic age acceleration, defined as the residual from regressing the corresponding biological age estimate on chronological age.

[b] Since traditional risk factors are already intrinsic to the inflammaging clock2 and proteomic ageing clock2, no additional adjustment for these covariates was performed to avoid over-adjustment.

# 4 DISCUSSION

## 4.1 Study 1: Comparison of four frailty measurements

### 4.1.1 Summary of the key results

This study provides a comprehensive comparison of four widely used frailty measures (CFS, FFP, FI, and FS) in relation to aging, physical performance and all-cause mortality in a large population-based cohort of older adults. The findings underscore the differences among these measures regarding their predictive value for the different outcomes, and across sex and age groups.

### 4.1.2 Agreement of frailty classifications

Although the proportion of subjects classified as prefrail and frail was similar across the FFP, FI, and FS, and CFS measures, their inter-rater agreement was fair to moderate. This reflects the conceptual differences in frailty measurements. Each test has a unique methodological feature: The FS relies only on self-reported information, the CFS on clinical judgement of physicians, the FI only includes age-related deficits, and the FFP is the only instrument that includes physical assessments (grip strength and gait speed). The only instruments with a moderate agreement were the CFS and FI. This aligns with prior research suggesting that frailty measures capturing cumulative deficits (e.g., FI) tend to correlate more strongly with clinical frailty assessments (e.g., CFS) because for both, age-related diseases play an important role (Moreno-Ariño et al. 2020). In contrast, the weak agreement of the FS with other measures supports previous concerns regarding its limited overlap with more comprehensive frailty assessment methods (Oviedo-Briones et al. 2021). In the nursing home setting, the FI identified more than twice as many frail individuals as the FS (94% *vs.* 41%) (Oviedo-Briones et al. 2021).

### 4.1.3 Ageing and frailty

Frailty is increasingly recognized as a clinical manifestation of biological ageing, characterized by multisystem dysregulation (Fried et al. 2001). Within the geroscience framework, frailty reflects underlying processes such as chronic inflammation, mitochondrial dysfunction, and cellular senescence (Johnson et al. 2019). The extent to which frailty instruments capture these mechanisms varies (O'Caoimh et al. 2021). According to my results, the FFP and FI best reflected ageing processes, as the pre-frail and frailty prevalences estimated by these instruments increased consistently with age. This was not the case for the FS, which might be explained by its nature as a totally

subjective instrument. Self-perceived frailty may not correlate with biological age as well as the FFP, which is partly based on the objective measurements of grip strength and gait speed. The high correlation of FI with age is not surprising given that each item of the FI needs to correlate with age to be selected for the FI (Dzando et al. 2025).

### 4.1.4 Frailty and physical performance

All four frailty measures were strongly associated with poor physical performance, as assessed by the SPPB, with the FFP demonstrating the strongest association. This finding is expected given that the FFP directly incorporates components of physical function, which are also central to the SPPB.

Interestingly, the FI, which is based on the accumulation of deficits across multiple domains, showed a weaker association with poor physical performance compared with the FFP. This may indicate that the FI captures broader aspects of health deterioration beyond musculoskeletal decline, including cognitive and metabolic impairments, which are not directly measured by the SPPB. The CFS and FS also demonstrated weaker associations, which may reflect the fact that they do not include physical performance items in their frailty assessment.

Stratified analyses revealed notable sex and age differences in the strength of associations between frailty and poor physical performance. Among women, the FS exhibited the second strongest association after the FFP, suggesting that self-reported fatigue and weight loss may be particularly relevant predictors of functional decline in women (Lindgren et al. 2025; Tirosh et al. 2015). In contrast, the FI was the second strongest predictor among men, implying that cumulative health deficits play a more prominent role in male frailty-related functional decline. These findings align with prior research showing sex differences, with women more often physically frail and men accumulating more chronic disease-related deficits (Sharma et al. 2024).

Age-stratified analyses showed that the ranking of frailty measures in modelling poor physical performance differed in relatively young (63-74 years) and old (75-89 years) individuals of the study population. In younger participants, the CFS had the second strongest association after the FFP, whereas in older participants, the FS was more strongly associated with poor physical performance than the CFS. This may suggest that the clinical judgement of the CFS performed better in younger individuals than the self-rated FS assessment, whereas in later life, the self-reported health deteriorations as captured by the FS become more predictive of functional decline.

### 4.1.5 Frailty and mortality

The FI had the strongest association with mortality in the total population and all subgroups except men, which can be explained by its construct as a cumulative deficit model based on age-related items. This close link to ageing makes it a robust instrument for all-cause mortality prediction. Surprisingly, the FS outperformed the FI among men. All individual items of the FS are strong predictors for mortality and possibly some of them predict mortality better among men than women. For all other frailty instruments, the association of frailty with mortality was stronger among women than men, which is consistent with previous literature (Hoogendijk et al. 2020; Kaskirbayeva et al. 2023; Nicaise et al. 2024; Walker et al. 2025). However, the biological reason for this is unknown.

Interestingly, despite its strong association with physical performance, the FFP had the weakest association with mortality in the total study population. This may be due to its emphasis on short-term functional decline rather than broader systemic health deterioration (Zucchelli et al. 2018). Nevertheless, the FFP ranked second in women for mortality prediction, highlighting again sex-differences in frailty-health outcome-associations and suggesting that, in women, functional decline is more prevalent and occurs earlier in life, with lower baseline physical reserves contributing to reduced resilience against loss of function.

The frailty instruments had comparable associations with mortality in relatively young (63-74 years) and old (75-89 years) individuals with exception of the FFP, which showed a stronger association of frailty with mortality among younger than older individuals. This likely reflects that declines in physical performance are less common earlier in life, making them more potent indicators of mortality risk when present.

These findings have important clinical and public health implications. First, my results confirm that different frailty measures capture distinct aspects of frailty, with implications for their use in research and clinical practice. The FFP appears to be the most relevant for identifying individuals at risk for poor physical performance, while the FI provides the best assessment method for long-term mortality risk. The FS is the easiest measure to use in practice because it can be self-assessed, and it performed surprisingly well for poor physical performance (2$^{nd}$ rank) and mortality prediction (3$^{rd}$ rank). However, its construct validity is questionable because it did not show a clear trend of increasing frailty with ageing like the other frailty instruments.

Sex- and age-specific differences in frailty associations underscore the need for tailored frailty assessment strategies. In men, self-reported health indicators (FS) may provide valuable insights into functional decline, whereas in women, deficit accumulation (FI) may be more informative.

Similarly, younger individuals may benefit from clinical assessments such as the CFS, while in older adults, subjective frailty measures (FS) may be more predictive of adverse outcomes. Routine frailty assessments, particularly using tools tailored to sex and age, could facilitate early interventions to mitigate adverse outcomes and improve the healthspan in ageing populations.

### 4.1.6 Strengths and limitations

There are several strengths, including the use of a large, well-characterized, population-based cohort that supports robust statistical analysis. The comprehensive assessment of four frailty measures alongside an objective physical performance test and complete mortality assessment allowed a robust analysis of high internal validity. Moreover, the comparison of four widely used frailty measures provides practical guidance for clinicians and researchers.

However, several limitations warrant consideration. Confounding cannot be excluded. I used a limited number of covariates in the models because it was the aim of this study to compare the predictive value of frailty instruments and not to precisely quantify causal associations of frailty with poor physical performance and mortality. The study population was healthier than the general population because institutionalized adults and less healthy non-participants were excluded. This does not affect the relative effect estimates (ORs and HRs) shown in this study but might have underestimated the (pre-)frailty prevalence. Some interaction and discrimination tests were borderline significant, but the point estimates were consistent with the main conclusions, likely reflecting limited power. Finally, the results on physical performance were cross-sectional and future studies should adopt a longitudinal design.

## 4.2 Study 2: Integrated proteomic and metabolomic profiling for frailty

### 4.2.1 Main findings and comparison with previous studies

In two large community-based prospective cohorts with standardized multi-omics profiling, I developed and externally validated the TRIM models for identifying and predicting frailty in older adults. Incorporation of 17 inflammation-related proteins and 9 NMR-derived metabolites into a conventional risk factor model consistently improved discrimination, calibration, reclassification, and clinical net benefit for both prevalent frailty and 8-year frailty incidence. An interactive web-based calculator was additionally developed to facilitate broader accessibility and translational application of the models.

My findings extend previous investigations linking proteomic and metabolomic data to frailty in several important ways. Large-scale biobank analyses have identified numerous cross-sectional associations between circulating proteins and frailty and have linked proteomic profiles to ageing-related outcomes (Sathyan et al. 2025; Xu et al. 2025). However, these studies have primarily focused on biomarker discovery rather than the development of parsimonious and externally validated prediction models. Similarly, metabolomic studies have characterized frailty-associated signatures but have less frequently quantified incremental predictive value beyond established risk factors using complementary performance metrics (Basisty et al. 2020; Liu et al. 2025b; Mutz et al. 2024; Tang et al. 2025; Tseng et al. 2024; Zhang et al. 2024). Furthermore, no previous study has integrated proteomics and metabolomics within a single validated framework. I derived a stable biomarker panel using bootstrap-based selection, modeled incident frailty under interval censoring with death as a competing risk, and evaluated performance across discrimination, calibration, reclassification, and decision-curve net benefit, with validation of predictive gains in an independent cohort. In line with TRIPOD principles (Efthimiou et al. 2024), optimism was minimized using bootstrap resampling with cross-validated penalization rather than split-sample approaches, which, together with external validation, support the stability and transportability of the model. This combination of methodological rigor and external reproducibility addresses the gap between molecular association and clinically interpretable prospective risk prediction. Notably, although proteomic biomarkers alone yielded larger performance gains than metabolomic markers alone, their combination produced greater improvement than either omics biomarker group individually, indicating complementary and partially orthogonal biological signals that together capture the multidimensional systems architecture underlying frailty.

### 4.2.2 Biological interpretation of identified biomarkers

Conceptually, frailty represents the clinical consolidation of distributed molecular dysregulation across organ systems. Circulating multi-omics biomarkers reflect inflammatory, metabolic, and endocrine stress that precedes overt functional decline. These biomarkers provide a mechanistic foundation for earlier risk identification beyond conventional clinical factors. Although biomarkers were selected from prevalent frailty identification, the ability of the same panel to predict incident frailty suggests that these signals capture biologically meaningful vulnerability rather than merely concurrent functional deficits. The comparatively weaker performance of the incident frailty-derived model may be explained by the much lower number of available incident frailty cases (N=250) compared to prevalent cases (N=1,863) in UKB.

Although causal inference cannot be established in observational studies, the selected biomarkers converge on three interrelated biological domains that are central to frailty pathophysiology. First, inflammatory and immune signaling pathways, reflected by *IL6* (Soysal et al. 2016) and *CCL20* (Li et al. 2023), indicate persistent low-grade immune activation and altered leukocyte trafficking, processes that have been repeatedly linked to diminished physiological reserve. Second, markers of metabolic and nutritional status, including glucose (Zaslavsky et al. 2016) and albumin (Danesh et al. 2004), capture systemic bioenergetic stress and nutritional inflammatory imbalance. Third, growth factor and endocrine-related pathways, including *FGF23* (Beben et al. 2016; Egund et al. 2023; Jovanovich et al. 2021; Paul et al. 2021; Sharma et al. 2021), *HGF* (Elgaabari et al. 2024), *TGFA* (Gerber et al. 2016), and *CDCP1* (Lehallier et al. 2019; Oh et al. 2023), suggest ongoing tissue remodeling and stress-related intercellular signaling. Together, these findings support the concept of inflammaging and suggest that integrated proteomic and metabolomic profiling captures biological vulnerability that precedes the onset of frailty.

### 4.2.3 Clinical and public health implications

From a translational and public health perspective, biologically informed identification of individuals at increased frailty risk is particularly relevant in settings where effective preventive interventions are available but resource-intensive. The decision curve analysis showed consistently greater net benefit for the omics-enhanced model than the conventional risk factor model across decision thresholds from 0.10 to 0.30 in both cohorts. These findings support the potential utility of the omics-enhanced model for identifying and prioritizing individuals at elevated frailty risk who may be considered for preventive and management strategies, including structured exercise programs, nutritional counselling, and medication reviews to reduce polypharmacy, as recommended in international clinical practice guidelines (Dent et al. 2019b). The developed TRIM models could also be used in randomized controlled trials testing preventive interventions against frailty by enriching trial populations with individuals at elevated risk of frailty. The identified proteins and metabolites could be added as secondary endpoints to evaluate whether the intervention is also effective against frailty on the molecular level. The web-based TRIM application may further support implementation of precision frailty prevention strategies by enabling individualized risk stratification in clinical and public health settings.

### 4.2.4 Strengths and limitations

Major strengths of this work include the use of two large population-based cohorts, external validation in an independent study population, and the application of rigorous statistical methods for biomarker selection and model evaluation.

Several limitations merit consideration. First, biomarkers were measured at a single time point at baseline, precluding assessment of temporal trajectories and limiting inference on biological dynamics. Second, differences in biospecimen type (plasma in UKB and serum in ESTHER), fasting status (few blood samples in UKB and most in ESTHER were collected after fasting overnight), and proteomic platforms (OLINK Explore *vs.* OLINK Target Inflammation panel) between cohorts may have introduced measurement heterogeneity, although harmonized log-transformation and z-score standardization, together with consistent findings across cohorts, support the robustness of the results. Third, participants were predominantly of European ancestry, and validation in more diverse populations is warranted. Finally, further studies are needed to evaluate implementation feasibility, cost-effectiveness, and whether the observed improvements in prediction performance translate into improved clinical decision-making in real-world settings.

## 4.3 Study 3: Metabolomics-based biological ageing clocks for mortality

### 4.3.1 Main findings

By integrating high-throughput NMR metabolomics profiling with conventional risk factors in two large prospective cohorts, I identified 68 circulating metabolomic biomarkers independently associated with 10-year all-cause mortality and validated them across both cohorts. These biomarkers mapped to key lipid metabolism and disease-related pathways. Additionally, I developed sex- and age-specific mortality prediction models, demonstrating consistent improvements in risk prediction across internal and external validation cohorts. The inclusion of metabolomic biomarkers significantly enhanced the mortality prediction, with the greatest benefit observed for short-term (5-year) mortality prediction. Moreover, I developed the MetaboMR clock1 and clock2, which captured the biological age well because each year of calculated age acceleration was associated with an 8% (MetaboMR clock1) or 9% (MetaboMR clock2) higher mortality in the external validation set than could be expected based on the chronological age alone.

### 4.3.2 Biological interpretation of mortality-associated metabolites

After adjusting for conventional risk factors, the metabolome-wide association analysis identified and externally validated key metabolomic biomarkers previously associated with mortality, including more than half of the 14 biomarkers reported by Deelen et al. (Deelen et al. 2019) and all four identified by Fischer et al. (Fischer et al. 2014). These include albumin, GlycA, citrate, histidine, and several lipid metabolites. From a mechanistic perspective, albumin is widely recognized as an indicator of systemic inflammation and nutritional status, reflecting overall protein homeostasis, which declines with ageing (Cabrerizo et al. 2015). Lower albumin levels have been associated with increased mortality risk, potentially due to their role in immune function, oxidative stress modulation, and chronic disease progression (Cabrerizo et al. 2015; Claudel and Verma 2025; Di Rosa et al. 2024). GlycA has been widely validated as a robust marker of systemic inflammation, with strong associations with cardiovascular and autoimmune diseases (Connelly et al. 2017; Crick et al. 2024; Guo et al. 2024; Ritchie et al. 2015), contributing to disease-related mortality (Chandler et al. 2016; Kettunen et al. 2018; Lawler et al. 2016; Seeling et al. 2023). Citrate, a key intermediate in the tricarboxylic acid (TCA) cycle, has been implicated in cancer metabolism, where altered energy utilization promotes tumor progression (Pavlova and Thompson 2016; Zhao et al. 2022). Similarly, lipids involved in cholesterol metabolism play a critical role in cardiovascular and metabolic disease-related mortality (Würtz et al. 2015). In contrast, histidine, another metabolite involved in central carbon metabolism, has been inversely associated with all-cause mortality, possibly due to its role in antioxidant activity, nitric oxide production, and immune regulation (Yan et al. 2022).

Using hierarchical clustering analysis, I identified four metabolic clusters linked to mortality, primarily composed of lipoprotein particles (HDL, IDL, LDL, and VLDL). KEGG-based pathway analysis revealed associations with lipid metabolism, energy metabolism, and systemic inflammation, highlighting their role in cardiovascular disease, neurodegenerative diseases and metabolic dysfunction. Additionally, the enrichment of pathways related to endocrine regulation (e.g., thyroid hormone synthesis and insulin signaling) suggests that metabolic perturbations may contribute to mortality through systemic effects (Bianco et al. 2014). Beyond the shared clusters, sex- and age-specific pathway enrichment was also observed. In younger men, enrichment of cancer metabolism, insulin resistance, and glucocorticoid pathways may underlie their higher midlife risk of metabolic and stress-related disorders as well as malignancy (Hanahan and Weinberg 2011). In older men, bile acid metabolism and viral infection pathways are consistent with age-related changes in lipid regulation and immune function (Hu et al. 2025; Nie et al. 2025). In younger

women, enrichment of branched-chain amino acid metabolism supports prior evidence linking elevated BCAAs to obesity and diabetes (Würtz et al. 2015). In older women, pathways related to inflammation and microbial metabolism align with the role of inflammaging and gut dysbiosis in late-life morbidity and mortality (Franceschi et al. 2018; O'Toole and Jeffery 2015). These subgroup-specific patterns provide biological plausibility for the observed heterogeneity in metabolite-mortality associations.

The observation that GlycA, the Omega-6/Omega-3 ratio, VLDL size, LA-pct, valine, and albumin showed large positive or negative weights in two or more subgroups, further underscores the central role of inflammation, energy metabolism, and lipid regulation in ageing biology and mortality risk. While this study provides strong observational evidence, additional functional studies and Mendelian randomization analyses are required to establish causal relationships and to evaluate whether interventions targeting these pathways could modify ageing trajectories and mortality risk.

### 4.3.3 Incremental value of metabolomics for mortality prediction

Integrating NMR-based metabolomic data with clinical indicators and sociodemographic factors to develop a mortality risk score is a feasible approach (Deelen et al. 2019; Fischer et al. 2014; Lian and Vardhanabhuti 2024). The improvement of the C-statistic by adding metabolites to a traditional risk factor model in this study was consistent with those reported from general population samples by Fischer et al. (+0.031 for 5-year mortality) (Fischer et al. 2014) and Deelen et al. (+0.065 and +0.040 for 5- and 10-year mortality, respectively) (Deelen et al. 2019). Similarly, Lian et al. found that incorporating 40 NMR-derived metabolomic biomarkers with conventional risk factors enhanced long-term risk prediction (AUC increase from 0.813 to 0.833) in 110,000 UKB participants (Lian and Vardhanabhuti 2024).

Additionally, the categorical NRI in this study was 14.7% for 10-year mortality and 19.1% for 5-year mortality prediction in the total study population of the external validation cohort. The IDI was also statistically significant in these analyses. Comparable risk reclassification statistics have been reported by previous studies (Deelen et al. 2019; Fischer et al. 2014; Lian and Vardhanabhuti 2024). However, these studies lacked external validation sets (Fischer et al. 2014; Lau et al. 2024; Lian and Vardhanabhuti 2024; Mutz et al. 2024; Yang et al. 2024; Zeng et al. 2024; Zhang et al. 2024), which usually result in lower C-statistics and reclassification statistics than in internal validation sets, limiting their generalizability.

Although some previous studies performed post hoc sex- or age-stratified analyses (Deelen et al. 2019; Wang et al. 2023; Yan et al. 2022), none has derived sex- and age-specific biomarker models

during model development. This distinction is critical, given biological differences in metabolic pathways across age and sex (Auro et al. 2014; da Silva and Schumacher 2021; Reicher et al. 2024; Sen et al. 2016). Thus, the results are novel and not directly comparable with previous research.

### 4.3.4 Metabolomics-based biological ageing clocks

I further explored the utility of metabolomic biomarkers in predicting the biological age by developing two metabolomics-based ageing clocks. The clock combining the selected metabolomic biomarkers and traditional risk factors (MetaboMR clock2) outperformed the clock based on metabolites only (MetaboMR clock1) and especially chronological age, when they were used to predict 10-year all-cause mortality alone without any other variables in the model. This shows that the MetaboMRclocks capture the biological age better than the chronological age. Surprisingly, the MetaboMRclock2 did not substantially outperform the MetaboMR clock1 in terms of the correlation with chronological age and the association of the metabolic age acceleration with all-cause mortality. This underscores the utility of metabolomic biomarkers in estimating the biological age acceleration. Biologically, the blood metabolome might mirror the physiological alterations caused by the traditional risk factors so well, that adding the traditional risk factors adds little additional risk information.

The observed 8% increase in 10-year all-cause mortality per year of biological age acceleration, as measured by MetaboMR clock1, indicates that each additional year of higher biological age compared to the chronological age is associated with an 8% higher risk of death over the following 10 years, highlighting its considerable predictive value. The risk for cancer mortality (9%) was comparable to that for all-cause mortality, whereas the risk of death from CVD was particularly strongly increased (13% for each year of higher biological age), which underscores the importance of metabolites in pathophysiological processes leading to CVD, while their role as indicators of an increased risk for cancer mortality should also not be neglected.

The clocks developed in this study could serve as clinical endpoints to evaluate the efficacy of such interventions or as motivational tools to encourage individuals to adopt lifestyle changes that improve their metabolic profile (e.g., dietary or physical activity interventions). To facilitate real-world applications, I developed a web-based calculator for the MetaboMR clocks that requires only the input of the metabolomic measurements for the MetaboMR clock1, and additional basic questionnaire-based mortality risk factor information for the MetaboMR clock2. This tool can readily be used in routine care.

The age- and sex-specific derivation approach distinguishes the MetaboMR clocks from other metabolomic ageing clocks previously developed in UKB. The clocks of Zhang et al. (Zhang et al. 2024), Mutz et al. (Mutz et al. 2024), Jia et al. (Jia et al. 2024) and Hao et al. (Hao et al. 2025) showed strong associations with ageing and mortality in UKB. However, these previous studies lacked external validation using large-scale metabolomic data. Comparative evaluation of UKB-derived metabolomic ageing clocks in external cohorts remains an important direction for future research.

### 4.3.5 Strengths and limitations

The strengths of this work include its prospective design, large sample size, extended follow-up, and independent external validation. The use of the same standardized NMR metabolomics platform across both cohorts ensured consistency of the measurements. However, certain limitations must be acknowledged. The NMR platform primarily quantifies lipid metabolites, limiting assessment of other metabolic pathways (Wishart et al. 2018). As an observational study, causality cannot be established, though several of Hill's criteria (temporality, strength, consistency, and biological plausibility) are met (Hill 2015). Furthermore, metabolomics data were collected only at baseline, and studies with repeated measurements are needed to assess how dynamic changes in metabolite profiles are related to mortality risk (Reicher et al. 2024; Shen et al. 2024). Despite differences in sample type (plasma in UKB *vs.* serum in ESTHER), my results remained highly concordant, emphasizing the reproducibility of identified biomarkers. This aligns with findings from Deelen et al., who observed consistent biomarker-mortality associations across 12 cohorts, regardless of sample type (Deelen et al. 2019). Moreover, although UKB had a higher baseline prevalence of treated dyslipidemia than ESTHER, sensitivity analyses excluding participants receiving lipid-lowering therapy at baseline produced materially unchanged results, implying that pharmacologic lipid modulation does not have a relevant impact on the performance of the MetaboMR clocks. Finally, as the models were derived and tested in a mostly Caucasian population from Western Europe, additional validation across diverse populations is needed.

## 4.4 Study 4: Inflammaging and proteomic ageing clocks for mortality

### 4.4.1 Summary of the key results

In this study, I developed and externally validated sex-specific inflammaging and proteomic ageing clocks for prediction of 10-year mortality from all causes. The clock combining proteins from

the broader proteomic panel and traditional risk factors (ProtAge2) consistently achieved the highest predictive performance across cohorts and outperformed models based on traditional risk factors only or models combining inflammatory biomarkers and traditional risk factors. The ProtAge2 clock captured the biological age well because each one-year increase in age acceleration (calculated as the residual from regressing ProtAge2 on chronological age), was associated with 22% higher all-cause mortality.

### 4.4.2 Biological basis of the proteomic and inflammaging ageing clocks

The superior predictive performance of the proteomic ageing clocks likely reflects their ability to capture biological processes beyond inflammation. Proteins selected from the broader proteomic panels were enriched in pathways spanning multiple biological domains, including PI3K-Akt signaling, proteoglycan signaling in cancer, MAPK signaling, HIF-1 signaling, focal adhesion, efferocytosis, lipid and atherosclerosis, and neurodegeneration-related processes. These pathways have been implicated in ageing-related dysregulation affecting multiple physiological systems, consistent with current systems-level models of ageing biology (Lehallier et al. 2019; López-Otín et al. 2013). In particular, the proteomic ageing clocks included important proteins involved in cardiovascular haemodynamic stress (*N-terminal prohormone of brain natriuretic peptide [NT-proBNP]*), neuroaxonal injury (*neurofilament light polypeptide [NEFL]*), matrix remodeling (*macrophage metalloelastase [MMP12]*), systemic stress signaling (*growth/differentiation factor 15 [GDF15]*), epithelial injury and fibrotic remodeling (*WAP four-disulfide core domain protein 2 [WFDC2]*), inflammaging-associated immune activation (*IL6*), and proliferative and regenerative signaling (*EGFR* and *amphiregulin [AREG]*), several of which have previously been implicated in biological ageing and age-related diseases (Argentieri et al. 2024; Lehallier et al. 2019; Wang et al. 2026).

By contrast, proteins selected from the targeted inflammation panel were enriched predominantly in cytokine-cytokine receptor interaction, IL-17 signaling, chemokine signaling, Toll-like receptor signaling, and TNF-associated pathways, consistent with the concept of inflammaging, in which chronic low-grade inflammation has been proposed as a downstream manifestation linking several hallmarks of ageing (Franceschi and Campisi 2014). Despite relying on substantially fewer proteins, the inflammatory biomarker extended traditional risk factor model predicted 10-year all-cause mortality significantly better than traditional risk factors alone. This shows that a relatively compact inflammatory protein network captures information relevant to multiple ageing-related biological processes beyond classic risk factors like smoking and obesity. This is consistent with

epidemiological evidence linking chronic inflammation to frailty, multimorbidity, and late-life mortality (Ershler and Keller 2000; Furman et al. 2019).

### 4.4.3 Sex differences and comparison with previous molecular ageing clocks

Despite some differences in the selected proteins and protein-protein interaction networks, the predictive performance of the inflammaging and proteomic ageing clocks was broadly comparable between men and women. Previous studies have consistently demonstrated sex differences in endocrine, immune, and metabolic ageing trajectories (Austad and Fischer 2016; Klein and Flanagan 2016; Márquez et al. 2020). Consistent with these observations, proteins selected from the broader proteomic panel in men showed higher enrichment of pathways related to PI3K-Akt signaling, proteoglycan-mediated remodeling, and *EGFR*-centered network connectivity than in women, suggesting greater involvement of growth-factor signaling, tissue remodeling, and stress-response pathways in men. By contrast, proteins selected in women showed relatively stronger enrichment of cytokine-mediated immune pathways, chemokine signaling, and inflammatory network organization, which is consistent with previous evidence that ageing in women is more closely linked to immune-inflammatory regulation and chronic inflammatory activation (Klein and Flanagan 2016). The comparable predictive performance observed across sexes therefore suggests that different molecular pathways may provide similar prognostic information regarding biological ageing and mortality risk (Dearden et al. 2018; Tower 2017; Tower et al. 2020). The different underlying biology of ageing of men and women also requires the development of sex specific biological ageing clocks, as done in my study.

Previous biological ageing clocks have been developed from proteomic, metabolomic, and DNA methylation data using various training targets, including chronological age, mortality risk, and ageing-related phenotypes (Argentieri et al. 2024; Belsky et al. 2022; Kuo et al. 2024; Levine et al. 2018; Lu et al. 2019; Peng et al. 2026). Several of these clocks have shown strong associations with mortality. For example, individuals in the highest 5% of proteomic age acceleration according to the chronological age-trained proteomic clock developed by Argentieri et al. had a 1.9-fold higher mortality risk than those whose proteomic and chronological ages were equivalent (Argentieri et al. 2024). In a corresponding analysis in my external validation cohort, individuals in the highest 5% of ProtAge2 acceleration had a 2.18-fold higher all-cause mortality risk (95% CI, 1.39–3.44). Kuo et al., who used mortality risk as a proxy for biological age, reported a 9.7% higher all-cause mortality risk per one-year increase in proteomic age (Kuo et al. 2024). Among DNA methylation-based measures, each 1-SD increment in DunedinPACE was associated with 26% and 65% higher mortality risks in the independent VA Normative Aging Study and

Framingham Heart Study Offspring cohort, respectively (Belsky et al. 2022). In my external validation, each one-year increase in ProtAge2 acceleration was associated with a 22% higher all-cause mortality risk, which was greater than the increase reported per year of proteomic age by Kuo et al. and similar in magnitude to that reported per SD of DunedinPACE in the VA Normative Aging Study, although the latter was measured on a different scale.

My clocks were developed using a mortality-informed approach, with proteins selected according to their predictive value for mortality separately in men and women. As all-cause mortality reflects the cumulative consequences of ageing across multiple physiological systems (Kennedy et al. 2014; López-Otín et al. 2023; Ukraintseva et al. 2021), it provides a relevant endpoint for biological age estimation. The strong associations of age acceleration derived from my clocks with all-cause, cardiovascular, and cancer mortality further support this approach.

### 4.4.4 Clinical implications and practical applications

The developed clocks can already be used for the assessment of the biological age in a research context. Such a validated measure of the biological age could be useful as an endpoint in clinical trials to assess the efficacy of interventions in geriatric medicine. I would recommend using the best performing ProtAge2 model. This would require a short questionnaire for the traditional risk factors and the measurement of the Olink Explore or Olink Explore HT panel in the total study population and subsequent z-transformation of the protein NPX values. The equations provided in **Appendix, Table A11** could then be applied to the standardized protein measurements and traditional risk factors to estimate biological age for each study participant.

When successful interventions in the developing field of longevity are available, the biological age assessment could also be useful in clinical practice as a screening tool, to identify those individuals who would profit most from these interventions or for monitoring the success of such treatments. For a potential application of the developed clocks in clinical practice, the 19 proteins for men and 15 proteins for women would need to be measured in absolute concentrations using sex-specific Olink Focus panels, which can measure up to 27 proteins. However, these Olink Focus panels need yet to be developed, the clock algorithms newly calibrated and made commercially available, which was beyond the scope of this project. Once available, this low number of up to 27 proteins can be measured at much lower costs compared to the large Olink Explore panels with > 3000 proteins.

### 4.4.5 Strengths and limitations

This study has several strengths. I developed sex-specific biological ageing clocks and tested them in an external validation set. In addition to discrimination, I systematically evaluated reclassification statistics and model calibration.

Nevertheless, several limitations should be considered when interpreting these findings. First, the observational design precludes causal inference regarding whether the identified proteins directly contribute to biological ageing or primarily reflect downstream consequences of ageing-related pathological processes. Second, proteomic measurements were available only at baseline, precluding assessment of longitudinal changes in biological ageing. Third, although external validation supported the robustness of the proposed clocks, both cohorts consisted predominantly of individuals of European ancestry, and results cannot be generalized to other populations.

## 4.5 Integrated interpretation across the four studies

Biological ageing is a multidimensional process characterized by the progressive decline of physiological reserve and resilience across multiple organ systems. Although chronological age remains the most widely used indicator of ageing, individuals of the same chronological age often differ substantially in their functional capacity, susceptibility to disease, and risk of adverse health outcomes. This heterogeneity reflects the complexity of biological ageing and highlights the need for complementary approaches that integrate clinical phenotypes with molecular characteristics. The four studies included in this dissertation collectively address this challenge by investigating biological ageing across different but interconnected levels of organization, ranging from clinical manifestations to molecular mechanisms and integrative biomarkers.

The first study established frailty as the clinical foundation of the dissertation. By comparing four widely used frailty instruments, it demonstrated that frailty is not a single entity but a multidimensional clinical construct. Different instruments captured distinct aspects of ageing-related decline, with varying abilities to identify impaired physical function and predict mortality. These findings indicate that no individual frailty measure can fully represent biological ageing. Instead, frailty should be viewed as the observable clinical expression of cumulative physiological dysregulation, reflecting the combined effects of multiple biological processes rather than a single pathological pathway.

While clinical frailty effectively identifies vulnerable individuals, it provides only limited insight into the biological mechanisms responsible for functional decline. This limitation motivated the second study, which integrated proteomic and metabolomic profiling to characterize the molecular

basis of frailty. The identification of reproducible inflammatory and metabolic signatures, together with the successful development and external validation of the TRIM models, demonstrates that clinically observable frailty is accompanied by measurable molecular alterations. Importantly, proteomic and metabolomic biomarkers contributed complementary information, suggesting that biological vulnerability is distributed across multiple interconnected biological systems rather than confined to a single pathway. These observations support the concept that molecular dysregulation precedes or accompanies the development of clinical frailty and provides a mechanistic explanation for inter-individual differences in ageing trajectories.

Building upon these findings, the third and fourth studies extended the focus from individual biomarkers to quantitative measures of biological ageing. Instead of examining isolated molecular features, metabolomic and proteomic data were integrated into biological ageing clocks capable of estimating biological age and predicting subsequent mortality. Although developed using different molecular platforms, both ageing clocks consistently demonstrated that individuals with accelerated biological ageing experienced substantially higher mortality than expected based on chronological age alone. These findings indicate that molecular ageing reflects cumulative biological damage across multiple physiological systems and captures important aspects of ageing that are not represented by chronological age. Rather than competing approaches, metabolomic and proteomic ageing clocks provide complementary perspectives on the biological processes underlying ageing and survival.

Taken together, the four studies describe a coherent continuum of biological ageing. Frailty represents the clinical manifestation of declining physiological reserve, whereas proteomic and metabolomic profiling characterize the molecular alterations that accompany this decline. Biological ageing clocks subsequently integrate these molecular alterations into quantitative indicators of biological age that provide prognostic information beyond conventional clinical assessment. This progression from phenotype to molecular mechanisms and finally to integrated biomarkers illustrates how different levels of biological organization contribute complementary information to the assessment of ageing.

Importantly, the present findings do not suggest that molecular biomarkers should replace established clinical assessment. Instead, clinical and molecular measures should be regarded as complementary components of a unified framework for evaluating biological ageing. Clinical assessment remains indispensable for identifying functional impairment and guiding patient management, whereas molecular profiling improves understanding of the biological processes underlying these phenotypes and enables earlier identification of individuals at increased risk before overt

functional decline becomes apparent. Integrating these complementary dimensions therefore provides a more comprehensive characterization of biological ageing than either approach alone.

Collectively, this dissertation supports a conceptual shift from viewing ageing primarily as a function of chronological time towards understanding ageing as a dynamic biological process that can be quantified at both the clinical and molecular levels. By bridging frailty assessment, multi-omics profiling, and molecular ageing clocks within a single framework, the present work provides a more comprehensive strategy for assessing biological ageing.

## 4.6 Overall strengths and limitations

A major strength of this dissertation lies in its coherent conceptual framework. Rather than addressing isolated research questions, the four studies were designed to investigate biological ageing from complementary perspectives spanning clinical phenotypes, molecular characteristics, and integrated biological ageing measures. This sequential approach enabled the findings from individual studies to be interpreted within a broader framework of biological ageing, providing a more comprehensive understanding than could be achieved by considering each study independently.

A second strength is the complementary use of phenotypic and high-dimensional molecular data. Clinical frailty assessment, proteomic profiling, metabolomic profiling, and molecular ageing clocks each capture different aspects of the ageing process. By combining these complementary dimensions, this dissertation moves beyond the traditional distinction between clinical and molecular epidemiology and illustrates how functional decline and molecular dysregulation can be jointly incorporated into the assessment of biological ageing. The consistency of findings across these different levels of investigation strengthens the overall interpretation of the results.

The dissertation also benefits from the use of large, well-characterized population-based cohorts and the emphasis on external validation throughout the molecular studies. Rather than focusing solely on biomarker discovery, the molecular analyses were designed to evaluate reproducibility, predictive performance, and potential clinical applicability in independent populations. This emphasis on validation increases the robustness of the proposed models and enhances their potential for future translation into epidemiological and clinical research.

Several limitations should nevertheless be considered when interpreting the findings as a whole. First, all studies were observational in nature. Although temporal relationships could be established for prospective analyses and extensive adjustment for potential confounding was performed where appropriate, causal relationships cannot be inferred. Consequently, the identified molecular

signatures and biological ageing measures should primarily be interpreted as indicators of biological ageing rather than direct determinants of adverse health outcomes.

Second, molecular profiling was performed only at baseline. Biological ageing is a dynamic process, whereas the present analyses relied on single measurements of circulating proteins and metabolites. Longitudinal studies incorporating repeated omics assessments will be necessary to determine how molecular ageing trajectories change over time and whether these measures are sufficiently responsive to interventions designed to promote healthy ageing.

Third, although the included cohorts differed in design and biospecimen collection, the study populations consisted predominantly of individuals of European ancestry. While external validation across independent cohorts supports the generalizability of the findings within similar populations, further validation in more diverse ethnic and environmental settings will be required before the proposed models can be considered broadly applicable.

Finally, mortality was selected as the principal outcome for the development and evaluation of the molecular ageing measures because it represents an objective and clinically important endpoint that integrates the cumulative consequences of biological ageing. Nevertheless, healthy ageing encompasses a broader spectrum of outcomes, including maintenance of physical function, cognitive performance, psychological well-being, resilience, and quality of life. Future studies should therefore evaluate whether the molecular signatures and biological ageing clocks developed in this dissertation similarly predict these multidimensional aspects of ageing.

## 4.7 Overall implications and future directions

The findings of this dissertation have implications for both ageing research and the assessment of ageing-related vulnerability. A central challenge is to determine whether biologically informed measures can improve risk stratification and clinical decision-making beyond information that is already readily available in routine practice. Addressing this question will require moving from biomarker discovery and prediction towards longitudinal evaluation, intervention studies, and assessment of real-world clinical utility.

From a clinical and public health perspective, the potential value of biologically informed ageing assessment lies primarily in improved risk stratification. Frailty measures are inexpensive and readily applicable, whereas molecular profiling may provide additional information on biological vulnerability before severe functional impairment becomes evident. Their relative value will therefore depend on the setting and purpose of assessment. In population screening or routine care,

simple clinical measures may remain the most practical first-line approach, while molecular information could be most useful when more refined risk stratification is required. Whether the additional prognostic information provided by omics-based measures is sufficient to justify their costs and complexity, however, remains to be established. Demonstrating predictive improvement is an important step, but it does not by itself establish clinical utility.

A particularly important direction for future research is the longitudinal assessment of biological ageing. The molecular analyses in this dissertation were based primarily on measurements obtained at a single time point, whereas ageing is inherently dynamic. Repeated proteomic and metabolomic measurements would allow investigation of whether molecular age acceleration remains stable, progresses gradually, or changes in response to alterations in health, lifestyle, or disease. Linking these trajectories to changes in frailty, physical function, multimorbidity, and other ageing-related outcomes could also clarify whether molecular ageing measures capture ongoing biological deterioration rather than simply reflecting accumulated risk at the time of measurement.

This longitudinal perspective is also essential for determining whether molecular ageing measures can be used to evaluate interventions. A biomarker that predicts mortality is not necessarily a valid surrogate marker of the ageing process or of treatment response. Before ageing clocks can be used as intermediate endpoints in trials of lifestyle, pharmacological, or geroprotective interventions, it will be necessary to establish whether changes in these measures correspond to meaningful changes in subsequent health outcomes. Intervention studies with repeated molecular profiling could therefore help distinguish biomarkers that primarily predict risk from those that are responsive to modification of the underlying ageing process.

Further work should also address the generalizability and biological interpretation of molecular ageing measures. Validation in populations with greater diversity in ancestry, socioeconomic circumstances, environmental exposures, and disease burden will be necessary to determine whether the associations and prediction models observed in the present cohorts are transportable to other settings. At the same time, integrating proteomic and metabolomic information with additional molecular layers, such as epigenomic or genomic data, may provide a more complete representation of ageing biology. Such analyses should be guided by biological plausibility and gains in reproducibility and prediction rather than by increasing molecular complexity alone.

Ultimately, translation into practice will depend on whether biologically informed assessment leads to decisions that improve health outcomes. Future studies should therefore move beyond measures of statistical performance and examine clinical utility, feasibility, cost-effectiveness, and

the consequences of using biological ageing measures to guide prevention or treatment. The longer-term goal is not to replace chronological age or established clinical assessment with molecular tests, but to determine when additional biological information meaningfully improves the identification and management of individuals at increased risk of age-related decline. Progress towards this goal would move ageing research from the description and prediction of biological ageing towards testing whether ageing trajectories can be monitored and, ultimately, modified.

## 4.8 Overall conclusions

Biological ageing is characterized by substantial interindividual heterogeneity that cannot be adequately captured by chronological age alone. The findings of this dissertation show that this heterogeneity is evident across multiple levels, from clinically manifest vulnerability to circulating molecular alterations and mortality-related biological age. At the clinical level, commonly used frailty instruments captured partly different aspects of vulnerability, with their performance depending on the outcome of interest, sex, and age. At the molecular level, proteomic and metabolomic profiles provided complementary information on frailty and improved its identification and prediction beyond conventional risk factors. Molecular ageing clocks further captured differences in mortality risk among individuals of similar chronological age, with broader proteomic profiling providing additional information beyond inflammation-focused biomarkers.

These findings indicate that no single clinical measure, biomarker profile, or molecular ageing clock provides a complete representation of biological ageing. The different approaches capture complementary manifestations of physiological decline and molecular dysregulation. Biological age should therefore be understood as a multidimensional construct that can be characterized from different, but related, perspectives rather than as a single quantity that is fully represented by any one measure.

Overall, this dissertation supports a multilevel approach to the study of biological ageing, progressing from the clinical assessment of frailty to its molecular characterization and the estimation of mortality-related biological age. This perspective provides a more differentiated understanding of why individuals of the same chronological age may experience markedly different trajectories of functional decline and survival. Future studies integrating longitudinal clinical and multi-omics measurements may further clarify how these dimensions evolve and interact over time and whether their combined assessment can improve the identification of individuals at increased risk of adverse ageing-related outcomes.

# 5 SUMMARY AND ZUSAMMENFASSUNG

## 5.1 English summary

Population ageing is accompanied by increasing burdens of chronic disease, functional decline, disability, and mortality. Although chronological age is widely used to characterize ageing, individuals of the same age vary considerably in their health, functional capacity, and mortality risk. Frailty captures clinically manifest vulnerability in older adults, whereas proteomic and metabolomic profiling offers the opportunity to characterize biological alterations underlying ageing. Combining these clinical and molecular perspectives may therefore provide a more comprehensive assessment of biological ageing and age-related health risks.

The overall aim of this dissertation was to improve the assessment of biological ageing by integrating clinical frailty measures with proteomic and metabolomic profiling in large population-based cohorts, and to evaluate their complementary value for characterizing biological ageing and predicting adverse ageing-related outcomes. Four complementary studies were conducted, progressing from the comparison of established clinical frailty measures to the identification of molecular signatures of frailty and the development of metabolomic and proteomic measures of biological ageing. The analyses were primarily based on data from the German ESTHER study and the UK Biobank, allowing molecular models developed in the UK Biobank to be externally validated in an independent population.

The first study compared four commonly used frailty instruments (the Clinical Frailty Scale, Fried Frailty Phenotype, Frailty Index, and FRAIL Scale) in 2,000 older adults. Frailty prevalence was similar across the four instruments, ranging from 18.1% to 20.2%, and frailty defined by each instrument was associated with both poor physical performance and all-cause mortality. However, their performance differed according to the outcome of interest. The Fried Frailty Phenotype showed the strongest association with poor physical performance, whereas the Frailty Index showed the strongest association with mortality. Performance also varied by sex and age, indicating that these instruments capture partly different aspects of age-related vulnerability and that their suitability depends on the population and outcome under investigation.

Given these differences in the clinical assessment of frailty, the second study investigated whether proteomic and metabolomic profiling could provide additional information for its identification and prediction. Among 19,151 UK Biobank participants, 17 inflammation-related proteins and 9 metabolites were selected for prevalent frailty. Six of these proteins and three metabolites also predicted incident frailty, with consistent directions of association in the UK Biobank and ESTHER study. In external validation, adding the selected biomarkers to conventional risk factors increased the area under the curve for prevalent frailty from 0.801 to 0.834 and the C-statistic for incident frailty from 0.699 to 0.746. The omics-enhanced models were well calibrated and provided greater net benefit than models based on conventional risk factors alone. Thus, molecular

profiling complemented the clinical characterization of frailty and improved the identification of prevalent frailty as well as the prediction of its future development.

Beyond identifying molecular signatures of frailty, metabolomic profiling was further used to develop quantitative measures of biological ageing and mortality risk. In the third study, nuclear magnetic resonance metabolomic data from 209,144 UK Biobank participants were analyzed and the resulting models were externally validated in 6,820 ESTHER participants. Sixty-eight metabolomic biomarkers, predominantly related to lipid metabolism and disease-related pathways, were associated with all-cause mortality in both cohorts. Sex- and age-specific models incorporating selected metabolites improved 10-year mortality prediction beyond traditional risk factors, with increases in the C-statistic ranging from 0.036 to 0.084 across subgroups. Each one-year increase in biological age acceleration derived from the metabolomics-based ageing clock was associated with 9% higher 10-year mortality in internal validation and 8% higher mortality in external validation. Metabolomic profiling therefore captured variation in mortality risk that was not reflected by chronological age or traditional risk factors alone.

The fourth study extended this approach to the proteome and specifically examined the contribution of inflammaging relative to broader proteomic alterations. Sex-specific ageing clocks were developed using proteomic data from 34,852 UK Biobank participants and externally validated in the ESTHER study. The inflammaging clocks included 11 proteins in men and 9 in women, whereas the broader proteomic ageing clocks included 19 and 15 proteins, respectively. In external validation, adding inflammatory proteins to conventional risk factors increased the C-statistic for 10-year all-cause mortality prediction from 0.720 to 0.760, while broader proteomic profiling increased it from 0.731 to 0.844. Each one-year increase in age acceleration derived from the inflammaging and broader proteomic ageing clocks was associated with 14% and 22% higher all-cause mortality, respectively. Both types of clocks captured interleukin-6-centered cytokine networks, while the broader proteomic clocks additionally incorporated proteins from a wider range of mortality-related pathways, including PI3K-Akt and proteoglycan signaling. These results suggest that inflammation represents an important component of biological ageing, but that broader proteomic alterations capture additional ageing-related processes relevant to mortality.

Taken together, the four studies characterize biological ageing from complementary clinical and molecular perspectives. Frailty instruments capture clinically manifest vulnerability but differ in their relationships with physical performance and mortality. Proteomic and metabolomic biomarkers provide additional information for identifying and predicting frailty, while molecular ageing clocks capture variation in mortality risk beyond chronological age and traditional risk factors. The progression from clinical frailty assessment to molecular characterization and mortality-based ageing measures illustrates that no single measure fully represents the complexity of biological ageing. Rather, clinical frailty measures and molecular biomarkers capture complementary dimensions of the ageing process. Their combined consideration may contribute to a more comprehensive assessment of ageing-related vulnerability and provide a basis for improved risk stratification in ageing research.

## 5.2 Deutsche Zusammenfassung

Die Alterung der Bevölkerung geht mit einer zunehmenden Belastung durch chronische Erkrankungen, funktionelle Einschränkungen, Behinderungen und Mortalität einher. Das chronologische Alter bildet die erhebliche Heterogenität des Gesundheitszustands und Mortalitätsrisikos gleichaltriger Personen jedoch nur unzureichend ab. Während Frailty die klinisch manifeste Vulnerabilität älterer Menschen erfasst, können proteomische und metabolomische Profile biologische Veränderungen des Alterns abbilden. Die Kombination klinischer und molekularer Ansätze könnte daher eine umfassendere Beurteilung des biologischen Alterns ermöglichen.

Ziel dieser Dissertation war es, die Erfassung des biologischen Alterns durch die Verknüpfung klinischer Frailty-Maße mit proteomischen und metabolomischen Profilen in großen populationsbasierten Kohorten zu verbessern und deren komplementären Nutzen für die Charakterisierung des biologischen Alterns und die Vorhersage altersassoziierter Gesundheitsereignisse zu untersuchen. Vier komplementäre Studien wurden hauptsächlich anhand von Daten der deutschen ESTHER-Studie und der UK Biobank durchgeführt. Dabei reichten die Untersuchungen vom Vergleich klinischer Frailty-Instrumente über die Identifizierung molekularer Frailty-Signaturen bis zur Entwicklung metabolomischer und proteomischer Maße des biologischen Alterns.

In der ersten Studie wurden die Clinical Frailty Scale, der Fried Frailty Phenotype, der Frailty Index und die FRAIL Scale bei 2.000 älteren Erwachsenen verglichen. Die Frailty-Prävalenz war mit 18,1–20,2 % bei allen Instrumenten ähnlich, und Frailty war unabhängig vom verwendeten Instrument sowohl mit eingeschränkter körperlicher Leistungsfähigkeit als auch mit Gesamtmortalität assoziiert. Der Fried Frailty Phenotype zeigte den stärksten Zusammenhang mit eingeschränkter körperlicher Leistungsfähigkeit, während der Frailty Index am stärksten mit Mortalität assoziiert war. Zudem variierte die Leistungsfähigkeit der Instrumente nach Geschlecht und Alter. Somit erfassen die Instrumente teilweise unterschiedliche Aspekte altersassoziierter Vulnerabilität, was bei ihrer Auswahl berücksichtigt werden sollte.

Die zweite Studie untersuchte, ob proteomische und metabolomische Biomarker zusätzliche Informationen für die Identifizierung und Vorhersage von Frailty liefern. Unter 19.151 Teilnehmenden der UK Biobank wurden 17 inflammationsbezogene Proteine und 9 Metaboliten für prävalente Frailty ausgewählt; 6 Proteine und 3 Metaboliten sagten zudem inzidente Frailty mit übereinstimmenden Assoziationsrichtungen in der UK Biobank und der ESTHER-Studie voraus. In der externen Validierung erhöhte die Ergänzung konventioneller Risikofaktoren um die ausgewählten Biomarker die Fläche unter der Kurve für prävalente Frailty von 0,801 auf 0,834 und die C-Statistik für inzidente Frailty von 0,699 auf 0,746. Proteomische und metabolomische Biomarker

verbesserten somit die Identifizierung und Vorhersage von Frailty über konventionelle Risikofaktoren hinaus.

Aufbauend auf diesen Ergebnissen wurden in der dritten Studie metabolomische Profile zur Entwicklung von Maßen des biologischen Alterns und zur Mortalitätsvorhersage genutzt. Hierzu wurden Daten von 209.144 Teilnehmenden der UK Biobank analysiert und die entwickelten Modelle bei 6.820 Teilnehmenden der ESTHER-Studie extern validiert. Insgesamt waren 68 metabolomische Biomarker, überwiegend aus dem Lipidstoffwechsel und krankheitsbezogenen Stoffwechselwegen, in beiden Kohorten mit Gesamtmortalität assoziiert. Geschlechts- und altersspezifische Modelle mit ausgewählten Metaboliten verbesserten die Vorhersage der 10-Jahres-Mortalität gegenüber traditionellen Risikofaktoren mit einer Zunahme der C-Statistik um 0,036–0,084. Jedes zusätzliche Jahr der metabolomisch bestimmten Beschleunigung des biologischen Alterns war mit einer um 9 % höheren 10-Jahres-Mortalität in der internen und einer um 8 % höheren Mortalität in der externen Validierung assoziiert.

In der vierten Studie wurde dieser Ansatz auf inflammationsbezogene und breitere proteomische Profile ausgeweitet. Geschlechtsspezifische Alterungsuhren wurden anhand von 34.852 Teilnehmenden der UK Biobank entwickelt und in der ESTHER-Studie extern validiert. Die Ergänzung konventioneller Risikofaktoren um inflammatorische Proteine erhöhte die C-Statistik für die 10-Jahres-Gesamtmortalität von 0,720 auf 0,760, während die Einbeziehung breiterer proteomischer Profile einen Anstieg von 0,731 auf 0,844 ergab. Jedes zusätzliche Jahr der anhand der Inflammaging- bzw. proteomischen Alterungsuhren bestimmten Beschleunigung des biologischen Alterns war mit einer um 14 % bzw. 22 % höheren Gesamtmortalität assoziiert. Beide Alterungsuhren erfassten inflammatorische Prozesse, während die breiteren proteomischen Profile zusätzliche mortalitätsrelevante Signalwege abbildeten und eine höhere Vorhersageleistung zeigten.

Zusammenfassend zeigen die vier Studien, dass klinische Frailty-Maße sowie proteomische und metabolomische Biomarker komplementäre Dimensionen des biologischen Alterns erfassen. Frailty-Instrumente unterscheiden sich hinsichtlich ihrer Beziehungen zu körperlicher Leistungsfähigkeit und Mortalität, während molekulare Biomarker zusätzliche Informationen für die Identifizierung und Vorhersage von Frailty liefern. Metabolomische und proteomische Alterungsmaße erfassen darüber hinaus Unterschiede im Mortalitätsrisiko, die durch das chronologische Alter und traditionelle Risikofaktoren allein nicht vollständig abgebildet werden. Die gemeinsame Betrachtung klinischer und molekularer Informationen kann daher zu einer umfassenderen Charakterisierung des biologischen Alterns beitragen.

# 6 REFERENCES


Abellan van Kan, G., Rolland, Y. M., Morley, J. E. and Vellas, B. (2008). **Frailty: toward a clinical definition**. J Am Med Dir Assoc *9 (2)*, 71-72, doi: 10.1016/j.jamda.2007.11.005.

Aebersold, R. and Mann, M. (2016). **Mass-spectrometric exploration of proteome structure and function**. Nature *537 (7620)*, 347-355, doi: 10.1038/nature19949.

Afilalo, J., Alexander, K. P., Mack, M. J., Maurer, M. S., Green, P., Allen, L. A., Popma, J. J., Ferrucci, L. and Forman, D. E. (2014). **Frailty assessment in the cardiovascular care of older adults**. J Am Coll Cardiol *63 (8)*, 747-762, doi: 10.1016/j.jacc.2013.09.070.

Altman, D. G. and Bland, J. M. (2005). **Standard deviations and standard errors**. BMJ *331 (7521)*, 903, doi: 10.1136/bmj.331.7521.903.

Amuthavalli Thiyagarajan, J., Mikton, C., Harwood, R. H., Gichu, M., Gaigbe-Togbe, V., Jhamba, T., Pokorna, D., Stoevska, V., Hada, R., Steffan, G. S., Liena, A., Rocard, E. and Diaz, T. (2022). **The UN Decade of healthy ageing: strengthening measurement for monitoring health and wellbeing of older people**. Age Ageing *51 (7)*, doi: 10.1093/ageing/afac147.

Antolini, L., Nam, B. H. and D'Agostino, R. B. (2004). **Inference on correlated discrimination measures in survival analysis: a nonparametric approach**. Commun Stat-Theor M *33 (9)*, 2117-2135, doi: 10.1081/sta-200026579.

Argentieri, M. A., Xiao, S., Bennett, D., Winchester, L., Nevado-Holgado, A. J., Ghose, U., Albukhari, A., Yao, P., Mazidi, M., Lv, J., Millwood, I., Fry, H., Rodosthenous, R. S., Partanen, J., Zheng, Z., Kurki, M., Daly, M. J., Palotie, A., Adams, C. J., Li, L., Clarke, R., Amin, N., Chen, Z. and van Duijn, C. M. (2024). **Proteomic aging clock predicts mortality and risk of common age-related diseases in diverse populations**. Nat Med *30 (9)*, 2450-2460, doi: 10.1038/s41591-024-03164-7.

Arteaga, A. S., Aguilar, L. T., González, J. T., Boza, A. S., Muñoz-Cruzado, V. D., Ciuró, F. P. and Ruíz, J. P. (2021). **Impact of frailty in surgical emergencies. A comparison of four frailty scales**. Eur J Trauma Emerg Surg *47 (5)*, 1613-1619, doi: 10.1007/s00068-020-01314-3.

Assarsson, E., Lundberg, M., Holmquist, G., Björkesten, J., Thorsen, S. B., Ekman, D., Eriksson, A., Rennel Dickens, E., Ohlsson, S., Edfeldt, G., Andersson, A. C., Lindstedt, P., Stenvang, J., Gullberg, M. and Fredriksson, S. (2014). **Homogenous 96-plex PEA immunoassay exhibiting high sensitivity, specificity, and excellent scalability**. PLoS One *9 (4)*, e95192, doi: 10.1371/journal.pone.0095192.

Auro, K., Joensuu, A., Fischer, K., Kettunen, J., Salo, P., Mattsson, H., Niironen, M., Kaprio, J., Eriksson, J. G., Lehtimäki, T., Raitakari, O., Jula, A., Tiitinen, A., Jauhiainen, M., Soininen, P., Kangas, A. J., Kähönen, M., Havulinna, A. S., Ala-Korpela, M., Salomaa, V., Metspalu, A. and Perola, M. (2014). **A metabolic view on menopause and ageing**. Nat Commun *5*, 4708, doi: 10.1038/ncomms5708.

Austad, S. N. and Fischer, K. E. (2016). **Sex Differences in Lifespan**. Cell Metab *23 (6)*, 1022-1033, doi: 10.1016/j.cmet.2016.05.019.

Ayala-Lopez, N. (2020). **Nuclear magnetic resonance metabolomics-enabled biomarker discovery for all-cause mortality**. Clin Chem *66 (2)*, 400, doi: 10.1093/clinchem/hvz014.

Balasubramanian, R., Paynter, N. P., Giulianini, F., Manson, J. E., Zhao, Y., Chen, J. C., Vitolins, M. Z., Albert, C. A., Clish, C. and Rexrode, K. M. (2020). **Metabolomic profiles associated with all-cause mortality in the Women's Health Initiative**. Int J Epidemiol *49 (1)*, 289-300, doi: 10.1093/ije/dyz211.

Basisty, N., Kale, A., Jeon, O. H., Kuehnemann, C., Payne, T., Rao, C., Holtz, A., Shah, S., Sharma, V., Ferrucci, L., Campisi, J. and Schilling, B. (2020). **A proteomic atlas of senescence-associated secretomes for aging biomarker development**. PLoS Biol *18 (1)*, e3000599, doi: 10.1371/journal.pbio.3000599.

Beard, J. R., Officer, A., de Carvalho, I. A., Sadana, R., Pot, A. M., Michel, J. P., Lloyd-Sherlock, P., Epping-Jordan, J. E., Peeters, G., Mahanani, W. R., Thiyagarajan, J. A. and Chatterji, S. (2016). **The World report on ageing and health: a policy framework for healthy ageing**. Lancet *387 (10033)*, 2145-2154, doi: 10.1016/s0140-6736(15)00516-4.

Beben, T., Ix, J. H., Shlipak, M. G., Sarnak, M. J., Fried, L. F., Hoofnagle, A. N., Chonchol, M., Kestenbaum, B. R., de Boer, I. H. and Rifkin, D. E. (2016). **Fibroblast growth factor-23 and frailty in elderly community-dwelling individuals: the Cardiovascular Health Study**. J Am Geriatr Soc *64 (2)*, 270-276, doi: 10.1111/jgs.13951.

Belga, S., Majumdar, S. R., Kahlon, S., Pederson, J., Lau, D., Padwal, R. S., Forhan, M., Bakal, J. A. and McAlister, F. A. (2016). **Comparing three different measures of frailty in medical inpatients: Multicenter prospective cohort study examining 30-day risk of readmission or death**. J Hosp Med *11 (8)*, 556-562, doi: 10.1002/jhm.2607.

Belsky, D. W., Caspi, A., Corcoran, D. L., Sugden, K., Poulton, R., Arseneault, L., Baccarelli, A., Chamarti, K., Gao, X., Hannon, E., Harrington, H. L., Houts, R., Kothari, M., Kwon, D., Mill, J., Schwartz, J., Vokonas, P., Wang, C., Williams, B. S. and Moffitt, T. E. (2022). **DunedinPACE, a DNA methylation biomarker of the pace of aging**. Elife *11*, doi: 10.7554/eLife.73420.

Benjamini, Y. and Hochberg, Y. (1995). **Controlling the false discovery rate: a practical and powerful approach to multiple testing**. J R Stat Soc B *57 (1)*, 289-300, doi: 10.1111/j.2517-6161.1995.tb02031.x.

Bianco, A. C., Anderson, G., Forrest, D., Galton, V. A., Gereben, B., Kim, B. W., Kopp, P. A., Liao, X. H., Obregon, M. J., Peeters, R. P., Refetoff, S., Sharlin, D. S., Simonides, W. S., Weiss, R. E. and Williams, G. R. (2014). **American thyroid association guide to investigating thyroid hormone economy and action in rodent and cell models**. Thyroid *24 (1)*, 88-168, doi: 10.1089/thy.2013.0109.

Bland, J. M. and Altman, D. G. (1996). **Transformations, means, and confidence intervals**. BMJ *312 (7038)*, 1079, doi: 10.1136/bmj.312.7038.1079.

Bountziouka, V., Nelson, C. P., Codd, V., Wang, Q., Musicha, C., Allara, E., Kaptoge, S., Di Angelantonio, E., Butterworth, A. S., Thompson, J. R., Curtis, E. M., Wood, A. M., Danesh, J. N., Harvey, N. C., Cooper, C. and Samani, N. J. (2022). **Association of shorter leucocyte telomere length with risk of frailty**. J Cachexia Sarcopenia Muscle *13 (3)*, 1741-1751, doi: 10.1002/jcsm.12971.

Burke, D. F., Bryant, P., Barrio-Hernandez, I., Memon, D., Pozzati, G., Shenoy, A., Zhu, W., Dunham, A. S., Albanese, P., Keller, A., Scheltema, R. A., Bruce, J. E., Leitner, A., Kundrotas, P., Beltrao, P. and Elofsson, A. (2023). **Towards a structurally resolved human protein interaction network**. Nat Struct Mol Biol *30 (2)*, 216-225, doi: 10.1038/s41594-022-00910-8.

Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L. T., Sharp, K., Motyer, A., Vukcevic, D., Delaneau, O., O'Connell, J., Cortes, A., Welsh, S., Young, A., Effingham, M., McVean, G., Leslie, S., Allen, N., Donnelly, P. and Marchini, J. (2018). **The UK Biobank resource with deep phenotyping and genomic data**. Nature *562 (7726)*, 203-209, doi: 10.1038/s41586-018-0579-z.

Cabrerizo, S., Cuadras, D., Gomez-Busto, F., Artaza-Artabe, I., Marín-Ciancas, F. and Malafarina, V. (2015). **Serum albumin and health in older people: Review and meta analysis**. Maturitas *81 (1)*, 17-27, doi: 10.1016/j.maturitas.2015.02.009.

Campisi, J., Kapahi, P., Lithgow, G. J., Melov, S., Newman, J. C. and Verdin, E. (2019). **From discoveries in ageing research to therapeutics for healthy ageing**. Nature *571 (7764)*, 183-192, doi: 10.1038/s41586-019-1365-2.

Cao, Z., Chen, H., Min, J. and Xu, C. (2025). **Proteomic landscape of multidimensional aging phenotypes**. Genome Med *17 (1)*, 122, doi: 10.1186/s13073-025-01558-x.

Cesari, M., Gambassi, G., van Kan, G. A. and Vellas, B. (2014). **The frailty phenotype and the frailty index: different instruments for different purposes**. Age Ageing *43 (1)*, 10-12, doi: 10.1093/ageing/aft160.

Cesari, M., Prince, M., Thiyagarajan, J. A., De Carvalho, I. A., Bernabei, R., Chan, P., Gutierrez-Robledo, L. M., Michel, J. P., Morley, J. E., Ong, P., Rodriguez Manas, L., Sinclair, A., Won, C. W., Beard, J. and Vellas, B. (2016). **Frailty: An Emerging Public Health Priority**. J Am Med Dir Assoc *17 (3)*, 188-192, doi: 10.1016/j.jamda.2015.12.016.

Chandler, P. D., Akinkuolie, A. O., Tobias, D. K., Lawler, P. R., Li, C., Moorthy, M. V., Wang, L., Duprez, D. A., Jacobs, D. R., Glynn, R. J., Otvos, J., Connelly, M. A., Post, W. S., Ridker, P. M., Manson, J. E., Buring, J. E., Lee, I. M. and Mora, S. (2016). **Association of N-linked glycoprotein acetyls and colorectal cancer incidence and mortality**. PLoS One *11 (11)*, e0165615, doi: 10.1371/journal.pone.0165615.

Checa-Lopez, M., Rodriguez-Laso, A., Carnicero, J. A., Solano-Jaurrieta, J. J., Saavedra Obermans, O., Sinclair, A., Landi, F., Scuteri, A., Álvarez-Bustos, A., Sepúlveda-Loyola, W. and Rodriguez-Manas, L. (2023). **Differential utility of various frailty diagnostic tools in non-geriatric hospital departments of several countries: A longitudinal study**. Eur J Clin Invest *53 (7)*, e13979, doi: 10.1111/eci.13979.

Chen, L. J., Sha, S., Brenner, H. and Schöttker, B. (2024). **Longitudinal associations of polypharmacy and frailty with major cardiovascular events and mortality among more than half a million middle-aged participants of the UK Biobank**. Maturitas *185*, 107998, doi: 10.1016/j.maturitas.2024.107998.

Chong, E., Ho, E., Baldevarona-Llego, J., Chan, M., Wu, L., Tay, L., Ding, Y. Y. and Lim, W. S. (2018). **Frailty in Hospitalized Older Adults: Comparing Different Frailty Measures in Predicting Short- and Long-term Patient Outcomes**. J Am Med Dir Assoc *19 (5)*, 450-457.e453, doi: 10.1016/j.jamda.2017.10.006.

Church, S., Rogers, E., Rockwood, K. and Theou, O. (2020). **A scoping review of the Clinical Frailty Scale**. BMC Geriatr *20 (1)*, 393, doi: 10.1186/s12877-020-01801-7.

Claudel, S. E. and Verma, A. (2025). **Albuminuria in cardiovascular, kidney, and metabolic disorders: a state-of-the-art review**. Circulation *151 (10)*, 716-732, doi: 10.1161/circulationaha.124.071079.

Clegg, A., Young, J., Iliffe, S., Rikkert, M. O. and Rockwood, K. (2013). **Frailty in elderly people**. Lancet *381 (9868)*, 752-762, doi: 10.1016/s0140-6736(12)62167-9.

Coenen, L., Lehallier, B., de Vries, H. E. and Middeldorp, J. (2023). **Markers of aging: Unsupervised integrated analyses of the human plasma proteome**. Front Aging *4*, 1112109, doi: 10.3389/fragi.2023.1112109.

Connelly, M. A., Otvos, J. D., Shalaurova, I., Playford, M. P. and Mehta, N. N. (2017). **GlycA, a novel biomarker of systemic inflammation and cardiovascular disease risk**. J Transl Med *15 (1)*, 219, doi: 10.1186/s12967-017-1321-6.

Crick, D. C. P., Khandaker, G. M., Halligan, S. L., Burgner, D., Mansell, T. and Fraser, A. (2024). **Comparison of the stability of glycoprotein acetyls and high sensitivity C-reactive protein as markers of chronic inflammation**. Immunology *171 (4)*, 497-512, doi: 10.1111/imm.13739.

da Silva, P. F. L. and Schumacher, B. (2021). **Principles of the molecular and cellular mechanisms of aging**. J Invest Dermatol *141 (4s)*, 951-960, doi: 10.1016/j.jid.2020.11.018.

Danesh, J., Wheeler, J. G., Hirschfield, G. M., Eda, S., Eiriksdottir, G., Rumley, A., Lowe, G. D., Pepys, M. B. and Gudnason, V. (2004). **C-reactive protein and other circulating markers of inflammation in the prediction of coronary heart disease**. N Engl J Med *350 (14)*, 1387-1397, doi: 10.1056/NEJMoa032804.

Dearden, L., Bouret, S. G. and Ozanne, S. E. (2018). **Sex and gender differences in developmental programming of metabolism**. Mol Metab *15*, 8-19, doi: 10.1016/j.molmet.2018.04.007.

Deelen, J., Kettunen, J., Fischer, K., van der Spek, A., Trompet, S., Kastenmüller, G., Boyd, A., Zierer, J., van den Akker, E. B., Ala-Korpela, M., Amin, N., Demirkan, A., Ghanbari, M., van Heemst, D., Ikram, M. A., van Klinken, J. B., Mooijaart, S. P., Peters, A., Salomaa, V., Sattar, N., Spector, T. D., Tiemeier, H., Verhoeven, A., Waldenberger, M., Würtz, P., Davey Smith, G., Metspalu, A., Perola, M., Menni, C., Geleijnse, J. M., Drenos, F., Beekman, M., Jukema, J. W., van Duijn, C. M. and Slagboom, P. E. (2019). **A metabolic profile of all-cause mortality risk identified in an observational study of 44,168 individuals**. Nat Commun *10 (1)*, 3346, doi: 10.1038/s41467-019-11311-9.

Delong, E. R., Delong, D. M. and Clarkepearson, D. I. (1988). **Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach**. Biometrics *44 (3)*, 837-845, doi: 10.2307/2531595.

Dent, E., Martin, F. C., Bergman, H., Woo, J., Romero-Ortuno, R. and Walston, J. D. (2019a). **Management of frailty: opportunities, challenges, and future directions**. Lancet *394 (10206)*, 1376-1386, doi: 10.1016/s0140-6736(19)31785-4.

Dent, E., Morley, J. E., Cruz-Jentoft, A. J., Woodhouse, L., Rodríguez-Mañas, L., Fried, L. P., Woo, J., Aprahamian, I., Sanford, A., Lundy, J., Landi, F., Beilby, J., Martin, F. C., Bauer, J. M., Ferrucci, L., Merchant, R. A., Dong, B., Arai, H., Hoogendijk, E. O., Won, C. W., Abbatecola, A., Cederholm, T., Strandberg, T., Gutiérrez Robledo, L. M., Flicker, L., Bhasin, S., Aubertin-Leheudre, M., Bischoff-Ferrari, H. A., Guralnik, J. M., Muscedere, J., Pahor, M., Ruiz, J., Negm, A. M., Reginster, J. Y., Waters, D. L. and Vellas, B. (2019b). **Physical frailty: ICFSR international clinical practice guidelines for identification and management**. J Nutr Health Aging *23 (9)*, 771-787, doi: 10.1007/s12603-019-1273-z.

Di, D., Zhou, H., Cui, Z., Zhang, J., Liu, Q., Yuan, T., Zhou, T., Luo, X., Ling, D. and Wang, Q. (2024). **Frailty phenotype as mediator between systemic inflammation and osteoporosis and fracture risks: A prospective study**. J Cachexia Sarcopenia Muscle *15 (3)*, 897-906, doi: 10.1002/jcsm.13447.

Di Rosa, M., Sabbatinelli, J., Giuliani, A., Carella, M., Magro, D., Biscetti, L., Soraci, L., Spannella, F., Fedecostante, M., Lenci, F., Tortato, E., Pimpini, L., Burattini, M., Cecchini, S., Cherubini, A., Bonfigli, A. R., Capalbo, M., Procopio, A. D., Balistreri, C. R. and Olivieri, F. (2024). **Correction: Inflammation scores based on C-reactive protein and albumin predict mortality in hospitalized older patients independent of the admission diagnosis**. Immun Ageing *21 (1)*, 79, doi: 10.1186/s12979-024-00483-8.

Ding, Y., Zuo, Y., Zhang, B., Fan, Y., Xu, G., Cheng, Z., Ma, S., Fang, S., Tian, A., Gao, D., Xu, X., Wang, Q., Jing, Y., Jiang, M., Xiong, M., Li, J., Han, Z., Sun, S., Wang, S., He, F., Yang, J., Qu, J., Zhang, W. and Liu, G. H. (2025). **Comprehensive human proteome profiles across a 50-year lifespan reveal aging trajectories and signatures**. Cell *188 (20)*, 5763-5784.e5726, doi: 10.1016/j.cell.2025.06.047.

Dzando, G., Ward, P. R., Mwanri, L., Moussa, R. K., Aheto, J. M. K. and Ambagtsheer, R. C. (2025). **Determining frailty index thresholds for older people across multiple countries in sub-Saharan Africa**. Commun Med (Lond) *5 (1)*, 237, doi: 10.1038/s43856-025-00952-1.

Efthimiou, O., Seo, M., Chalkou, K., Debray, T., Egger, M. and Salanti, G. (2024). **Developing clinical prediction models: a step-by-step guide**. BMJ *386*, e078276, doi: 10.1136/bmj-2023-078276.

Egund, L., Paulin, T. K., Ekstubbe, H., Bartosch, P. and Malmgren, L. (2023). **Longitudinal measurements of FGF23, sarcopenia, frailty and fracture in older community-dwelling women**. J Frailty Aging *12 (3)*, 166-174, doi: 10.14283/jfa.2023.22.

Eldjarn, G. H., Ferkingstad, E., Lund, S. H., Helgason, H., Magnusson, O. T., Gunnarsdottir, K., Olafsdottir, T. A., Halldorsson, B. V., Olason, P. I., Zink, F., Gudjonsson, S. A., Sveinbjornsson, G., Magnusson, M. I., Helgason, A., Oddsson, A., Halldorsson, G. H., Magnusson, M. K., Saevarsdottir, S., Eiriksdottir, T., Masson, G., Stefansson, H., Jonsdottir, I., Holm, H., Rafnar, T., Melsted, P., Saemundsdottir, J., Norddahl, G. L., Thorleifsson, G., Ulfarsson, M. O., Gudbjartsson, D. F., Thorsteinsdottir, U., Sulem, P. and Stefansson, K. (2023). **Large-scale plasma proteomics comparisons through genetics and disease associations**. Nature *622 (7982)*, 348-358, doi: 10.1038/s41586-023-06563-x.

Elgaabari, A., Imatomi, N., Kido, H., Nakashima, T., Okuda, S., Manabe, Y., Sawano, S., Mizunoya, W., Kaneko, R., Tanaka, S., Maeno, T., Matsuyoshi, Y., Seki, M., Kuwakado, S., Zushi, K., Daneshvar, N., Nakamura, M., Suzuki, T., Sunagawa, K., Anderson, J. E., Allen, R. E. and Tatsumi, R. (2024). **Age-related nitration/dysfunction of myogenic stem cell activator HGF**. Aging Cell *23 (2)*, e14041, doi: 10.1111/acel.14041.

Ersbøll, M., Valeur, N., Mogensen, U. M., Andersen, M. J., Møller, J. E., Velazquez, E. J., Hassager, C., Søgaard, P. and Køber, L. (2013). **Prediction of all-cause mortality and heart failure admissions from global left ventricular longitudinal strain in patients with acute myocardial infarction and preserved left ventricular ejection fraction**. J Am Coll Cardiol *61 (23)*, 2365-2373, doi: 10.1016/j.jacc.2013.02.061.

Ershler, W. B. and Keller, E. T. (2000). **Age-associated increased interleukin-6 gene expression, late-life diseases, and frailty**. Annu Rev Med *51*, 245-270, doi: 10.1146/annurev.med.51.1.245.

Ferrucci, L. and Fabbri, E. (2018). **Inflammageing: chronic inflammation in ageing, cardiovascular disease, and frailty**. Nat Rev Cardiol *15 (9)*, 505-522, doi: 10.1038/s41569-018-0064-2.

Ferrucci, L., Gonzalez-Freire, M., Fabbri, E., Simonsick, E., Tanaka, T., Moore, Z., Salimi, S., Sierra, F. and de Cabo, R. (2020). **Measuring biological aging in humans: A quest**. Aging Cell *19 (2)*, e13080, doi: 10.1111/acel.13080.

Fischer, K., Kettunen, J., Würtz, P., Haller, T., Havulinna, A. S., Kangas, A. J., Soininen, P., Esko, T., Tammesoo, M. L., Mägi, R., Smit, S., Palotie, A., Ripatti, S., Salomaa, V., Ala-Korpela, M., Perola, M. and Metspalu, A. (2014). **Biomarker profiling by nuclear magnetic resonance spectroscopy for the prediction of all-cause mortality: an observational study of 17,345 persons**. PLoS Med *11 (2)*, e1001606, doi: 10.1371/journal.pmed.1001606.

Franceschi, C. and Campisi, J. (2014). **Chronic inflammation (inflammaging) and its potential contribution to age-associated diseases**. J Gerontol A Biol Sci Med Sci *69 Suppl 1*, S4-9, doi: 10.1093/gerona/glu057.

Franceschi, C., Garagnani, P., Parini, P., Giuliani, C. and Santoro, A. (2018). **Inflammaging: a new immune-metabolic viewpoint for age-related diseases**. Nat Rev Endocrinol *14 (10)*, 576-590, doi: 10.1038/s41574-018-0059-4.

Franck, M., Tanner, K. T., Tennyson, R. L., Daunizeau, C., Ferrucci, L., Bandinelli, S., Trumble, B. C., Kaplan, H. S., Aronoff, J. E., Stieglitz, J., Kraft, T. S., Lea, A. J., Venkataraman, V. V., Wallace, I. J., Lim, Y. A. L., Ng, K. S., Yeong, J. P. S., Ho, R., Lim, X., Mehrjerd, A., Charalambous, E. G., Aiello, A. E., Pawelec, G., Franceschi, C., Hertel, J., Fülöp, T., Lemoine, M., Gurven, M. and Cohen, A. A. (2025). **Nonuniversality of inflammaging across human populations**. Nat Aging *5 (8)*, 1471-1480, doi: 10.1038/s43587-025-00888-0.

Fried, L. P., Ferrucci, L., Darer, J., Williamson, J. D. and Anderson, G. (2004). **Untangling the concepts of disability, frailty, and comorbidity: implications for improved targeting and care**. J Gerontol A Biol Sci Med Sci *59 (3)*, 255-263, doi: 10.1093/gerona/59.3.m255.

Fried, L. P., Tangen, C. M., Walston, J., Newman, A. B., Hirsch, C., Gottdiener, J., Seeman, T., Tracy, R., Kop, W. J., Burke, G. and McBurnie, M. A. (2001). **Frailty in older adults: evidence for a phenotype**. J Gerontol A Biol Sci Med Sci *56 (3)*, M146-156, doi: 10.1093/gerona/56.3.m146.

Furman, D., Campisi, J., Verdin, E., Carrera-Bastos, P., Targ, S., Franceschi, C., Ferrucci, L., Gilroy, D. W., Fasano, A., Miller, G. W., Miller, A. H., Mantovani, A., Weyand, C. M., Barzilai, N., Goronzy, J. J., Rando, T. A., Effros, R. B., Lucia, A., Kleinstreuer, N. and Slavich, G. M. (2019). **Chronic inflammation in the etiology of disease across the life span**. Nat Med *25 (12)*, 1822-1832, doi: 10.1038/s41591-019-0675-0.

Gafita, A., Calais, J., Grogan, T. R., Hadaschik, B., Wang, H., Weber, M., Sandhu, S., Kratochwil, C., Esfandiari, R., Tauber, R., Zeldin, A., Rathke, H., Armstrong, W. R., Robertson, A., Thin, P., D'Alessandria, C., Rettig, M. B., Delpassand, E. S., Haberkorn, U., Elashoff, D., Herrmann, K., Czernin, J., Hofman, M. S., Fendler, W. P. and Eiber, M. (2021). **Nomograms to predict outcomes after (177)Lu-PSMA therapy in men with metastatic castration-resistant prostate cancer: an international, multicentre, retrospective study**. Lancet Oncol *22 (8)*, 1115-1125, doi: 10.1016/s1470-2045(21)00274-6.

Ganna, A. and Ingelsson, E. (2015). **5 year mortality predictors in 498,103 UK Biobank participants: a prospective population-based study**. Lancet *386 (9993)*, 533-540, doi: 10.1016/s0140-6736(15)60175-1.

Gao, X., Huang, N., Jiang, M., Holleczek, B., Schöttker, B., Huang, T. and Brenner, H. (2022). **Mortality and morbidity risk prediction for older former smokers based on a score of smoking history: evidence from UK Biobank and ESTHER cohorts**. Age Ageing *51 (7)*, doi: 10.1093/ageing/afac154.

Gerber, P. A., Buhren, B. A., Schrumpf, H., Hevezi, P., Bölke, E., Sohn, D., Jänicke, R. U., Belum, V. R., Robert, C., Lacouture, M. E. and Homey, B. (2016). **Mechanisms of skin aging induced by EGFR inhibitors**. Support Care Cancer *24 (10)*, 4241-4248, doi: 10.1007/s00520-016-3254-7.

Geyer, P. E., Kulak, N. A., Pichler, G., Holdt, L. M., Teupser, D. and Mann, M. (2016). **Plasma Proteome Profiling to Assess Human Health and Disease**. Cell Syst *2 (3)*, 185-195, doi: 10.1016/j.cels.2016.02.015.

Glickman, M. E., Rao, S. R. and Schultz, M. R. (2014). **False discovery rate control is a recommended alternative to Bonferroni-type adjustments in health studies**. J Clin Epidemiol *67 (8)*, 850-857, doi: 10.1016/j.jclinepi.2014.03.012.

Grambsch, P. M. and Therneau, T. M. (1994). **Proportional hazards tests and diagnostics based on weighted residuals**. Biometrika *81 (3)*, 515-526.

Guo, Y., Liu, Q., Zheng, Z., Qing, M., Yao, T., Wang, B., Zhou, M., Wang, D., Ke, Q., Ma, J., Shan, Z. and Chen, W. (2024). **Genetic association of inflammatory marker GlycA with lung function and respiratory diseases**. Nat Commun *15 (1)*, 3751, doi: 10.1038/s41467-024-47845-w.

Hanahan, D. and Weinberg, R. A. (2011). **Hallmarks of cancer: the next generation**. Cell *144 (5)*, 646-674, doi: 10.1016/j.cell.2011.02.013.

Hanley, J. A. and McNeil, B. J. (1982). **The meaning and use of the area under a receiver operating characteristic (ROC) curve**. Radiology *143 (1)*, 29-36, doi: 10.1148/radiology.143.1.7063747.

Hao, M., Zhang, H., Wu, J., Huang, Y., Li, X., Wang, M., Wang, S., Wang, J., Chen, J., Jun Bao, Z., Jin, L., Wang, X., Hu, Z., Jiang, S. and Li, Y. (2025). **Gompertz Law-Based Biological Age (GOLD BioAge): A Simple and Practical Measurement of Biological Ageing to Capture Morbidity and Mortality Risks**. Adv Sci (Weinh), e01765, doi: 10.1002/advs.202501765.

Hasin, Y., Seldin, M. and Lusis, A. (2017). **Multi-omics approaches to disease**. Genome Biol *18 (1)*, 83, doi: 10.1186/s13059-017-1215-1.

Hill, A. B. (2015). **The environment and disease: association or causation? 1965**. J R Soc Med *108 (1)*, 32-37, doi: 10.1177/0141076814562718.

Hoogendijk, E. O., Afilalo, J., Ensrud, K. E., Kowal, P., Onder, G. and Fried, L. P. (2019). **Frailty: implications for clinical practice and public health**. Lancet *394 (10206)*, 1365-1375, doi: 10.1016/s0140-6736(19)31786-6.

Hoogendijk, E. O., Stenholm, S., Ferrucci, L., Bandinelli, S., Inzitari, M. and Cesari, M. (2020). **Operationalization of a frailty index among older adults in the InCHIANTI study: predictive ability for all-cause and cardiovascular disease mortality**. Aging Clin Exp Res *32 (6)*, 1025-1034, doi: 10.1007/s40520-020-01478-3.

Horvath, S. (2013). **DNA methylation age of human tissues and cell types**. Genome Biol *14 (10)*, R115, doi: 10.1186/gb-2013-14-10-r115.

Horvath, S. and Raj, K. (2018). **DNA methylation-based biomarkers and the epigenetic clock theory of ageing**. Nat Rev Genet *19 (6)*, 371-384, doi: 10.1038/s41576-018-0004-3.

Hu, M., Yang, J., Chen, Y., Zhang, Y., Liu, C. H., Feng, Y., Ye, B. and Li, D. (2025). **Celastrol ameliorates metabolic dysfunction associated steatohepatitis by regulating the CYP7B1-mediated alternative bile acid synthetic pathway**. Phytomedicine *147*, 157172, doi: 10.1016/j.phymed.2025.157172.

Huang, J., Weinstein, S. J., Moore, S. C., Derkach, A., Hua, X., Liao, L. M., Gu, F., Mondul, A. M., Sampson, J. N. and Albanes, D. (2018). **Serum metabolomic profiling of all-cause mortality: a prospective analysis in the alpha-tocopherol, beta-carotene cancer prevention (ATBC) study cohort**. Am J Epidemiol *187 (8)*, 1721-1732, doi: 10.1093/aje/kwy017.

Jia, X., Fan, J., Wu, X., Cao, X., Ma, L., Abdelrahman, Z., Zhao, F., Zhu, H., Bizzarri, D., Akker, E., Slagboom, P. E., Deelen, J., Zhou, D. and Liu, Z. (2024). **A novel metabolomic aging clock predicting health outcomes and its genetic and modifiable factors**. Adv Sci (Weinh) *11 (43)*, e2406670, doi: 10.1002/advs.202406670.

Jia, X., Gao, W., Hagelin, H., Zhao, Y., Zhang, J., Cao, X., Zhang, L., Wu, Y., Ma, L., Chen, L., Sun, L., Guo, H., Zhang, C., Jylhävä, J., Hu, Z., Hoogendijk, E. O., Hägg, S. and Liu, Z. (2026). **Plasma proteomic signature of frailty in 50,506 adults**. Cell Metab *38 (5)*, 1012-1028.e1016, doi: 10.1016/j.cmet.2026.02.013.

Jian, X., Sun, W., Zhang, J., Zhang, Q., Meng, X., Lu, H., Zheng, D., Wu, L. and Wang, Y. (2024). **Frailty mediating the causality between leucocyte telomere length and mortality: a cohort study of 440,551 UK Biobank participants**. Epma j *15 (1)*, 99-110, doi: 10.1007/s13167-024-00355-7.

Johnson, A. A., Shokhirev, M. N. and Shoshitaishvili, B. (2019). **Revamping the evolutionary theories of aging**. Ageing Res Rev *55*, 100947, doi: 10.1016/j.arr.2019.100947.

Johnson, A. A., Shokhirev, M. N., Wyss-Coray, T. and Lehallier, B. (2020). **Systematic review and analysis of human proteomics aging studies unveils a novel proteomic aging clock and identifies key processes that change with age**. Ageing Res Rev *60*, 101070, doi: 10.1016/j.arr.2020.101070.

Johnson, C. H., Ivanisevic, J. and Siuzdak, G. (2016). **Metabolomics: beyond biomarkers and towards mechanisms**. Nat Rev Mol Cell Biol *17 (7)*, 451-459, doi: 10.1038/nrm.2016.25.

Jovanovich, A., Ginsberg, C., You, Z., Katz, R., Ambrosius, W. T., Berlowitz, D., Cheung, A. K., Cho, M., Lee, A. K., Punzi, H., Rehman, S., Roumie, C., Supiano, M. A., Wright, C. B., Shlipak, M., Ix, J. H. and Chonchol, M. (2021). **FGF23, frailty, and falls in SPRINT**. J Am Geriatr Soc *69 (2)*, 467-473, doi: 10.1111/jgs.16895.

Julkunen, H., Cichońska, A., Tiainen, M., Koskela, H., Nybo, K., Mäkelä, V., Nokso-Koivisto, J., Kristiansson, K., Perola, M., Salomaa, V., Jousilahti, P., Lundqvist, A., Kangas, A. J., Soininen, P., Barrett, J. C. and Würtz, P. (2023). **Atlas of plasma NMR biomarkers for health and disease in 118,461 individuals from the UK Biobank**. Nat Commun *14 (1)*, 604, doi: 10.1038/s41467-023-36231-7.

Jylhävä, J., Pedersen, N. L. and Hägg, S. (2017). **Biological Age Predictors**. EBioMedicine *21*, 29-36, doi: 10.1016/j.ebiom.2017.03.046.

Kanehisa, M., Goto, S., Sato, Y., Kawashima, M., Furumichi, M. and Tanabe, M. (2014). **Data, information, knowledge and principle: back to metabolism in KEGG**. Nucleic Acids Res *42 (Database issue)*, D199-205, doi: 10.1093/nar/gkt1076.

Karczewski, K. J. and Snyder, M. P. (2018). **Integrative omics for health and disease**. Nat Rev Genet *19 (5)*, 299-310, doi: 10.1038/nrg.2018.4.

Kaskirbayeva, D., West, R., Jaafari, H., King, N., Howdon, D., Shuweihdi, F., Clegg, A. and Nikolova, S. (2023). **Progression of frailty as measured by a cumulative deficit index: A systematic review**. Ageing Res Rev *84*, 101789, doi: 10.1016/j.arr.2022.101789.

Keating, N. (2022). **A research framework for the United Nations Decade of Healthy Ageing (2021-2030)**. Eur J Ageing *19 (3)*, 775-787, doi: 10.1007/s10433-021-00679-7.

Kennedy, B. K., Berger, S. L., Brunet, A., Campisi, J., Cuervo, A. M., Epel, E. S., Franceschi, C., Lithgow, G. J., Morimoto, R. I., Pessin, J. E., Rando, T. A., Richardson, A., Schadt, E. E., Wyss-Coray, T. and Sierra, F. (2014). **Geroscience: linking aging to chronic disease**. Cell *159 (4)*, 709-713, doi: 10.1016/j.cell.2014.10.039.

Kerr, K. F., Wang, Z., Janes, H., McClelland, R. L., Psaty, B. M. and Pepe, M. S. (2014). **Net reclassification indices for evaluating risk prediction instruments: a critical review**. Epidemiology *25 (1)*, 114-121, doi: 10.1097/ede.0000000000000018.

Kettunen, J., Ritchie, S. C., Anufrieva, O., Lyytikäinen, L. P., Hernesniemi, J., Karhunen, P. J., Kuukasjärvi, P., Laurikka, J., Kähönen, M., Lehtimäki, T., Havulinna, A. S., Salomaa, V., Männistö, S., Ala-Korpela, M., Perola, M., Inouye, M. and Würtz, P. (2018). **Biomarker glycoprotein acetyls is associated with the risk of a wide spectrum of incident diseases and stratifies mortality risk in angiography patients**. Circ-Genom Precis Me *11 (11)*, e002234, doi: 10.1161/circgen.118.002234.

Kim, D. H. and Rockwood, K. (2024). **Frailty in older adults**. N Engl J Med *391 (6)*, 538-548, doi: 10.1056/NEJMra2301292.

Klein, S. L. and Flanagan, K. L. (2016). **Sex differences in immune responses**. Nat Rev Immunol *16 (10)*, 626-638, doi: 10.1038/nri.2016.90.

Kuo, C. L., Chen, Z., Liu, P., Pilling, L. C., Atkins, J. L., Fortinsky, R. H., Kuchel, G. A. and Diniz, B. S. (2024). **Proteomic aging clock (PAC) predicts age-related outcomes in middle-aged and older adults**. Aging Cell *23 (8)*, e14195, doi: 10.1111/acel.14195.

Langfelder, P. and Horvath, S. (2008). **WGCNA: an R package for weighted correlation network analysis**. BMC Bioinformatics *9*, 559, doi: 10.1186/1471-2105-9-559.

Lau, C. E., Manou, M., Markozannes, G., Ala-Korpela, M., Ben-Shlomo, Y., Chaturvedi, N., Engmann, J., Gentry-Maharaj, A., Herzig, K. H., Hingorani, A., Järvelin, M. R., Kähönen, M., Kivimäki, M., Lehtimäki, T., Marttila, S., Menon, U., Munroe, P. B., Palaniswamy, S., Providencia, R., Raitakari, O., Schmidt, A. F., Sebert, S., Wong, A., Vineis, P., Tzoulaki, I. and Robinson, O. (2024). **NMR metabolomic modeling of age and lifespan: a multicohort analysis**. Aging Cell *23 (7)*, e14164, doi: 10.1111/acel.14164.

Lawler, P. R., Akinkuolie, A. O., Chandler, P. D., Moorthy, M. V., Vandenburgh, M. J., Schaumberg, D. A., Lee, I. M., Glynn, R. J., Ridker, P. M., Buring, J. E. and Mora, S. (2016). **Circulating N-linked glycoprotein acetyls and longitudinal mortality risk**. Circ Res *118 (7)*, 1106-1115, doi: 10.1161/circresaha.115.308078.

Lehallier, B., Gate, D., Schaum, N., Nanasi, T., Lee, S. E., Yousef, H., Moran Losada, P., Berdnik, D., Keller, A., Verghese, J., Sathyan, S., Franceschi, C., Milman, S., Barzilai, N. and Wyss-Coray, T. (2019). **Undulating changes in human plasma proteome profiles across the lifespan**. Nat Med *25 (12)*, 1843-1850, doi: 10.1038/s41591-019-0673-2.

Lehallier, B., Shokhirev, M. N., Wyss-Coray, T. and Johnson, A. A. (2020). **Data mining of human plasma proteins generates a multitude of highly predictive aging clocks that reflect different aspects of aging**. Aging Cell *19 (11)*, e13256, doi: 10.1111/acel.13256.

Levine, M. E. (2013). **Modeling the rate of senescence: can estimated biological age predict mortality more accurately than chronological age?** J Gerontol A Biol Sci Med Sci *68 (6)*, 667-674, doi: 10.1093/gerona/gls233.

Levine, M. E., Lu, A. T., Quach, A., Chen, B. H., Assimes, T. L., Bandinelli, S., Hou, L., Baccarelli, A. A., Stewart, J. D., Li, Y., Whitsel, E. A., Wilson, J. G., Reiner, A. P., Aviv, A., Lohman, K., Liu, Y., Ferrucci, L. and Horvath, S. (2018). **An epigenetic biomarker of aging for lifespan and healthspan**. Aging (Albany NY) *10 (4)*, 573-591, doi: 10.18632/aging.101414.

Li, J. J., Jiang, S., Zhu, M. L., Liu, X. H., Sun, X. H. and Zhao, S. Q. (2021). **Comparison of Three Frailty Scales for Prediction of Adverse Outcomes among Older Adults: A Prospective Cohort Study**. J Nutr Health Aging *25 (4)*, 419-424, doi: 10.1007/s12603-020-1534-x.

Li, X., Li, C., Zhang, W., Wang, Y., Qian, P. and Huang, H. (2023). **Inflammation and aging: signaling pathways and intervention therapies**. Signal Transduct Target Ther *8 (1)*, 239, doi: 10.1038/s41392-023-01502-8.

Lian, J. and Vardhanabhuti, V. (2024). **Metabolic biomarkers using nuclear magnetic resonance metabolomics assay for the prediction of aging-related disease risk and mortality: a prospective, longitudinal, observational, cohort study based on the UK Biobank**. Geroscience *46 (2)*, 1515-1526, doi: 10.1007/s11357-023-00918-y.

Lindgren, L. H., Hammer, N. M., Flurey, C. A., Jensen, K. V., Andersen, L. and Esbensen, B. A. (2025). **Gender differences in illness acceptance and coping strategies among patients with rheumatoid arthritis, psoriatic arthritis, and axial spondyloarthritis: a cross-sectional survey study**. Rheumatol Int *45 (3)*, 52, doi: 10.1007/s00296-025-05805-7.

Liu, R., Yang, S., Zhong, X., Zhu, Z., Huang, W. and Wang, W. (2025a). **Metabolomic signature of retinal ageing, polygenetic susceptibility, and major health outcomes**. Br J Ophthalmol *109 (5)*, 619-627, doi: 10.1136/bjo-2024-325846.

Liu, W. S., You, J., Chen, S. D., Zhang, Y., Feng, J. F., Xu, Y. M., Yu, J. T. and Cheng, W. (2025b). **Plasma proteomics identify biomarkers and undulating changes of brain aging**. Nat Aging *5 (1)*, 99-112, doi: 10.1038/s43587-024-00753-6.

López-Otín, C., Blasco, M. A., Partridge, L., Serrano, M. and Kroemer, G. (2013). **The hallmarks of aging**. Cell *153 (6)*, 1194-1217, doi: 10.1016/j.cell.2013.05.039.

López-Otín, C., Blasco, M. A., Partridge, L., Serrano, M. and Kroemer, G. (2023). **Hallmarks of aging: An expanding universe**. Cell *186 (2)*, 243-278, doi: 10.1016/j.cell.2022.11.001.

Lu, A. T., Quach, A., Wilson, J. G., Reiner, A. P., Aviv, A., Raj, K., Hou, L., Baccarelli, A. A., Li, Y., Stewart, J. D., Whitsel, E. A., Assimes, T. L., Ferrucci, L. and Horvath, S. (2019). **DNA methylation GrimAge strongly predicts lifespan and healthspan**. Aging (Albany NY) *11 (2)*, 303-327, doi: 10.18632/aging.101684.

Lundberg, M., Eriksson, A., Tran, B., Assarsson, E. and Fredriksson, S. (2011). **Homogeneous antibody-based proximity extension assays provide sensitive and specific detection of low-abundant proteins in human blood**. Nucleic Acids Res *39 (15)*, e102, doi: 10.1093/nar/gkr424.

Ma, L. Z., Liu, W. S., He, Y., Zhang, Y., You, J., Feng, J. F., Tan, L., Cheng, W. and Yu, J. T. (2026). **Plasma proteomics identify novel biomarkers and dynamic patterns of biological aging**. J Adv Res *80*, 691-704, doi: 10.1016/j.jare.2025.05.004.

Ma, T., He, L., Luo, Y., Fu, D., Huang, J., Zhang, G., Cheng, X. and Bai, Y. (2023). **Frailty, an Independent Risk Factor in Progression Trajectory of Cardiometabolic Multimorbidity: A Prospective Study of UK Biobank**. J Gerontol A Biol Sci Med Sci *78 (11)*, 2127-2135, doi: 10.1093/gerona/glad125.

Macdonald-Dunlop, E., Taba, N., Klarić, L., Frkatović, A., Walker, R., Hayward, C., Esko, T., Haley, C., Fischer, K., Wilson, J. F. and Joshi, P. K. (2022). **A catalogue of omics biological ageing clocks reveals substantial commonality and associations with disease risk**. Aging (Albany NY) *14 (2)*, 623-659, doi: 10.18632/aging.203847.

Mak, J. K. L., Kananen, L., Qin, C., Kuja-Halkola, R., Tang, B., Lin, J., Wang, Y., Jääskeläinen, T., Koskinen, S., Lu, Y., Magnusson, P. K. E., Hägg, S. and Jylhävä, J. (2023). **Unraveling the metabolic underpinnings of frailty using multicohort observational and Mendelian randomization analyses**. Aging Cell *22 (8)*, e13868, doi: 10.1111/acel.13868.

Mak, J. K. L., Qin, C., Krüger, M., Kuukka, A., Hägg, S., Lin, J. and Jylhävä, J. (2025). **Large-scale genome-wide analyses with proteomics integration reveal novel loci and biological insights into frailty**. Nat Aging *5 (8)*, 1589-1600, doi: 10.1038/s43587-025-00925-y.

Malmstrom, T. K., Miller, D. K. and Morley, J. E. (2014). **A comparison of four frailty models**. J Am Geriatr Soc *62 (4)*, 721-726, doi: 10.1111/jgs.12735.

Márquez, E. J., Chung, C. H., Marches, R., Rossi, R. J., Nehar-Belaid, D., Eroglu, A., Mellert, D. J., Kuchel, G. A., Banchereau, J. and Ucar, D. (2020). **Sexual-dimorphism in human immune system aging**. Nat Commun *11 (1)*, 751, doi: 10.1038/s41467-020-14396-9.

Miller, M. D., Paradis, C. F., Houck, P. R., Mazumdar, S., Stack, J. A., Rifai, A. H., Mulsant, B. and Reynolds, C. F., 3rd (1992). **Rating chronic medical illness burden in geropsychiatric practice and research: application of the Cumulative Illness Rating Scale**. Psychiatry Res *41 (3)*, 237-248, doi: 10.1016/0165-1781(92)90005-n.

Mitnitski, A. B., Mogilner, A. J. and Rockwood, K. (2001). **Accumulation of deficits as a proxy measure of aging**. ScientificWorldJournal *1*, 323-336, doi: 10.1100/tsw.2001.58.

Moreno-Ariño, M., Torrente Jiménez, I., Cartanyà Gutiérrez, A., Oliva Morera, J. C. and Comet, R. (2020). **Assessing the strengths and weaknesses of the Clinical Frailty Scale through correlation with a frailty index**. Aging Clin Exp Res *32 (11)*, 2225-2232, doi: 10.1007/s40520-019-01450-w.

Morley, J. E., Malmstrom, T. K. and Miller, D. K. (2012). **A simple frailty questionnaire (FRAIL) predicts outcomes in middle aged African Americans**. J Nutr Health Aging *16 (7)*, 601-608, doi: 10.1007/s12603-012-0084-2.

Mutz, J., Iniesta, R. and Lewis, C. M. (2024). **Metabolomic age (MileAge) predicts health and life span: A comparison of multiple machine learning algorithms**. Sci Adv *10 (51)*, eadp3743, doi: 10.1126/sciadv.adp3743.

Nicaise, E. H., Palmateer, G., Schmeusser, B. N., Futral, C., Liu, Y., Goyal, S., Nabavizadeh, R., Kooby, D. A., Maithel, S. K., Sweeney, J. F., Sarmiento, J. M., Ogan, K. and Master, V. A. (2024). **Differences in preoperative frailty assessment of surgical candidates by sex, age, and race**. Surg Open Sci *19*, 172-177, doi: 10.1016/j.sopen.2024.05.003.

Nie, S., Cao, L., Su, L., Zhou, S., Li, Y., Chen, R., Luo, F., Gao, Q., Lin, Y., Guo, Z., Xu, X. and Wang, G. (2025). **Prophylactic antiviral therapy and all-cause mortality in cancer patients with hepatitis B e antigen-negative chronic hepatitis B virus infection receiving immunosuppressive therapy**. J Natl Cancer Inst *117 (8)*, 1605-1613, doi: 10.1093/jnci/djaf032.

O'Caoimh, R., Sezgin, D., O'Donovan, M. R., Molloy, D. W., Clegg, A., Rockwood, K. and Liew, A. (2021). **Prevalence of frailty in 62 countries across the world: a systematic review and meta-analysis of population-level studies**. Age Ageing *50 (1)*, 96-104, doi: 10.1093/ageing/afaa219.

O'Toole, P. W. and Jeffery, I. B. (2015). **Gut microbiota and aging**. Science *350 (6265)*, 1214-1215, doi: 10.1126/science.aac8469.

Oh, H. S., Rutledge, J., Nachun, D., Pálovics, R., Abiose, O., Moran-Losada, P., Channappa, D., Urey, D. Y., Kim, K., Sung, Y. J., Wang, L., Timsina, J., Western, D., Liu, M., Kohlfeld, P., Budde, J., Wilson, E. N., Guen, Y., Maurer, T. M., Haney, M., Yang, A. C., He, Z., Greicius, M. D., Andreasson, K. I., Sathyan, S., Weiss, E. F., Milman, S., Barzilai, N., Cruchaga, C., Wagner, A. D., Mormino, E., Lehallier, B., Henderson, V. W., Longo, F. M., Montgomery, S. B. and Wyss-Coray, T. (2023). **Organ aging signatures in the plasma proteome track health and disease**. Nature *624 (7990)*, 164-172, doi: 10.1038/s41586-023-06802-1.

Op Het Veld, L. P. M., Beurskens, A., de Vet, H. C. W., van Kuijk, S. M. J., Hajema, K., Kempen, G. and van Rossum, E. (2019). **The ability of four frailty screening instruments to predict mortality, hospitalization and dependency in (instrumental) activities of daily living**. Eur J Ageing *16 (3)*, 387-394, doi: 10.1007/s10433-019-00502-4.

Oviedo-Briones, M., Laso Á, R., Carnicero, J. A., Cesari, M., Grodzicki, T., Gryglewska, B., Sinclair, A., Landi, F., Vellas, B., Checa-López, M. and Rodriguez-Mañas, L. (2021). **A Comparison of Frailty Assessment Instruments in Different Clinical and Social Care Settings: The Frailtools Project**. J Am Med Dir Assoc *22 (3)*, 607.e607-607.e612, doi: 10.1016/j.jamda.2020.09.024.

Oviedo-Briones, M., Rodríguez-Laso, Á., Carnicero, J. A., Gryglewska, B., Sinclair, A. J., Landi, F., Vellas, B., Rodríguez Artalejo, F., Checa-López, M. and Rodriguez-Mañas, L. (2022). **The ability of eight frailty instruments to identify adverse outcomes across different settings: the FRAILTOOLS project**. J Cachexia Sarcopenia Muscle *13 (3)*, 1487-1501, doi: 10.1002/jcsm.12990.

Park, J., Bakoyannis, G. and Yiannoutsos, C. T. (2019). **Semiparametric competing risks regression under interval censoring using the R package intccr**. Comput Methods Programs Biomed *173*, 167-176, doi: 10.1016/j.cmpb.2019.03.002.

Patti, G. J., Yanes, O. and Siuzdak, G. (2012). **Innovation: Metabolomics: the apogee of the omics trilogy**. Nat Rev Mol Cell Biol *13 (4)*, 263-269, doi: 10.1038/nrm3314.

Paul, S., Wong, M., Akhabue, E., Mehta, R. C., Kramer, H., Isakova, T., Carnethon, M. R., Wolf, M. and Gutiérrez, O. M. (2021). **Fibroblast growth factor 23 and incident cardiovascular disease and mortality in middle-aged adults**. J Am Heart Assoc *10 (16)*, e020196, doi: 10.1161/jaha.120.020196.

Pavasini, R., Guralnik, J., Brown, J. C., di Bari, M., Cesari, M., Landi, F., Vaes, B., Legrand, D., Verghese, J., Wang, C., Stenholm, S., Ferrucci, L., Lai, J. C., Bartes, A. A., Espaulella, J., Ferrer, M., Lim, J. Y., Ensrud, K. E., Cawthon, P., Turusheva, A., Frolova, E., Rolland, Y., Lauwers, V., Corsonello, A., Kirk, G. D., Ferrari, R., Volpato, S. and Campo, G. (2016). **Short Physical Performance Battery and all-cause mortality: systematic review and meta-analysis**. BMC Med *14 (1)*, 215, doi: 10.1186/s12916-016-0763-7.

Pavlova, N. N. and Thompson, C. B. (2016). **The emerging hallmarks of cancer metabolism**. Cell Metab *23 (1)*, 27-47, doi: 10.1016/j.cmet.2015.12.006.

Pencina, M. J., D'Agostino, R. B., Sr., D'Agostino, R. B., Jr. and Vasan, R. S. (2008). **Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond**. Stat Med *27 (2)*, 157-172; discussion 207-112, doi: 10.1002/sim.2929.

Pencina, M. J., D'Agostino, R. B., Sr. and Steyerberg, E. W. (2011). **Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers**. Stat Med *30 (1)*, 11-21, doi: 10.1002/sim.4085.

Peng, L., Xie, R., Holleczek, B., Brenner, H. and Schöttker, B. (2026). **Development of Age- and Sex-Specific Metabolomics-Based Biological Ageing Clocks for 10-Year Mortality Prediction**. Adv Sci (Weinh) *13 (1)*, e10189, doi: 10.1002/advs.202510189.

Pu, Y., Qi, X., Huang, L., Wu, W., Zhou, X., Li, W., Yang, Z., Kong, M., Shen, J., Qi, W., Sun, Z., Mei, Z., Wang, Q., Yang, W., Gao, X., Wang, X., Liu, Z., Yuan, C. and Zheng, Y. (2025). **Gut metagenome and plasma metabolome profiles in older adults suggest pyruvate metabolism as a link between sleep quality and frailty**. Cell Rep *44 (10)*, 116297, doi: 10.1016/j.celrep.2025.116297.

Pu, Y., Sun, Z., Zhang, H., Huang, Q., Wang, Z., Mei, Z., Wang, P., Kong, M., Yang, W., Lin, C., Zhou, X., Lin, S., Huang, Q., Huang, L., Sun, L., Yuan, C., Xu, Q., Tang, H., Wang, X. and Zheng, Y. (2024). **Gut microbial features and circulating metabolomic signatures of frailty in older adults**. Nat Aging *4 (9)*, 1249-1262, doi: 10.1038/s43587-024-00678-0.

Rattray, N. J. W., Trivedi, D. K., Xu, Y., Chandola, T., Johnson, C. H., Marshall, A. D., Mekli, K., Rattray, Z., Tampubolon, G., Vanhoutte, B., White, I. R., Wu, F. C. W., Pendleton, N., Nazroo, J. and Goodacre, R. (2019). **Metabolic dysregulation in vitamin E and carnitine shuttle energy mechanisms associate with human frailty**. Nat Commun *10 (1)*, 5027, doi: 10.1038/s41467-019-12716-2.

Ravindrarajah, R., Lee, D. M., Pye, S. R., Gielen, E., Boonen, S., Vanderschueren, D., Pendleton, N., Finn, J. D., Tajar, A., O'Connell, M. D., Rockwood, K., Bartfai, G., Casanueva, F. F., Forti, G., Giwercman, A., Han, T. S., Huhtaniemi, I. T., Kula, K., Lean, M. E., Punab, M., Wu, F. C. and O'Neill, T. W. (2013). **The ability of three different models of frailty to predict all-cause mortality: results from the European Male Aging Study (EMAS)**. Arch Gerontol Geriatr *57 (3)*, 360-368, doi: 10.1016/j.archger.2013.06.010.

Reicher, L., Bar, N., Godneva, A., Reisner, Y., Zahavi, L., Shahaf, N., Dhir, R., Weinberger, A. and Segal, E. (2024). **Phenome-wide associations of human aging uncover sex-specific dynamics**. Nat Aging *4 (11)*, 1643-1655, doi: 10.1038/s43587-024-00734-9.

Ritchie, S. C., Würtz, P., Nath, A. P., Abraham, G., Havulinna, A. S., Fearnley, L. G., Sarin, A. P., Kangas, A. J., Soininen, P., Aalto, K., Seppälä, I., Raitoharju, E., Salmi, M., Maksimow, M., Männistö, S., Kähönen, M., Juonala, M., Ripatti, S., Lehtimäki, T., Jalkanen, S., Perola, M., Raitakari, O., Salomaa, V., Ala-Korpela, M., Kettunen, J. and Inouye, M. (2015). **The biomarker GlycA is associated with chronic inflammation and predicts long-term risk of severe infection**. Cell Syst *1 (4)*, 293-301, doi: 10.1016/j.cels.2015.09.007.

Ritt, M., Schwarz, C., Kronawitter, V., Delinic, A., Bollheimer, L. C., Gassmann, K. G. and Sieber, C. C. (2015). **Analysis of Rockwood et Al's Clinical Frailty Scale and Fried et Al's Frailty Phenotype as Predictors of Mortality and Other Clinical Outcomes in Older Patients Who Were Admitted to a Geriatric Ward**. J Nutr Health Aging *19 (10)*, 1043-1048, doi: 10.1007/s12603-015-0667-9.

Robinson, O. and Lau, C. E. (2020). **Measuring biological age using metabolomics**. Aging (Albany NY) *12 (22)*, 22352-22353, doi: 10.18632/aging.104216.

Rockwood, K. and Mitnitski, A. (2007). **Frailty in relation to the accumulation of deficits**. J Gerontol A Biol Sci Med Sci *62 (7)*, 722-727, doi: 10.1093/gerona/62.7.722.

Rockwood, K., Song, X., MacKnight, C., Bergman, H., Hogan, D. B., McDowell, I. and Mitnitski, A. (2005). **A global clinical measure of fitness and frailty in elderly people**. CMAJ *173 (5)*, 489-495, doi: 10.1503/cmaj.050051.

Sathyan, S., Liu, F., Tanaka, T., Ferrucci, L., Ayers, E., Windham, B. G., Gao, T., Weiss, E. F., Candia, J., Coresh, J., Barzilai, N., Milman, S., Walker, K. A. and Verghese, J. (2025). **A frailty-based plasma proteomic signature capturing overall health and well-being in older adults**. Aging Cell *24 (9)*, e70144, doi: 10.1111/acel.70144.

Saum, K. U., Dieffenbach, A. K., Jansen, E. H., Schöttker, B., Holleczek, B., Hauer, K. and Brenner, H. (2015). **Association between Oxidative Stress and Frailty in an Elderly German Population: Results from the ESTHER Cohort Study**. Gerontology *61 (5)*, 407-415, doi: 10.1159/000380881.

Saum, K. U., Dieffenbach, A. K., Müller, H., Holleczek, B., Hauer, K. and Brenner, H. (2014). **Frailty prevalence and 10-year survival in community-dwelling older adults: results from the ESTHER cohort study**. Eur J Epidemiol *29 (3)*, 171-179, doi: 10.1007/s10654-014-9891-6.

Saum, K. U., Müller, H., Stegmaier, C., Hauer, K., Raum, E. and Brenner, H. (2012). **Development and evaluation of a modification of the Fried frailty criteria using population-independent cutpoints**. J Am Geriatr Soc *60 (11)*, 2110-2115, doi: 10.1111/j.1532-5415.2012.04192.x.

Sayed, N., Huang, Y., Nguyen, K., Krejciova-Rajaniemi, Z., Grawe, A. P., Gao, T., Tibshirani, R., Hastie, T., Alpert, A., Cui, L., Kuznetsova, T., Rosenberg-Hasson, Y., Ostan, R., Monti, D., Lehallier, B., Shen-Orr, S. S., Maecker, H. T., Dekker, C. L., Wyss-Coray, T., Franceschi, C., Jojic, V., Haddad, F., Montoya, J. G., Wu, J. C., Davis, M. M. and Furman, D. (2021). **An inflammatory aging clock (iAge) based on deep learning tracks multimorbidity, immunosenescence, frailty and cardiovascular aging**. Nat Aging *1*, 598-615, doi: 10.1038/s43587-021-00082-y.

Schöttker, B., Haug, U., Schomburg, L., Köhrle, J., Perna, L., Müller, H., Holleczek, B. and Brenner, H. (2013). **Strong associations of 25-hydroxyvitamin D concentrations with all-cause, cardiovascular, cancer, and respiratory disease mortality in a large cohort study**. Am J Clin Nutr *97 (4)*, 782-793, doi: 10.3945/ajcn.112.047712.

Schöttker, B., Larsen, E. L., Weimann, A., Henriksen, T., Brenner, H. and Poulsen, H. E. (2022). **Associations of urinary metabolites of oxidized DNA and RNA with the incidence of diabetes mellitus using UPLC-MS/MS and ELISA methods**. Free Radic Biol Med *183*, 51-59, doi: 10.1016/j.freeradbiomed.2022.03.007.

Schöttker, B., Saum, K. U., Perna, L., Ordóñez-Mena, J. M., Holleczek, B. and Brenner, H. (2014). **Is vitamin D deficiency a cause of increased morbidity and mortality at older age or simply an indicator of poor health?** Eur J Epidemiol *29 (3)*, 199-210, doi: 10.1007/s10654-014-9894-3.

Searle, S. D., Mitnitski, A., Gahbauer, E. A., Gill, T. M. and Rockwood, K. (2008). **A standard procedure for creating a frailty index**. BMC Geriatr *8*, 24, doi: 10.1186/1471-2318-8-24.

Seeling, K. S., Hehl, L., Vell, M. S., Rendel, M. D., Creasy, K. T., Trautwein, C., Mehler, D. M. A., Keszthelyi, D., Schneider, K. M. and Schneider, C. V. (2023). **Comorbidities, biomarkers and cause specific mortality in patients with irritable bowel syndrome: a phenome-wide association study**. United Eur Gastroent *11 (5)*, 458-470, doi: 10.1002/ueg2.12397.

Sen, P., Shah, P. P., Nativio, R. and Berger, S. L. (2016). **Epigenetic mechanisms of longevity and aging**. Cell *166 (4)*, 822-839, doi: 10.1016/j.cell.2016.07.050.

Shang, N., Liu, H., Wang, N., Guo, S. and Ma, L. (2022). **Comparison of three frailty screening instruments for prediction of adverse outcomes among older adults in the emergency department**. Geriatr Gerontol Int *22 (10)*, 851-856, doi: 10.1111/ggi.14469.

Sharma, M., Anand, A., Chattopadhyay, A. and Goswami, I. (2024). **Gender differentials in cognitive frailty among older adults in India: a multivariate decomposition approach**. Sci Rep *14 (1)*, 24597, doi: 10.1038/s41598-024-74584-1.

Sharma, S., Katz, R., Dubin, R. F., Drew, D. A., Gutierrez, O. M., Shlipak, M. G., Sarnak, M. J. and Ix, J. H. (2021). **FGF23 and cause-specific mortality in community-living individuals: the Health, Aging, and Body Composition Study**. J Am Geriatr Soc *69 (3)*, 711-717, doi: 10.1111/jgs.16910.

Shen, X., Wang, C., Zhou, X., Zhou, W., Hornburg, D., Wu, S. and Snyder, M. P. (2024). **Nonlinear dynamics of multi-omics profiles during human aging**. Nat Aging *4 (11)*, 1619-1634, doi: 10.1038/s43587-024-00692-2.

Sjölander, A. and Dickman, P. W. (2024). **Why test for proportional hazards-or any other model assumptions?** Am J Epidemiol *193 (6)*, 926-927, doi: 10.1093/aje/kwae002.

Soininen, P., Kangas, A. J., Würtz, P., Tukiainen, T., Tynkkynen, T., Laatikainen, R., Järvelin, M. R., Kähönen, M., Lehtimäki, T., Viikari, J., Raitakari, O. T., Savolainen, M. J. and Ala-Korpela, M. (2009). **High-throughput serum NMR metabonomics for cost-effective holistic studies on systemic metabolism**. Analyst *134 (9)*, 1781-1785, doi: 10.1039/b910205a.

Soysal, P., Stubbs, B., Lucato, P., Luchini, C., Solmi, M., Peluso, R., Sergi, G., Isik, A. T., Manzato, E., Maggi, S., Maggio, M., Prina, A. M., Cosco, T. D., Wu, Y. T. and Veronese, N. (2016). **Inflammation and frailty in the elderly: A systematic review and meta-analysis**. Ageing Res Rev *31*, 1-8, doi: 10.1016/j.arr.2016.08.006.

Stekhoven, D. J. and Bühlmann, P. (2012). **MissForest--non-parametric missing value imputation for mixed-type data**. Bioinformatics *28 (1)*, 112-118, doi: 10.1093/bioinformatics/btr597.

Stensrud, M. J. and Hernán, M. A. (2020). **Why test for proportional hazards?** JAMA *323 (14)*, 1401-1402, doi: 10.1001/jama.2020.1267.

Subramanian, I., Verma, S., Kumar, S., Jere, A. and Anamika, K. (2020). **Multi-omics Data Integration, Interpretation, and Its Application**. Bioinform Biol Insights *14*, 1177932219899051, doi: 10.1177/1177932219899051.

Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., Downey, P., Elliott, P., Green, J., Landray, M., Liu, B., Matthews, P., Ong, G., Pell, J., Silman, A., Young, A., Sprosen, T., Peakman, T. and Collins, R. (2015). **UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age**. PLoS Med *12 (3)*, e1001779, doi: 10.1371/journal.pmed.1001779.

Sun, B. B., Chiou, J., Traylor, M., Benner, C., Hsu, Y. H., Richardson, T. G., Surendran, P., Mahajan, A., Robins, C., Vasquez-Grinnell, S. G., Hou, L., Kvikstad, E. M., Burren, O. S., Davitte, J., Ferber, K. L., Gillies, C. E., Hedman Å, K., Hu, S., Lin, T., Mikkilineni, R., Pendergrass, R. K., Pickering, C., Prins, B., Baird, D., Chen, C. Y., Ward, L. D., Deaton, A. M., Welsh, S., Willis, C. M., Lehner, N., Arnold, M., Wörheide, M. A., Suhre, K., Kastenmüller, G., Sethi, A., Cule, M., Raj, A., Burkitt-Gray, L., Melamud, E., Black, M. H., Fauman, E. B., Howson, J. M. M., Kang, H. M., McCarthy, M. I., Nioi, P., Petrovski, S., Scott, R. A., Smith, E. N., Szalma, S., Waterworth, D. M., Mitnaul, L. J., Szustakowski, J. D., Gibson, B. W., Miller, M. R. and Whelan, C. D. (2023). **Plasma proteomic associations with genetics and health in the UK Biobank**. Nature *622 (7982)*, 329-338, doi: 10.1038/s41586-023-06592-6.

Sun, B. B., Maranville, J. C., Peters, J. E., Stacey, D., Staley, J. R., Blackshaw, J., Burgess, S., Jiang, T., Paige, E., Surendran, P., Oliver-Williams, C., Kamat, M. A., Prins, B. P., Wilcox, S. K., Zimmerman, E. S., Chi, A., Bansal, N., Spain, S. L., Wood, A. M., Morrell, N. W., Bradley, J. R., Janjic, N., Roberts, D. J., Ouwehand, W. H., Todd, J. A., Soranzo, N., Suhre, K., Paul, D. S., Fox, C. S., Plenge, R. M., Danesh, J., Runz, H. and Butterworth, A. S. (2018). **Genomic atlas of the human plasma proteome**. Nature *558 (7708)*, 73-79, doi: 10.1038/s41586-018-0175-2.

Takeda, C., Angioni, D., Setphan, E., Macaron, T., De Souto Barreto, P., Sourdet, S., Sierra, F. and Vellas, B. (2020). **Age-related frailty: a clinical model for geroscience?** J Nutr Health Aging *24 (10)*, 1140-1143, doi: 10.1007/s12603-020-1491-4.

Tanaka, T., Basisty, N., Fantoni, G., Candia, J., Moore, A. Z., Biancotto, A., Schilling, B., Bandinelli, S. and Ferrucci, L. (2020). **Plasma proteomic biomarker signature of age predicts health and life span**. Elife *9*, doi: 10.7554/eLife.61073.

Tang, J., Yue, L., Xu, Y., Xu, F., Cai, X., Fu, Y., Miao, Z., Gou, W., Hu, W., Xue, Z., Deng, K., Shen, L., Jiang, Z., Shuai, M., Liang, X., Xiao, C., Xie, Y., Guo, T., Chen, Y. M. and Zheng, J. S. (2025). **Longitudinal serum proteome mapping reveals biomarkers for healthy ageing and related cardiometabolic diseases**. Nat Metab *7 (1)*, 166-181, doi: 10.1038/s42255-024-01185-7.

Theou, O., Brothers, T. D., Mitnitski, A. and Rockwood, K. (2013). **Operationalization of frailty using eight commonly used scales and comparison of their ability to predict all-cause mortality**. J Am Geriatr Soc *61 (9)*, 1537-1551, doi: 10.1111/jgs.12420.

Tibshirani, R. (1997). **The lasso method for variable selection in the Cox model**. Stat Med *16 (4)*, 385-395, doi: 10.1002/(sici)1097-0258(19970228)16:4<385::aid-sim380>3.0.co;2-3.

Tirosh, A., de Souza, R. J., Sacks, F., Bray, G. A., Smith, S. R. and LeBoff, M. S. (2015). **Sex Differences in the Effects of Weight Loss Diets on Bone Mineral Density and Body Composition: POUNDS LOST Trial**. J Clin Endocrinol Metab *100 (6)*, 2463-2471, doi: 10.1210/jc.2015-1050.

Tower, J. (2017). **Sex-Specific Gene Expression and Life Span Regulation**. Trends Endocrinol Metab *28 (10)*, 735-747, doi: 10.1016/j.tem.2017.07.002.

Tower, J., Pomatto, L. C. D. and Davies, K. J. A. (2020). **Sex differences in the response to oxidative and proteolytic stress**. Redox Biol *31*, 101488, doi: 10.1016/j.redox.2020.101488.

Trares, K., Bhardwaj, M., Perna, L., Stocker, H., Petrera, A., Hauck, S. M., Beyreuther, K., Brenner, H. and Schöttker, B. (2022). **Association of the inflammation-related proteome with dementia development at older age: results from a large, prospective, population-based cohort study**. Alzheimers Res Ther *14 (1)*, 128, doi: 10.1186/s13195-022-01063-y.

Trares, K., Wiesenfarth, M., Stocker, H., Perna, L., Petrera, A., Hauck, S. M., Beyreuther, K., Brenner, H. and Schöttker, B. (2024). **Addition of inflammation-related biomarkers to the CAIDE model for risk prediction of all-cause dementia, Alzheimer's disease and vascular dementia in a prospective study**. Immun Ageing *21 (1)*, 23, doi: 10.1186/s12979-024-00427-2.

Tseng, W. H., Chattopadhyay, A., Phan, N. N., Chuang, E. Y. and Lee, O. K. (2024). **Utilizing multimodal approach to identify candidate pathways and biomarkers and predicting frailty syndrome in individuals from UK Biobank**. Geroscience *46 (1)*, 1211-1228, doi: 10.1007/s11357-023-00874-7.

Ukraintseva, S., Arbeev, K., Duan, M., Akushevich, I., Kulminski, A., Stallard, E. and Yashin, A. (2021). **Decline in biological resilience as key manifestation of aging: Potential mechanisms and role in health and longevity**. Mech Ageing Dev *194*, 111418, doi: 10.1016/j.mad.2020.111418.

Uno, H., Cai, T., Pencina, M. J., D'Agostino, R. B. and Wei, L. J. (2011). **On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data**. Stat Med *30 (10)*, 1105-1117, doi: 10.1002/sim.4154.

Van Calster, B., McLernon, D. J., van Smeden, M., Wynants, L. and Steyerberg, E. W. (2019). **Calibration: the Achilles heel of predictive analytics**. BMC Med *17 (1)*, 230, doi: 10.1186/s12916-019-1466-7.

Vickers, A. J. and Elkin, E. B. (2006). **Decision curve analysis: a novel method for evaluating prediction models**. Med Decis Making *26 (6)*, 565-574, doi: 10.1177/0272989x06295361.

Walker, E., Winter, R. and Hodgson, L. E. (2025). **Frailty in older adults admitted to hospital: a six-year dual-centre retrospective study of over 53 000 clinical frailty scale assessments**. Age Ageing *54 (6)*, doi: 10.1093/ageing/afaf137.

Wang, F., Tessier, A. J., Liang, L., Wittenbecher, C., Haslam, D. E., Fernández-Duval, G., Heather Eliassen, A., Rexrode, K. M., Tobias, D. K., Li, J., Zeleznik, O., Grodstein, F., Martínez-González, M. A., Salas-Salvadó, J., Clish, C., Lee, K. H., Sun, Q., Stampfer, M. J., Hu, F. B. and Guasch-Ferré, M. (2023). **Plasma metabolomic profiles associated with mortality and longevity in a prospective analysis of 13,512 individuals**. Nat Commun *14 (1)*, 5744, doi: 10.1038/s41467-023-41515-z.

Wang, S., Nan, B., Rosset, S. and Zhu, J. (2011). **Random LASSO**. Ann Appl Stat *5 (1)*, 468-485, doi: 10.1214/10-aoas377.

Wang, Y., Chen, S., Li, Y., Lu, Q., Li, R., Chen, J. X., Zhang, J. J., Yu, H., Xu, H., Xie, J., Li, X., Liu, G., Pan, A., Geng, T. and Liao, Y. F. (2025). **Metabolic profiling of frailty, associations with type 2 diabetes and interaction with genetic susceptibility**. Cardiovasc Diabetol *24 (1)*, 226, doi: 10.1186/s12933-025-02776-8.

Wang, Y., Xiao, S., Liu, B., Jiang, R., Liu, Y., Hang, Y., Chen, L., Chen, R., Vitiello, M. V., Bennett, D., Wang, B., Lv, J., Yu, C., Haslam, D. E., Zheng, Q., Gerszten, R. E., Bao, Y., Shi, J., Xie, J., Lu, L., Li, L., van Duijn, C. M., Wang, D. D., Chen, Z. and Chan, A. T. (2026). **Organ-specific proteomic aging clocks predict disease and longevity across diverse populations**. Nat Aging *6 (1)*, 162-180, doi: 10.1038/s43587-025-01016-8.

Widagdo, I. S., Pratt, N., Russell, M. and Roughead, E. E. (2015). **Predictive performance of four frailty measures in an older Australian population**. Age Ageing *44 (6)*, 967-972, doi: 10.1093/ageing/afv144.

Wishart, D. S. (2019). **Metabolomics for Investigating Physiological and Pathophysiological Processes**. Physiol Rev *99 (4)*, 1819-1875, doi: 10.1152/physrev.00035.2018.

Wishart, D. S., Feunang, Y. D., Marcu, A., Guo, A. C., Liang, K., Vázquez-Fresno, R., Sajed, T., Johnson, D., Li, C., Karu, N., Sayeeda, Z., Lo, E., Assempour, N., Berjanskii, M., Singhal, S., Arndt, D., Liang, Y., Badran, H., Grant, J., Serra-Cayuela, A., Liu, Y., Mandal, R., Neveu, V., Pon, A., Knox, C., Wilson, M., Manach, C. and Scalbert, A. (2018). **HMDB 4.0: the human metabolome database for 2018**. Nucleic Acids Res *46 (D1)*, d608-d617, doi: 10.1093/nar/gkx1089.

Woo, J., Leung, J. and Morley, J. E. (2012). **Comparison of frailty indicators based on clinical phenotype and the multiple deficit approach in predicting mortality and physical limitation**. J Am Geriatr Soc *60 (8)*, 1478-1486, doi: 10.1111/j.1532-5415.2012.04074.x.

Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., Fu, X., Liu, S., Bo, X. and Yu, G. (2021). **clusterProfiler 4.0: A universal enrichment tool for interpreting omics data**. Innovation (Camb) *2 (3)*, 100141, doi: 10.1016/j.xinn.2021.100141.

Würtz, P., Havulinna, A. S., Soininen, P., Tynkkynen, T., Prieto-Merino, D., Tillin, T., Ghorbani, A., Artati, A., Wang, Q., Tiainen, M., Kangas, A. J., Kettunen, J., Kaikkonen, J., Mikkilä, V., Jula, A., Kähönen, M., Lehtimäki, T., Lawlor, D. A., Gaunt, T. R., Hughes, A. D., Sattar, N., Illig, T., Adamski, J., Wang, T. J., Perola, M., Ripatti, S., Vasan, R. S., Raitakari, O. T., Gerszten, R. E., Casas, J. P., Chaturvedi, N., Ala-Korpela, M. and Salomaa, V. (2015). **Metabolite profiling and cardiovascular event risk: a prospective study of 3 population-based cohorts**. Circulation *131 (9)*, 774-785, doi: 10.1161/circulationaha.114.013116.

Würtz, P., Kangas, A. J., Soininen, P., Lawlor, D. A., Davey Smith, G. and Ala-Korpela, M. (2017). **Quantitative serum nuclear magnetic resonance metabolomics in large-scale epidemiology: a primer on -omic technologies**. Am J Epidemiol *186 (9)*, 1084-1096, doi: 10.1093/aje/kwx016.

Xiong, J., Zhu, X., Guo, Y., Tang, H., Dong, C., Wang, B., Liu, M., Li, Z. and Tu, Y. (2025). **Multi-omic underpinnings of heterogeneous aging across multiple organ systems**. Cell Genom *5 (12)*, 101032, doi: 10.1016/j.xgen.2025.101032.

Xu, J., Liu, J., Tang, J., Liao, J., Liu, X., Odden, M. C. and Wu, C. (2025). **Plasma proteomic signature of risk and prognosis of frailty in the UK Biobank**. Geroscience *47 (2)*, 2365-2381, doi: 10.1007/s11357-024-01415-6.

Xue, Q. L. (2011). **The frailty syndrome: definition and natural history**. Clin Geriatr Med *27 (1)*, 1-15, doi: 10.1016/j.cger.2010.08.009.

Yan, Y., Smith, E., Melander, O. and Ottosson, F. (2022). **The association between plasma metabolites and future risk of all-cause mortality**. J Intern Med *292 (5)*, 804-815, doi: 10.1111/joim.13540.

Yang, Z., Huang, N., Zhuang, Z., Jin, M., Zhang, Z., Song, Y., Cui, H., Zhang, S., Huang, T., Liu, X. and Li, N. (2024). **Earlier age at menopause, plasma metabolome, and risk of premature mortality**. Metabolites *14 (11)*, 571, doi: 10.3390/metabo14110571.

Yao, Z., Jia, X., Chen, Z., Zhang, T., Li, X., Zhang, L., Chen, F., Zhang, J., Zhang, Z., Liu, Z. and Chen, Z. (2024). **Dietary patterns, metabolomics and frailty in a large cohort of 120 000 participants**. Food Funct *15 (6)*, 3174-3185, doi: 10.1039/d3fo03575a.

Zaslavsky, O., Walker, R. L., Crane, P. K., Gray, S. L. and Larson, E. B. (2016). **Glucose levels and risk of frailty**. J Gerontol A Biol Sci Med Sci *71 (9)*, 1223-1229, doi: 10.1093/gerona/glw024.

Zeleznik, O. A., Eliassen, A. H., Kraft, P., Poole, E. M., Rosner, B. A., Jeanfavre, S., Deik, A. A., Bullock, K., Hitchcock, D. S., Avila-Pacheco, J., Clish, C. B. and Tworoger, S. S. (2020). **A prospective analysis of circulating plasma metabolites associated with ovarian cancer risk**. Cancer Res *80 (6)*, 1357-1367, doi: 10.1158/0008-5472.Can-19-2567.

Zeng, X., Chen, R., Shi, D., Zhang, X., Su, T., Wang, Y., Hu, Y., He, M., Yu, H. and Shang, X. (2024). **Association of metabolomic aging acceleration and body mass index phenotypes with mortality and obesity-related morbidities**. Aging Cell, e14435, doi: 10.1111/acel.14435.

Zhang, D., Tang, W., Dou, L. Y., Luo, J. and Sun, Y. (2022). **Four different frailty models predict health outcomes in older patients with stable chronic obstructive pulmonary disease**. BMC Geriatr *22 (1)*, 57, doi: 10.1186/s12877-022-02750-z.

Zhang, S., Wang, Z., Wang, Y., Zhu, Y., Zhou, Q., Jian, X., Zhao, G., Qiu, J., Xia, K., Tang, B., Mutz, J., Li, J. and Li, B. (2024). **A metabolomic profile of biological aging in 250,341 individuals from the UK Biobank**. Nat Commun *15 (1)*, 8081, doi: 10.1038/s41467-024-52310-9.

Zhang, X., Feng, X., Liu, W., Liu, R., Huang, Q., Zhang, P., Li, Z., Li, X., Mao, C. and Duan, C. (2026). **Frailty-related plasma metabolomic signatures predict long-term mortality risk and implicate systemic aging pathways: evidence from a prospective cohort study**. NPJ Aging *12 (1)*, 28, doi: 10.1038/s41514-025-00327-9.

Zhao, B., Gan, L., Graubard, B. I., Männistö, S., Albanes, D. and Huang, J. (2022). **Associations of dietary cholesterol, serum cholesterol, and egg consumption with overall and cause-specific mortality: systematic review and updated meta-analysis**. Circulation *145 (20)*, 1506-1520, doi: 10.1161/circulationaha.121.057642.

Zucchelli, A., Vetrano, D. L., Marengoni, A., Grande, G., Romanelli, G., Calderón-Larrañaga, A., Fratiglioni, L. and Rizzuto, D. (2018). **Frailty predicts short-term survival even in older adults without multimorbidity**. Eur J Intern Med *56*, 53-56, doi: 10.1016/j.ejim.2018.06.012.

# 7 PUBLICATIONS AND PERSONAL CONTRIBUTIONS

The work presented in this dissertation is based on four original research projects conducted as part of my doctoral studies. The corresponding publications and manuscripts are listed below.

**Publications related to this dissertation:**

1. Peng, L., Hauer, K., Holleczek, B., Calderazzo, S., Brenner, H., Schöttker, B. **Comparison of 4 Frailty Measurements in Relation to Ageing, Physical Performance, and All-Cause Mortality**. J Am Med Dir Assoc. 2026 May;27(5):106161. doi: 10.1016/j.jamda.2026.106161.

2. Peng, L., Xie, R., Holleczek, B., Brenner, H., Schöttker, B. **Development of Age- and Sex-Specific Metabolomics-Based Biological Ageing Clocks for 10-Year Mortality Prediction**. Adv Sci (Weinh). 2026 Jan;13(1):e10189. doi: 10.1002/advs.202510189.

**Manuscripts related to this dissertation:**

3. Peng, L., Carlsson, S. V., Hauer, K., Calderazzo, S., Xie, R., Vlaski, T., Holleczek, B., Brenner, H., Schöttker, B. **Integrated Proteomic and Metabolomic Profiling for Frailty Identification and Prediction**. [Manuscript under review at *npj Aging*]

4. Peng, L., Wankhede, D., Bhardwaj, M., Vlaski, T., Xie, R., Calderazzo, S., Holleczek, B., Brenner, H., Schöttker, B. **Novel Inflammaging and Proteomic Ageing Clocks as Surrogates of Biological Age Predict All-Cause Mortality**. [Manuscript in preparation]

Study 1 (Chapters 2.1, 3.1, and 4.1) corresponds to Publication #1, which has been published in the *Journal of the American Medical Directors Association*. Study 2 (Chapters 2.2, 3.2, and 4.2) corresponds to Manuscript #3, which is currently under review at *npj Aging*. Study 3 (Chapters 2.3, 3.3, and 4.3) corresponds to Publication #2, which has been published in *Advanced Science*. Study 4 (Chapters 2.4, 3.4, and 4.4) corresponds to Manuscript #4, which is currently being prepared for journal submission. The General Introduction (Chapter 1) and the integrative sections of the Discussion (Chapters 4.5–4.8) were written specifically for this dissertation to provide the overarching background, synthesis and interpretation of the findings, overall strengths and limitations, implications and future directions, and conclusions across the four research projects.

**Personal contributions:**

*Study 1: Comparison of Four Frailty Measurements in Relation to Ageing, Physical Performance and All-Cause Mortality* corresponds to Publication #1 and forms the basis of Chapters 2.1, 3.1, and 4.1. Prof. Dr. Ben Schöttker conceptualized and designed the study. I performed 100% of the statistical analyses, contributed to the interpretation of the results, prepared the figures and tables, and drafted the manuscript. Prof. Dr. Ben Schöttker supervised the study and made major revisions

to the manuscript. Prof. Dr. Hermann Brenner, Prof. Dr. Klaus Hauer, Prof. Dr. Bernd Holleczek, and Dr. Silvia Calderazzo contributed to the interpretation and discussion of the findings and critically revised the manuscript.

*Study 2: Integrated Proteomic and Metabolomic Profiling for Frailty Identification and Prediction* corresponds to Manuscript #3 and forms the basis of Chapters 2.2, 3.2, and 4.2. I conceptualized and designed the study together with my supervisor, Prof. Dr. Ben Schöttker. I performed 100% of the statistical analyses, contributed to the interpretation of the results, prepared the figures and tables, and drafted the manuscript. Prof. Dr. Ben Schöttker supervised the study and substantially revised the manuscript. Prof. Dr. Hermann Brenner, Prof. Dr. Sigrid Carlsson, Tomislav Vlaski, Prof. Dr. Klaus Hauer, Dr. Silvia Calderazzo, Dr. Ruijie Xie, and Prof. Dr. Bernd Holleczek contributed to the interpretation and discussion of the findings and critically revised the manuscript.

*Study 3: Age- and Sex-Specific Metabolomics-Based Biological Ageing Clocks for 10-Year Mortality Prediction* corresponds to Publication #2 and forms the basis of Chapters 2.3, 3.3, and 4.3. Prof. Dr. Ben Schöttker conceptualized and designed the study. I performed 100% of the statistical analyses, contributed to the interpretation of the results, prepared the figures and tables, and drafted the manuscript. Prof. Dr. Ben Schöttker supervised the study and made major revisions to the manuscript. Prof. Dr. Hermann Brenner, Prof. Dr. Bernd Holleczek, and Dr. Ruijie Xie contributed to the interpretation and discussion of the findings and critically revised the manuscript.

*Study 4: Inflammaging and Proteomic Ageing Clocks as Surrogates of Biological Age Predict All-Cause Mortality* corresponds to Manuscript #4 and forms the basis of Chapters 2.4, 3.4, and 4.4. I conceptualized and designed the study together with my supervisor, Prof. Dr. Ben Schöttker. I performed 100% of the statistical analyses, contributed to the interpretation of the results, prepared the figures and tables, and drafted the manuscript. Prof. Dr. Ben Schöttker supervised the study and substantially revised the manuscript. Prof. Dr. Hermann Brenner, Dr. Durgesh Wankhede, Dr. Megha Bhardwaj, Dr. Tomislav Vlaski, Dr. Ruijie Xie, Dr. Silvia Calderazzo, and Prof. Dr. Bernd Holleczek contributed to the interpretation and discussion of the findings and critically revised the manuscript.

The General Introduction (Chapter 1) and the integrative sections of the Discussion (Chapters 4.5–4.8) were written specifically for this dissertation. I developed the overarching framework integrating the four studies, synthesized and interpreted their findings in the broader context of biological ageing research, and drafted these sections under the supervision of Prof. Dr. Ben Schöttker.

## Additional publications not included in this dissertation:

5. Peng, L, Guo, Y, Gerhard, M, Gao, J. J., Liu, Z. C., Mejías-Luque, R., Zhang, L., Vieth, M., Ma, J. L., Liu, W. D., Li, Z. X., Zhou, T., Li, W. Q., You, W. C., Zhang, Y., Pan, K. F. **Metabolite Alterations and Interactions with Microbiota in Helicobacter pylori-Associated Gastric Lesions**. Microbiol Spectr. 2023 Aug 17;11(4):e0534722. doi: 10.1128/spectrum.05347-22.

6. Liu, Y., Zhang, H., Halimulati., M, Peng, L. **Vascular aging mediates the association between atherogenic lipid profiles and depression: survey-weighted evidence from NHANES 2005-2020**. J Affect Disord. 2026 Feb 15;395(Pt B):120749. doi: 10.1016/j.jad.2025.120749.

7. Xie, H., Halimulati, M., Dou, Y., Zhang, H., Jiang, X., Peng, L. **Systemic immune-inflammation states in US adults with seropositivity to infectious pathogens: A nutrient-wide association study**. JPEN J Parenter Enteral Nutr. 2025 Jan;49(1):94-102. doi: 10.1002/jpen.2695.

8. Wang, S. C., Peng, L. **Frailty as a predictor of mortality in lung cancer survivors: evidence from a nationally representative cohort NHIS 1997-2018**. Sci Rep. 2025 Nov 14;15(1):39898. doi: 10.1038/s41598-025-20785-1.

9. Bai, Y., Zhang, H., Yang, J., Peng, L. **The Association between Daily Total Dietary Nutrient Intake and Recent Glycemic Control States of Non-Pregnant Adults 20+ Years Old from NHANES 1999-2018 (Except for 2003-2004)**. Nutrients. 2021 Nov 21;13(11):4168. doi: 10.3390/nu13114168.

10. Xie, R., Sha, S., Peng, L., Holleczek, B., Brenner, H., Schöttker, B. **Metabolomics data improve 10-year cardiovascular risk prediction with the SCORE2 algorithm for the general population without cardiovascular disease or diabetes**. Eur J Prev Cardiol. 2025 Apr 24:zwaf254. doi: 10.1093/eurjpc/zwaf254.

11. Xie, R., Herder, C., Sha, S., Peng, L., Brenner, H., Schöttker, B. **Novel type 2 diabetes prediction score based on traditional risk factors and circulating metabolites: model derivation and validation in two large cohort studies**. EClinicalMedicine. 2024 Dec 6;79:102971. doi: 10.1016/j.eclinm.2024.102971.

## Awards and honors:

Stephan Weiland Prize 2026 for Best Publication, German Society for Epidemiology (DGEpi), awarded for Publication #2: "Development of Age- and Sex-Specific Metabolomics-Based Biological Ageing Clocks for 10-Year Mortality Prediction."

Academic Innovation Award, Peking University. The highest institutional honor awarded for exceptional postgraduate research and publication excellence (2022).

## Conference presentations:

IARC-DKFZ Conference,
December 3, 2025, Lyon, France
Poster: "Development of Age- and Sex-Specific Metabolomics-Based Biological Ageing Clocks for 10-Year Mortality Prediction"

26th GSAAM Annual Conference | Young Research Award Session,
June 12, 2026, Frankfurt am Main, Germany
Oral presentation: "Integrated Proteomic and Metabolomic Profiling for Frailty Identification and Prediction: Results from the German ESTHER Study and UK Biobank"

Science@DKFZ Lecture Series,
November 12, 2025, Heidelberg, Germany
Oral presentation: "Development of Age- and Sex-Specific Metabolomics-Based Biological Ageing Clocks for 10-Year Mortality Prediction"

# 8 APPENDIX

**Table A1. Quality control metrics for proteins of the OLINK Target 96 Inflammation panel measured in the ESTHER study**

| No. | Protein [a] | UniProt ID [b] | < LOD (%) | Missing (%) | QC [c] |
|---|---|---|---|---|---|
| 1 | *IL22RA1* | Q8N6P7 | 97.6 | 0.1 | Fail |
| 2 | *IL33* | O95760 | 94.1 | 0.0 | Fail |
| 3 | *IL2RB* | P14784 | 93.7 | 0.0 | Fail |
| 4 | *IL24* | Q13007 | 93.5 | 0.2 | Fail |
| 5 | *IL2* | P60568 | 93.3 | 0.0 | Fail |
| 6 | *IL13* | P35225 | 87.1 | 0.0 | Fail |
| 7 | *TSLP* | Q969D9 | 86.9 | 0.0 | Fail |
| 8 | *IL4* | P05112 | 85.6 | 0.0 | Fail |
| 9 | *ARTN* | Q5T4W7 | 73.7 | 0.0 | Fail |
| 10 | *IL1A* | P01583 | 66.7 | 0.0 | Fail |
| 11 | *IL20RA* | Q9UHF4 | 65.8 | 0.0 | Fail |
| 12 | *GDNF* | P39905 | 64.7 | 0.0 | Fail |
| 13 | *TNF* | P01375 | 63.2 | 35.2 | Fail |
| 14 | *IL15RA* | Q13261 | 62.7 | 0.0 | Fail |
| 15 | *IL17C* | Q9P0M4 | 61.6 | 0.0 | Fail |
| 16 | *LIF* | P15018 | 55.7 | 0.0 | Fail |
| 17 | *IL20* | Q9NYY1 | 54.8 | 0.0 | Fail |
| 18 | *NRTN* | Q99748 | 50.9 | 0.0 | Fail |
| 19 | *IL5* | P05113 | 48.5 | 0.0 | Fail |
| 20 | *IL17A* | Q16552 | 41.4 | 0.1 | Fail |
| 21 | *FGF5* | P12034 | 26.3 | 0.0 | Fail |
| 22 | *SLAMF1* | Q13291 | 25.1 | 0.0 | Fail |
| 23 | *IFNG* | P01579 | 23.6 | 35.2 | Fail |
| 24 | *FGF23* | Q9GZV9 | 18.6 | 0.0 | Pass |
| 25 | *SIRT2* | Q8IXJ6 | 16.9 | 0.0 | Pass |
| 26 | *NTF3* | P20783 | 15.8 | 0.0 | Pass |
| 27 | *CCL7* | P80098 | 11.5 | 0.0 | Pass |
| 28 | *IL10RA* | Q13651 | 9.9 | 0.0 | Pass but fail in UKB |
| 29 | *SULT1A1* | P50225 | 7.5 | 0.0 | Pass |
| 30 | *AXIN1* | O15169 | 7.2 | 0.0 | Pass |
| 31 | *CST5* | P28325 | 4.6 | 0.0 | Pass |
| 32 | *IL10* | P22301 | 3.6 | 0.0 | Pass |
| 33 | *CCL28* | Q9NRJ3 | 2.8 | 0.0 | Pass |
| 34 | *IL6* | P05231 | 0.5 | 0.0 | Pass |
| 35 | *CCL20* | P78556 | 0.3 | 0.0 | Pass |
| 36 | *CASP8* | Q14790 | 0.2 | 0.0 | Pass |
| 37 | *CD274* | Q9NZQ7 | 0.2 | 0.0 | Pass |
| 38 | *LIFR* | P42702 | 0.1 | 0.0 | Pass |
| 39 | *IL12B* | P29460 | 0.1 | 0.0 | Pass |
| 40 | *ADA* | P00813 | 0.1 | 0.0 | Pass |
| 41 | *CDCP1* | Q9H5V8 | 0.1 | 0.0 | Pass but fail in UKB |
| 42 | *NGF* | P01138 | 0.1 | 0.0 | Pass |
| 43 | *STAMBP* | O95630 | 0.1 | 0.0 | Pass |
| 44 | *TNFSF11* | O14788 | 0.0 | 0.0 | Pass |
| 45 | *S100A12* | P80511 | 0.0 | 0.0 | Pass |
| 46 | *CXCL9* | Q07325 | 0.0 | 0.0 | Pass |
| 47 | *FGF19* | O95750 | 0.0 | 0.0 | Pass |
| 48 | *LTA* | P01374 | 0.0 | 0.0 | Pass |
| 49 | *CXCL8* | P10145 | 0.0 | 0.0 | Pass |
| 50 | *CD8A* | P01732 | 0.0 | 0.0 | Pass |
| 51 | *CD6* | P30203 | 0.0 | 0.0 | Pass |

| No. | Protein [a] | UniProt ID [b] | < LOD (%) | Missing (%) | QC [c] |
|---|---|---|---|---|---|
| 52 | *CCL13* | Q99616 | 0.0 | 0.0 | Pass |
| 53 | *FGF21* | Q9NSA1 | 0.0 | 0.0 | Pass |
| 54 | *VEGFA* | P15692 | 0.0 | 0.0 | Pass |
| 55 | *CD244* | Q9BZW8 | 0.0 | 0.0 | Pass |
| 56 | *IL7* | P13232 | 0.0 | 0.0 | Pass |
| 57 | *TNFRSF11B* | O00300 | 0.0 | 1.1 | Pass |
| 58 | *TGFB1* | P01137 | 0.0 | 0.0 | Pass |
| 59 | *PLAU* | P00749 | 0.0 | 0.0 | Pass |
| 60 | *CCL2* | P13500 | 0.0 | 0.0 | Pass |
| 61 | *CXCL11* | O14625 | 0.0 | 0.0 | Pass |
| 62 | *TNFSF10* | P50591 | 0.0 | 0.0 | Pass |
| 63 | *OSM* | P13725 | 0.0 | 0.0 | Pass |
| 64 | *CXCL1* | P09341 | 0.0 | 0.0 | Pass |
| 65 | *CCL4* | P13236 | 0.0 | 0.0 | Pass |
| 66 | *KITLG* | P21583 | 0.0 | 0.0 | Pass |
| 67 | *IL18* | Q14116 | 0.0 | 0.0 | Pass |
| 68 | *TGFA* | P01135 | 0.0 | 0.0 | Pass |
| 69 | *CCL11* | P51671 | 0.0 | 0.0 | Pass |
| 70 | *TNFSF14* | O43557 | 0.0 | 0.0 | Pass |
| 71 | *MMP1* | P03956 | 0.0 | 0.0 | Pass |
| 72 | *CCL19* | Q99731 | 0.0 | 0.0 | Pass |
| 73 | *IL10RB* | Q08334 | 0.0 | 0.0 | Pass |
| 74 | *IL18R1* | Q13478 | 0.0 | 0.0 | Pass |
| 75 | *CXCL5* | P42830 | 0.0 | 0.0 | Pass |
| 76 | *HGF* | P14210 | 0.0 | 0.0 | Pass |
| 77 | *MMP10* | P09238 | 0.0 | 0.0 | Pass |
| 78 | *CCL23* | P55773 | 0.0 | 0.0 | Pass |
| 79 | *CD5* | P06127 | 0.0 | 0.0 | Pass |
| 80 | *CCL3* | P10147 | 0.0 | 0.0 | Pass |
| 81 | *FLT3LG* | P49771 | 0.0 | 0.0 | Pass |
| 82 | *CXCL6* | P80162 | 0.0 | 0.0 | Pass |
| 83 | *CXCL10* | P02778 | 0.0 | 0.0 | Pass |
| 84 | *EIF4EBP1* | Q13541 | 0.0 | 0.0 | Pass |
| 85 | *DNER* | Q8NFT8 | 0.0 | 0.0 | Pass |
| 86 | *CD40* | P25942 | 0.0 | 0.0 | Pass |
| 87 | *CCL8* | P80075 | 0.0 | 0.0 | Pass |
| 88 | *CCL25* | O15444 | 0.0 | 0.0 | Pass |
| 89 | *CX3CL1* | P78423 | 0.0 | 0.0 | Pass |
| 90 | *TNFRSF9* | Q07011 | 0.0 | 0.0 | Pass |
| 91 | *TNFSF12* | O43508 | 0.0 | 1.0 | Pass |
| 92 | *CSF1* | P09603 | 0.0 | 0.0 | Pass |

[a] Proteins were measured using the Olink® Target 96 Inflammation panel in the ESTHER cohort (n=8,082). Values represent the percentage of samples below the assay-specific limit of detection (LOD) and the proportion of missing measurements.

[b] UniProt identifiers are provided for protein annotation (https://www.uniprot.org/).

[c] QC status was defined as "Pass" for proteins with <25% of values below LOD and <20% missingness; all others were classified as "Fail". All 69 proteins that passed QC in the ESTHER cohort were available in the UK Biobank Olink Explore 3072 platform. However, 2 proteins (*IL10RA* and *CDCP1*) did not meet QC criteria in the UK Biobank (n=34,852) and needed to be excluded from analysis.

**Table A2. Abbreviations, full names, classes and distributions of 249 metabolomic biomarkers in the UK Biobank and ESTHER study**

| Metabolic biomarkers | Full Name | Class | UK Biobank n (%) | UK Biobank mean (SD) | ESTHER n (%) | ESTHER mean (SD) | Unit |
|---|---|---|---|---|---|---|---|
| Acetate | Acetate | Ketone bodies | 208,974 (99.9) | 0.02 (0.03) | 6,820 (100) | 0.04 (0.06) | mmol/l |
| Acetoacetate | Acetoacetate | Ketone bodies | 209,136 (99.9) | 0.01 (0.01) | 6,818 (99.9) | 0.01 (0.01) | mmol/l |
| Acetone | Acetone | Ketone bodies | 209,141 (99.9) | 0.01 (0.01) | 6,820 (100) | 0.02 (0.01) | mmol/l |
| Ala | Alanine | Amino acids | 209,061 (99.9) | 0.30 (0.08) | 6,775 (99.3) | 0.31 (0.08) | mmol/l |
| Albumin | Albumin | Fluid balance | 209,116 (99.9) | 39.24 (3.37) | 6,820 (100) | 44.70 (4.49) | g/l |
| ApoA1 | Apolipoprotein A1 | Apolipoproteins | 209,144 (100) | 1.47 (0.25) | 6,820 (100) | 1.52 (0.28) | g/l |
| ApoB | Apolipoprotein B | Apolipoproteins | 209,144 (100) | 0.86 (0.20) | 6,820 (100) | 1.10 (0.26) | g/l |
| ApoB/ApoA1 | Ratio of apolipoprotein B to apolipoprotein A1 | Apolipoproteins | 209,144 (100) | 0.60 (0.16) | 6,820 (100) | 0.75 (0.24) | ratio |
| bOHbutyrate | 3-Hydroxybutyrate | Ketone bodies | 205,167 (98.1) | 0.06 (0.06) | 6,769 (99.3) | 0.09 (0.07) | mmol/l |
| Cholines | Total cholines | Other lipids | 208,978 (99.9) | 2.60 (0.42) | 6,794 (99.6) | 2.94 (0.46) | mmol/l |
| Citrate | Citrate | Glycolysis related metabolites | 209,126 (99.9) | 0.07 (0.01) | 6,723 (98.6) | 0.06 (0.02) | mmol/l |
| Clinical-LDL-C | Clinical LDL cholesterol | Cholesterol | 209,144 (100) | 2.58 (0.75) | 6,820 (100) | 3.47 (0.99) | mmol/l |
| Creatinine | Creatinine | Fluid balance | 203,993 (97.5) | 0.07 (0.02) | 6,535 (96.1) | 0.06 (0.02) | mmol/l |
| DHA | Docosahexaenoic acid | Fatty acids | 208,978 (99.9) | 0.24 (0.09) | 6,794 (99.6) | 0.27 (0.08) | mmol/l |
| DHA-pct | Ratio of docosahexaenoic acid to total fatty acids | Fatty acid ratios | 208,978 (99.9) | 2.02 (0.69) | 6,794 (99.6) | 1.94 (0.51) | ratio |
| Gln | Glutamine | Amino acids | 208,489 (99.7) | 0.56 (0.08) | 6,627 (97.2) | 0.54 (0.11) | mmol/l |
| Glucose | Glucose | Glycolysis related metabolites | 208,766 (99.8) | 3.76 (1.22) | 6,584 (96.5) | 3.79 (2.17) | mmol/l |
| Gly | Glycine | Amino acids | 208,835 (99.8) | 0.17 (0.07) | 6,780 (99.4) | 0.36 (0.11) | mmol/l |
| GlycA | Glycoprotein acetyls | Inflammation | 209,143 (99.9) | 0.82 (0.12) | 6,820 (100) | 0.87 (0.13) | mmol/l |
| HDL-C | HDL cholesterol | Cholesterol | 209,144 (100) | 1.33 (0.33) | 6,820 (100) | 1.39 (0.34) | mmol/l |
| HDL-CE | Cholesteryl esters in HDL | Cholesteryl esters | 209,144 (100) | 1.04 (0.26) | 6,820 (100) | 1.07 (0.27) | mmol/l |
| HDL-FC | Free cholesterol in HDL | Free cholesterol | 209,144 (100) | 0.30 (0.07) | 6,820 (100) | 0.33 (0.07) | mmol/l |
| HDL-L | Total lipids in HDL | Total lipids | 209,144 (100) | 3.05 (0.65) | 6,820 (100) | 3.17 (0.69) | mmol/l |
| HDL-P | Concentration of HDL particles | Lipoprotein particle concentrations | 209,144 (100) | 0.02 (0.00) | 6,820 (100) | 0.02 (0.00) | mmol/l |
| HDL-PL | Phospholipids in HDL | Phospholipids | 209,144 (100) | 1.57 (0.32) | 6,820 (100) | 1.60 (0.35) | mmol/l |
| HDL-size | Average diameter for HDL particles | Lipoprotein particle sizes | 209,144 (100) | 9.64 (0.20) | 6,820 (100) | 9.67 (0.17) | nm |
| HDL-TG | Triglycerides in HDL | Triglycerides | 209,144 (100) | 0.15 (0.05) | 6,820 (100) | 0.17 (0.07) | mmol/l |
| His | Histidine | Amino acids | 208,869 (99.8) | 0.07 (0.01) | 6,785 (99.5) | 0.09 (0.02) | mmol/l |
| IDL-C | Cholesterol in IDL | IDL (average diameter 28.6 nm) | 209,144 (100) | 0.85 (0.22) | 6,820 (100) | 1.11 (0.30) | mmol/l |
| IDL-C-pct | Cholesterol to total lipids ratio in IDL | IDL ratios | 209,144 (100) | 67.58 (3.06) | 6,820 (100) | 67.57 (3.10) | % |
| IDL-CE | Cholesteryl esters in IDL | IDL (average diameter 28.6 nm) | 209,144 (100) | 0.62 (0.17) | 6,820 (100) | 0.82 (0.22) | mmol/l |
| IDL-CE-pct | Cholesteryl esters to total lipids ratio in IDL | IDL ratios | 209,144 (100) | 49.87 (2.49) | 6,820 (100) | 50.18 (2.66) | % |
| IDL-FC | Free cholesterol in IDL | IDL (average diameter 28.6 nm) | 209,144 (100) | 0.22 (0.06) | 6,820 (100) | 0.29 (0.08) | mmol/l |
| IDL-FC-pct | Free cholesterol to total lipids ratio in IDL | IDL ratios | 209,144 (100) | 17.71 (1.13) | 6,820 (100) | 17.40 (0.92) | % |
| IDL-L | Total lipids in IDL | IDL (average diameter 28.6 nm) | 209,144 (100) | 1.25 (0.30) | 6,820 (100) | 1.64 (0.41) | mmol/l |
| IDL-P | Concentration of IDL particles | IDL (average diameter 28.6 nm) | 209,144 (100) | 0.01 (0.01) | 6,820 (100) | 0.01 (0.01) | mmol/l |

| Metabolic biomarkers | Full Name | Class | UK Biobank n (%) | UK Biobank mean (SD) | ESTHER n (%) | ESTHER mean (SD) | Unit |
|---|---|---|---|---|---|---|---|
| IDL-PL | Phospholipids in IDL | IDL (average diameter 28.6 nm) | 209,144 (100) | 0.30 (0.07) | 6,820 (100) | 0.39 (0.10) | mmol/l |
| IDL-PL-pct | Phospholipids to total lipids ratio in IDL | IDL ratios | 209,144 (100) | 23.84 (0.92) | 6,820 (100) | 23.79 (0.78) | % |
| IDL-TG | Triglycerides in IDL | IDL (average diameter 28.6 nm) | 209,144 (100) | 0.10 (0.03) | 6,820 (100) | 0.14 (0.05) | mmol/l |
| IDL-TG-pct | Triglycerides to total lipids ratio in IDL | IDL ratios | 209,144 (100) | 8.59 (2.55) | 6,820 (100) | 8.64 (2.79) | % |
| Ile | Isoleucine | Branched−chain amino acids | 209,111 (99.9) | 0.05 (0.02) | 6,820 (100) | 0.07 (0.02) | mmol/l |
| L-HDL-C | Cholesterol in large HDL | Large HDL (average diameter 12.1 nm) | 209,144 (100) | 0.30 (0.17) | 6,820 (100) | 0.32 (0.15) | mmol/l |
| L-HDL-C-pct | Cholesterol to total lipids ratio in large HDL | Large HDL ratios | 209,144 (100) | 43.64 (5.64) | 6,820 (100) | 44.61 (4.81) | % |
| L-HDL-CE | Cholesteryl esters in large HDL | Large HDL (average diameter 12.1 nm) | 209,144 (100) | 0.23 (0.13) | 6,820 (100) | 0.25 (0.12) | mmol/l |
| L-HDL-CE-pct | Cholesteryl esters to total lipids ratio in large HDL | Large HDL ratios | 209,144 (100) | 33.68 (4.84) | 6,820 (100) | 34.02 (4.46) | % |
| L-HDL-FC | Free cholesterol in large HDL | Large HDL (average diameter 12.1 nm) | 209,144 (100) | 0.07 (0.04) | 6,820 (100) | 0.08 (0.03) | mmol/l |
| L-HDL-FC-pct | Free cholesterol to total lipids ratio in large HDL | Large HDL ratios | 209,144 (100) | 9.97 (1.21) | 6,820 (100) | 10.58 (1.02) | % |
| L-HDL-L | Total lipids in large HDL | Large HDL (average diameter 12.1 nm) | 209,144 (100) | 0.66 (0.33) | 6,820 (100) | 0.71 (0.30) | mmol/l |
| L-HDL-P | Concentration of large HDL particles | Large HDL (average diameter 12.1 nm) | 209,144 (100) | 0.01 (0.01) | 6,820 (100) | 0.01 (0.01) | mmol/l |
| L-HDL-PL | Phospholipids in large HDL | Large HDL (average diameter 12.1 nm) | 209,144 (100) | 0.33 (0.15) | 6,820 (100) | 0.35 (0.14) | mmol/l |
| L-HDL-PL-pct | Phospholipids to total lipids ratio in large HDL | Large HDL ratios | 209,144 (100) | 50.70 (3.12) | 6,820 (100) | 49.03 (2.57) | % |
| L-HDL-TG | Triglycerides in large HDL | Large HDL (average diameter 12.1 nm) | 209,144 (100) | 0.03 (0.01) | 6,820 (100) | 0.04 (0.02) | mmol/l |
| L-HDL-TG-pct | Triglycerides to total lipids ratio in large HDL | Large HDL ratios | 209,144 (100) | 5.66 (3.27) | 6,820 (100) | 6.35 (3.57) | % |
| L-LDL-C | Cholesterol in large LDL | Large LDL (average diameter 25.5 nm) | 209,144 (100) | 1.15 (0.29) | 6,820 (100) | 1.51 (0.38) | mmol/l |
| L-LDL-C-pct | Cholesterol to total lipids ratio in large LDL | Large LDL ratios | 209,144 (100) | 71.09 (1.93) | 6,820 (100) | 71.25 (2.39) | % |
| L-LDL-CE | Cholesteryl esters in large LDL | Large LDL (average diameter 25.5 nm) | 209,144 (100) | 0.85 (0.22) | 6,820 (100) | 1.11 (0.27) | mmol/l |
| L-LDL-CE-pct | Cholesteryl esters to total lipids ratio in large LDL | Large LDL ratios | 209,144 (100) | 52.64 (1.35) | 6,820 (100) | 52.55 (1.58) | % |
| L-LDL-FC | Free cholesterol in large LDL | Large LDL (average diameter 25.5 nm) | 209,144 (100) | 0.30 (0.08) | 6,820 (100) | 0.40 (0.11) | mmol/l |
| L-LDL-FC-pct | Free cholesterol to total lipids ratio in large LDL | Large LDL ratios | 209,144 (100) | 18.45 (1.35) | 6,820 (100) | 18.71 (1.28) | % |
| L-LDL-L | Total lipids in large LDL | Large LDL (average diameter 25.5 nm) | 209,144 (100) | 1.61 (0.39) | 6,820 (100) | 2.12 (0.50) | mmol/l |
| L-LDL-P | Concentration of large LDL particles | Large LDL (average diameter 25.5 nm) | 209,144 (100) | 0.01 (0.01) | 6,820 (100) | 0.01 (0.01) | mmol/l |
| L-LDL-PL | Phospholipids in large LDL | Large LDL (average diameter 25.5 nm) | 209,144 (100) | 0.36 (0.09) | 6,820 (100) | 0.47 (0.11) | mmol/l |
| L-LDL-PL-pct | Phospholipids to total lipids ratio in large LDL | Large LDL ratios | 209,144 (100) | 22.47 (0.80) | 6,820 (100) | 22.03 (0.62) | % |

| Metabolic biomarkers | Full Name | Class | UK Biobank n (%) | UK Biobank mean (SD) | ESTHER n (%) | ESTHER mean (SD) | Unit |
|---|---|---|---|---|---|---|---|
| L-LDL-TG | Triglycerides in large LDL | Large LDL (average diameter 25.5 nm) | 209,144 (100) | 0.10 (0.03) | 6,820 (100) | 0.14 (0.05) | mmol/l |
| L-LDL-TG-pct | Triglycerides to total lipids ratio in large LDL | Large LDL ratios | 209,144 (100) | 6.44 (1.86) | 6,820 (100) | 6.71 (2.39) | % |
| L-VLDL-C | Cholesterol in large VLDL | Large VLDL (average diameter 53.6 nm) | 209,144 (100) | 0.10 (0.05) | 6,820 (100) | 0.11 (0.05) | mmol/l |
| L-VLDL-C-pct | Cholesterol to total lipids ratio in large VLDL | Large VLDL ratios | 209,123 (99.9) | 30.20 (4.39) | 6,820 (100) | 33.54 (5.73) | % |
| L-VLDL-CE | Cholesteryl esters in large VLDL | Large VLDL (average diameter 53.6 nm) | 209,144 (100) | 0.05 (0.02) | 6,820 (100) | 0.06 (0.03) | mmol/l |
| L-VLDL-CE-pct | Cholesteryl esters to total lipids ratio in large VLDL | Large VLDL ratios | 209,123 (99.9) | 16.25 (3.50) | 6,820 (100) | 18.38 (4.26) | % |
| L-VLDL-FC | Free cholesterol in large VLDL | Large VLDL (average diameter 53.6 nm) | 209,144 (100) | 0.05 (0.03) | 6,820 (100) | 0.05 (0.03) | mmol/l |
| L-VLDL-FC-pct | Free cholesterol to total lipids ratio in large VLDL | Large VLDL ratios | 209,123 (99.9) | 13.95 (1.38) | 6,820 (100) | 14.93 (2.52) | % |
| L-VLDL-L | Total lipids in large VLDL | Large VLDL (average diameter 53.6 nm) | 209,144 (100) | 0.34 (0.18) | 6,820 (100) | 0.35 (0.19) | mmol/l |
| L-VLDL-P | Concentration of large VLDL particles | Large VLDL (average diameter 53.6 nm) | 209,144 (100) | 0.01 (0.01) | 6,820 (100) | 0.01 (0.01) | mmol/l |
| L-VLDL-PL | Phospholipids in large VLDL | Large VLDL (average diameter 53.6 nm) | 209,144 (100) | 0.07 (0.04) | 6,820 (100) | 0.08 (0.04) | mmol/l |
| L-VLDL-PL-pct | Phospholipids to total lipids ratio in large VLDL | Large VLDL ratios | 209,123 (99.9) | 19.13 (3.36) | 6,820 (100) | 20.86 (3.79) | % |
| L-VLDL-TG | Triglycerides in large VLDL | Large VLDL (average diameter 53.6 nm) | 209,144 (100) | 0.17 (0.10) | 6,820 (100) | 0.16 (0.10) | mmol/l |
| L-VLDL-TG-pct | Triglycerides to total lipids ratio in large VLDL | Large VLDL ratios | 209,123 (99.9) | 50.67 (5.92) | 6,820 (100) | 45.37 (7.82) | % |
| LA | Linoleic acid | Fatty acids | 208,978 (99.9) | 3.48 (0.70) | 6,794 (99.6) | 4.17 (0.80) | mmol/l |
| LA-pct | Ratio of linoleic acid to total fatty acids | Fatty acid ratios | 208,978 (99.9) | 28.65 (3.44) | 6,794 (99.6) | 29.84 (3.66) | % |
| Lactate | Lactate | Glycolysis related metabolites | 208,761 (99.8) | 3.94 (1.11) | 6,817 (99.9) | 4.45 (2.71) | mmol/l |
| LDL-C | LDL cholesterol | Cholesterol | 209,144 (100) | 1.76 (0.45) | 6,820 (100) | 2.30 (0.55) | mmol/l |
| LDL-CE | Cholesteryl esters in LDL | Cholesteryl esters | 209,144 (100) | 1.29 (0.33) | 6,820 (100) | 1.68 (0.40) | mmol/l |
| LDL-FC | Free cholesterol in LDL | Free cholesterol | 209,144 (100) | 0.47 (0.12) | 6,820 (100) | 0.62 (0.16) | mmol/l |
| LDL-L | Total lipids in LDL | Total lipids | 209,144 (100) | 2.53 (0.62) | 6,820 (100) | 3.30 (0.75) | mmol/l |
| LDL-P | Concentration of LDL particles | Lipoprotein particle concentra-tions | 209,144 (100) | 0.01 (0.01) | 6,820 (100) | 0.01 (0.01) | mmol/l |
| LDL-PL | Phospholipids in LDL | Phospholipids | 209,144 (100) | 0.61 (0.14) | 6,820 (100) | 0.79 (0.18) | mmol/l |
| LDL-size | Average diameter for LDL particles | Lipoprotein particle sizes | 209,144 (100) | 23.92 (0.09) | 6,820 (100) | 23.95 (0.11) | nm |
| LDL-TG | Triglycerides in LDL | Triglycerides | 209,144 (100) | 0.15 (0.04) | 6,820 (100) | 0.20 (0.07) | mmol/l |
| Leu | Leucine | Branched−chain amino acids | 209,122 (99.9) | 0.10 (0.03) | 6,820 (100) | 0.16 (0.04) | mmol/l |
| M-HDL-C | Cholesterol in medium HDL | Medium HDL (average diameter 10.9 nm) | 209,144 (100) | 0.50 (0.13) | 6,820 (100) | 0.51 (0.15) | mmol/l |
| M-HDL-C-pct | Cholesterol to total lipids ratio in medium HDL | Medium HDL ratios | 209,144 (100) | 47.48 (2.99) | 6,820 (100) | 47.21 (3.71) | % |

| Metabolic biomarkers | Full Name | Class | UK Biobank n (%) | UK Biobank mean (SD) | ESTHER n (%) | ESTHER mean (SD) | Unit |
|---|---|---|---|---|---|---|---|
| M-HDL-CE | Cholesteryl esters in medium HDL | Medium HDL (average diameter 10.9 nm) | 209,144 (100) | 0.41 (0.10) | 6,820 (100) | 0.41 (0.12) | mmol/l |
| M-HDL-CE-pct | Cholesteryl esters to total lipids ratio in medium HDL | Medium HDL ratios | 209,144 (100) | 39.14 (2.59) | 6,820 (100) | 38.38 (3.53) | % |
| M-HDL-FC | Free cholesterol in medium HDL | Medium HDL (average diameter 10.9 nm) | 209,144 (100) | 0.09 (0.03) | 6,820 (100) | 0.09 (0.03) | mmol/l |
| M-HDL-FC-pct | Free cholesterol to total lipids ratio in medium HDL | Medium HDL ratios | 209,144 (100) | 8.34 (0.70) | 6,820 (100) | 8.83 (0.63) | % |
| M-HDL-L | Total lipids in medium HDL | Medium HDL (average diameter 10.9 nm) | 209,144 (100) | 1.05 (0.23) | 6,820 (100) | 1.06 (0.27) | mmol/l |
| M-HDL-P | Concentration of medium HDL particles | Medium HDL (average diameter 10.9 nm) | 209,144 (100) | 0.01 (0.01) | 6,820 (100) | 0.01 (0.01) | mmol/l |
| M-HDL-PL | Phospholipids in medium HDL | Medium HDL (average diameter 10.9 nm) | 209,144 (100) | 0.50 (0.10) | 6,820 (100) | 0.50 (0.12) | mmol/l |
| M-HDL-PL-pct | Phospholipids to total lipids ratio in medium HDL | Medium HDL ratios | 209,144 (100) | 47.12 (1.28) | 6,820 (100) | 46.69 (1.43) | % |
| M-HDL-TG | Triglycerides in medium HDL | Medium HDL (average diameter 10.9 nm) | 209,144 (100) | 0.06 (0.02) | 6,820 (100) | 0.06 (0.03) | mmol/l |
| M-HDL-TG-pct | Triglycerides to total lipids ratio in medium HDL | Medium HDL ratios | 209,144 (100) | 5.40 (1.83) | 6,820 (100) | 6.11 (2.66) | % |
| M-LDL-C | Cholesterol in medium LDL | Medium LDL (average diameter 23 nm) | 209,144 (100) | 0.43 (0.12) | 6,820 (100) | 0.56 (0.14) | mmol/l |
| M-LDL-C-pct | Cholesterol to total lipids ratio in medium LDL | Medium LDL ratios | 209,144 (100) | 68.34 (1.74) | 6,820 (100) | 68.04 (2.01) | % |
| M-LDL-CE | Cholesteryl esters in medium LDL | Medium LDL (average diameter 23 nm) | 209,144 (100) | 0.31 (0.09) | 6,820 (100) | 0.40 (0.10) | mmol/l |
| M-LDL-CE-pct | Cholesteryl esters to total lipids ratio in medium LDL | Medium LDL ratios | 209,144 (100) | 49.15 (1.94) | 6,820 (100) | 48.63 (1.95) | % |
| M-LDL-FC | Free cholesterol in medium LDL | Medium LDL (average diameter 23 nm) | 209,144 (100) | 0.12 (0.03) | 6,820 (100) | 0.16 (0.04) | mmol/l |
| M-LDL-FC-pct | Free cholesterol to total lipids ratio in medium LDL | Medium LDL ratios | 209,144 (100) | 19.19 (1.94) | 6,820 (100) | 19.40 (2.12) | % |
| M-LDL-L | Total lipids in medium LDL | Medium LDL (average diameter 23 nm) | 209,144 (100) | 0.63 (0.17) | 6,820 (100) | 0.81 (0.20) | mmol/l |
| M-LDL-P | Concentration of medium LDL particles | Medium LDL (average diameter 23 nm) | 209,144 (100) | 0.01 (0.01) | 6,820 (100) | 0.01 (0.01) | mmol/l |
| M-LDL-PL | Phospholipids in medium LDL | Medium LDL (average diameter 23 nm) | 209,144 (100) | 0.16 (0.04) | 6,820 (100) | 0.21 (0.05) | mmol/l |
| M-LDL-PL-pct | Phospholipids to total lipids ratio in medium LDL | Medium LDL ratios | 209,144 (100) | 26.13 (0.82) | 6,820 (100) | 26.25 (0.93) | % |
| M-LDL-TG | Triglycerides in medium LDL | Medium LDL (average diameter 23 nm) | 209,144 (100) | 0.03 (0.01) | 6,820 (100) | 0.05 (0.02) | mmol/l |
| M-LDL-TG-pct | Triglycerides to total lipids ratio in medium LDL | Medium LDL ratios | 209,144 (100) | 5.53 (1.61) | 6,820 (100) | 5.71 (2.16) | % |

| Metabolic biomarkers | Full Name | Class | UK Biobank n (%) | UK Biobank mean (SD) | ESTHER n (%) | ESTHER mean (SD) | Unit |
|---|---|---|---|---|---|---|---|
| M-VLDL-C | Cholesterol in medium VLDL | Medium VLDL (average diameter 44.5 nm) | 209,144 (100) | 0.17 (0.07) | 6,820 (100) | 0.23 (0.08) | mmol/l |
| M-VLDL-C-pct | Cholesterol to total lipids ratio in medium VLDL | Medium VLDL ratios | 209,144 (100) | 29.75 (6.95) | 6,820 (100) | 33.61 (6.46) | % |
| M-VLDL-CE | Cholesteryl esters in medium VLDL | Medium VLDL (average diameter 44.5 nm) | 209,144 (100) | 0.09 (0.04) | 6,820 (100) | 0.13 (0.05) | mmol/l |
| M-VLDL-CE-pct | Cholesteryl esters to total lipids ratio in medium VLDL | Medium VLDL ratios | 209,144 (100) | 16.08 (5.20) | 6,820 (100) | 18.83 (4.87) | % |
| M-VLDL-FC | Free cholesterol in medium VLDL | Medium VLDL (average diameter 44.5 nm) | 209,144 (100) | 0.08 (0.03) | 6,820 (100) | 0.10 (0.03) | mmol/l |
| M-VLDL-FC-pct | Free cholesterol to total lipids ratio in medium VLDL | Medium VLDL ratios | 209,144 (100) | 13.67 (1.80) | 6,820 (100) | 14.78 (1.68) | % |
| M-VLDL-L | Total lipids in medium VLDL | Medium VLDL (average diameter 44.5 nm) | 209,144 (100) | 0.59 (0.21) | 6,820 (100) | 0.70 (0.23) | mmol/l |
| M-VLDL-P | Concentration of medium VLDL particles | Medium VLDL (average diameter 44.5 nm) | 209,144 (100) | 0.01 (0.01) | 6,820 (100) | 0.01 (0.01) | mmol/l |
| M-VLDL-PL | Phospholipids in medium VLDL | Medium VLDL (average diameter 44.5 nm) | 209,144 (100) | 0.13 (0.05) | 6,820 (100) | 0.17 (0.05) | mmol/l |
| M-VLDL-PL-pct | Phospholipids to total lipids ratio in medium VLDL | Medium VLDL ratios | 209,144 (100) | 22.21 (1.88) | 6,820 (100) | 23.95 (1.84) | % |
| M-VLDL-TG | Triglycerides in medium VLDL | Medium VLDL (average diameter 44.5 nm) | 209,144 (100) | 0.29 (0.12) | 6,820 (100) | 0.30 (0.13) | mmol/l |
| M-VLDL-TG-pct | Triglycerides to total lipids ratio in medium VLDL | Medium VLDL ratios | 209,144 (100) | 48.04 (8.71) | 6,820 (100) | 42.44 (8.14) | % |
| MUFA | Monounsaturated fatty acids | Fatty acids | 208,978 (99.9) | 2.95 (0.83) | 6,794 (99.6) | 3.36 (0.93) | mmol/l |
| MUFA-pct | Ratio of monounsaturated fatty acids to total fatty acids | Fatty acid ratios | 208,978 (99.9) | 23.82 (2.65) | 6,794 (99.6) | 23.59 (2.40) | % |
| non-HDL-C | Total cholesterol minus HDL-C | Cholesterol | 209,144 (100) | 3.34 (0.86) | 6,820 (100) | 4.33 (1.05) | mmol/l |
| Omega-3 | Omega-3 fatty acids | Fatty acids | 208,978 (99.9) | 0.55 (0.23) | 6,794 (99.6) | 0.63 (0.21) | mmol/l |
| Omega-3-pct | Ratio of omega-3 fatty acids to total fatty acids | Fatty acid ratios | 208,978 (99.9) | 4.51 (1.58) | 6,794 (99.6) | 4.46 (1.12) | % |
| Omega-6 | Omega-6 fatty acids | Fatty acids | 208,978 (99.9) | 4.54 (0.70) | 6,794 (99.6) | 5.18 (0.79) | mmol/l |
| Omega-6/Omega-3 | Ratio of omega-6 fatty acids to omega-3 fatty acids | Fatty acid ratios | 208,969 (99.9) | 9.45 (4.21) | 6,794 (99.6) | 8.90 (2.59) | ratio |
| Omega-6-pct | Ratio of omega-6 fatty acids to total fatty acids | Fatty acid ratios | 208,978 (99.9) | 37.58 (3.57) | 6,794 (99.6) | 37.21 (3.55) | % |
| Phe | Phenylalanine | Aromatic amino acids | 209,046 (99.9) | 0.05 (0.01) | 6,786 (99.5) | 0.10 (0.02) | mmol/l |
| Phosphatidylc | Phosphatidylcholines | Other lipids | 208,978 (99.9) | 2.13 (0.39) | 6,794 (99.6) | 2.38 (0.43) | mmol/l |
| Phosphoglyc | Phosphoglycerides | Other lipids | 208,978 (99.9) | 2.32 (0.41) | 6,794 (99.6) | 2.61 (0.45) | mmol/l |
| PUFA | Polyunsaturated fatty acids | Fatty acids | 208,978 (99.9) | 5.10 (0.82) | 6,794 (99.6) | 5.81 (0.92) | mmol/l |
| PUFA/MUFA | Ratio of polyunsaturated fatty acids to monoun-saturated fatty acids | Fatty acid ratios | 208,978 (99.9) | 1.80 (0.34) | 6,794 (99.6) | 1.80 (0.31) | ratio |
| PUFA-pct | Ratio of polyunsaturated fatty acids to total fatty acids | Fatty acid ratios | 208,978 (99.9) | 42.10 (3.75) | 6,794 (99.6) | 41.67 (3.57) | % |

| Metabolic biomarkers | Full Name | Class | UK Biobank n (%) | UK Biobank mean (SD) | ESTHER n (%) | ESTHER mean (SD) | Unit |
|---|---|---|---|---|---|---|---|
| Pyruvate | Pyruvate | Glycolysis related metabolites | 208,489 (99.7) | 0.08 (0.03) | 6,698 (98.2) | 0.20 (0.24) | mmol/l |
| Remnant-C | Remnant cholesterol (non-HDL, non-LDL -cholesterol) | Cholesterol | 209,144 (100) | 1.58 (0.43) | 6,820 (100) | 2.03 (0.52) | mmol/l |
| S-HDL-C | Cholesterol in small HDL | Small HDL (average diameter 8.7 nm) | 209,144 (100) | 0.45 (0.06) | 6,820 (100) | 0.47 (0.08) | mmol/l |
| S-HDL-C-pct | Cholesterol to total lipids ratio in small HDL | Small HDL ratios | 209,144 (100) | 38.31 (1.88) | 6,820 (100) | 39.11 (2.34) | % |
| S-HDL-CE | Cholesteryl esters in small HDL | Small HDL (average diameter 8.7 nm) | 209,144 (100) | 0.33 (0.05) | 6,820 (100) | 0.34 (0.06) | mmol/l |
| S-HDL-CE-pct | Cholesteryl esters to total lipids ratio in small HDL | Small HDL ratios | 209,144 (100) | 28.36 (1.87) | 6,820 (100) | 28.29 (2.54) | % |
| S-HDL-FC | Free cholesterol in small HDL | Small HDL (average diameter 8.7 nm) | 209,144 (100) | 0.12 (0.02) | 6,820 (100) | 0.13 (0.02) | mmol/l |
| S-HDL-FC-pct | Free cholesterol to total lipids ratio in small HDL | Small HDL ratios | 209,144 (100) | 9.95 (0.54) | 6,820 (100) | 10.82 (0.76) | % |
| S-HDL-L | Total lipids in small HDL | Small HDL (average diameter 8.7 nm) | 209,144 (100) | 1.18 (0.16) | 6,820 (100) | 1.20 (0.19) | mmol/l |
| S-HDL-P | Concentration of small HDL particles | Small HDL (average diameter 8.7 nm) | 209,144 (100) | 0.01 (0.00) | 6,820 (100) | 0.01 (0.00) | mmol/l |
| S-HDL-PL | Phospholipids in small HDL | Small HDL (average diameter 8.7 nm) | 209,144 (100) | 0.67 (0.10) | 6,820 (100) | 0.67 (0.12) | mmol/l |
| S-HDL-PL-pct | Phospholipids to total lipids ratio in small HDL | Small HDL ratios | 209,144 (100) | 57.09 (1.22) | 6,820 (100) | 55.74 (1.46) | % |
| S-HDL-TG | Triglycerides in small HDL | Small HDL (average diameter 8.7 nm) | 209,144 (100) | 0.05 (0.02) | 6,820 (100) | 0.06 (0.02) | mmol/l |
| S-HDL-TG-pct | Triglycerides to total lipids ratio in small HDL | Small HDL ratios | 209,144 (100) | 4.60 (1.33) | 6,820 (100) | 5.15 (1.89) | % |
| S-LDL-C | Cholesterol in small LDL | Small LDL (average diameter 18.7 nm) | 209,144 (100) | 0.18 (0.05) | 6,820 (100) | 0.23 (0.05) | mmol/l |
| S-LDL-C-pct | Cholesterol to total lipids ratio in small LDL | Small LDL ratios | 209,144 (100) | 63.18 (2.04) | 6,820 (100) | 63.65 (2.00) | % |
| S-LDL-CE | Cholesteryl esters in small LDL | Small LDL (average diameter 18.7 nm) | 209,144 (100) | 0.13 (0.03) | 6,820 (100) | 0.17 (0.04) | mmol/l |
| S-LDL-CE-pct | Cholesteryl esters to total lipids ratio in small LDL | Small LDL ratios | 209,144 (100) | 45.90 (2.07) | 6,820 (100) | 46.28 (1.91) | % |
| S-LDL-FC | Free cholesterol in small LDL | Small LDL (average diameter 18.7 nm) | 209,144 (100) | 0.05 (0.01) | 6,820 (100) | 0.06 (0.02) | mmol/l |
| S-LDL-FC-pct | Free cholesterol to total lipids ratio in small LDL | Small LDL ratios | 209,144 (100) | 17.28 (2.01) | 6,820 (100) | 17.37 (2.44) | % |
| S-LDL-L | Total lipids in small LDL | Small LDL (average diameter 18.7 nm) | 209,144 (100) | 0.29 (0.07) | 6,820 (100) | 0.36 (0.08) | mmol/l |
| S-LDL-P | Concentration of small LDL particles | Small LDL (average diameter 18.7 nm) | 209,144 (100) | 0.01 (0.01) | 6,820 (100) | 0.01 (0.01) | mmol/l |
| S-LDL-PL | Phospholipids in small LDL | Small LDL (average diameter 18.7 nm) | 209,144 (100) | 0.09 (0.02) | 6,820 (100) | 0.11 (0.02) | mmol/l |
| S-LDL-PL-pct | Phospholipids to total lipids ratio in small LDL | Small LDL ratios | 209,144 (100) | 31.23 (1.75) | 6,820 (100) | 30.95 (1.55) | % |
| S-LDL-TG | Triglycerides in small LDL | Small LDL (average diameter 18.7 nm) | 209,144 (100) | 0.02 (0.01) | 6,820 (100) | 0.02 (0.01) | mmol/l |

| Metabolic biomarkers | Full Name | Class | UK Biobank n (%) | UK Biobank mean (SD) | ESTHER n (%) | ESTHER mean (SD) | Unit |
|---|---|---|---|---|---|---|---|
| S-LDL-TG-pct | Triglycerides to total lipids ratio in small LDL | Small LDL ratios | 209,144 (100) | 5.59 (1.84) | 6,820 (100) | 5.29 (2.33) | % |
| S-VLDL-C | Cholesterol in small VLDL | Small VLDL (average diameter 36.8 nm) | 209,144 (100) | 0.16 (0.05) | 6,820 (100) | 0.21 (0.07) | mmol/l |
| S-VLDL-C-pct | Cholesterol to total lipids ratio in small VLDL | Small VLDL ratios | 209,144 (100) | 37.79 (4.78) | 6,820 (100) | 41.02 (4.64) | % |
| S-VLDL-CE | Cholesteryl esters in small VLDL | Small VLDL (average diameter 36.8 nm) | 209,144 (100) | 0.10 (0.03) | 6,820 (100) | 0.13 (0.04) | mmol/l |
| S-VLDL-CE-pct | Cholesteryl esters to total lipids ratio in small VLDL | Small VLDL ratios | 209,144 (100) | 23.55 (2.82) | 6,820 (100) | 25.73 (2.67) | % |
| S-VLDL-FC | Free cholesterol in small VLDL | Small VLDL (average diameter 36.8 nm) | 209,144 (100) | 0.06 (0.02) | 6,820 (100) | 0.08 (0.02) | mmol/l |
| S-VLDL-FC-pct | Free cholesterol to total lipids ratio in small VLDL | Small VLDL ratios | 209,144 (100) | 14.24 (2.26) | 6,820 (100) | 15.29 (2.29) | % |
| S-VLDL-L | Total lipids in small VLDL | Small VLDL (average diameter 36.8 nm) | 209,144 (100) | 0.42 (0.13) | 6,820 (100) | 0.52 (0.17) | mmol/l |
| S-VLDL-P | Concentration of small VLDL particles | Small VLDL (average diameter 36.8 nm) | 209,144 (100) | 0.01 (0.01) | 6,820 (100) | 0.01 (0.01) | mmol/l |
| S-VLDL-PL | Phospholipids in small VLDL | Small VLDL (average diameter 36.8 nm) | 209,144 (100) | 0.10 (0.03) | 6,820 (100) | 0.13 (0.04) | mmol/l |
| S-VLDL-PL-pct | Phospholipids to total lipids ratio in small VLDL | Small VLDL ratios | 209,144 (100) | 23.60 (2.11) | 6,820 (100) | 24.63 (2.27) | % |
| S-VLDL-TG | Triglycerides in small VLDL | Small VLDL (average diameter 36.8 nm) | 209,144 (100) | 0.16 (0.06) | 6,820 (100) | 0.18 (0.08) | mmol/l |
| S-VLDL-TG-pct | Triglycerides to total lipids ratio in small VLDL | Small VLDL ratios | 209,144 (100) | 38.61 (6.78) | 6,820 (100) | 34.35 (6.77) | % |
| SFA | Saturated fatty acids | Fatty acids | 208,978 (99.9) | 4.19 (0.97) | 6,794 (99.6) | 4.91 (1.12) | mmol/l |
| SFA-pct | Ratio of saturated fatty acids to total fatty acids | Fatty acid ratios | 208,978 (99.9) | 34.08 (1.96) | 6,794 (99.6) | 34.74 (1.86) | % |
| Sphingomyelins | Sphingomyelins | Other lipids | 208,973 (99.9) | 0.46 (0.08) | 6,794 (99.6) | 0.54 (0.08) | mmol/l |
| TG/PG | Ratio of triglycerides to phosphoglycerides | Other lipids | 208,978 (99.9) | 0.58 (0.23) | 6,794 (99.6) | 0.55 (0.22) | ratio |
| Total-BCAA | Total concentration of branched-chain amino acids (leucine + isoleucine + valine) | Branched−chain amino acids | 208,977 (99.9) | 0.37 (0.09) | 6,794 (99.6) | 0.49 (0.10) | mmol/l |
| Total-C | Total cholesterol | Cholesterol | 209,144 (100) | 4.67 (0.98) | 6,820 (100) | 5.72 (1.12) | mmol/l |
| Total-CE | Total esterified cholesterol | Cholesteryl esters | 209,144 (100) | 3.40 (0.70) | 6,820 (100) | 4.14 (0.81) | mmol/l |
| Total-FA | Total fatty acids | Fatty acids | 208,978 (99.9) | 12.23 (2.42) | 6,794 (99.6) | 14.08 (2.76) | mmol/l |
| Total-FC | Total free cholesterol | Free cholesterol | 209,144 (100) | 1.28 (0.28) | 6,820 (100) | 1.58 (0.32) | mmol/l |
| Total-L | Total lipids in lipoprotein particles | Total lipids | 209,144 (100) | 8.99 (1.69) | 6,820 (100) | 10.53 (1.90) | mmol/l |
| Total-P | Total concentration of lipoprotein particles | Lipoprotein particle concentrations | 209,144 (100) | 0.02 (0.00) | 6,820 (100) | 0.02 (0.00) | mmol/l |
| Total-PL | Total phospholipids in lipoprotein particles | Phospholipids | 209,144 (100) | 2.97 (0.49) | 6,820 (100) | 3.36 (0.54) | mmol/l |
| Total-TG | Total triglycerides | Triglycerides | 209,144 (100) | 1.35 (0.59) | 6,820 (100) | 1.45 (0.66) | mmol/l |
| Tyr | Tyrosine | Aromatic amino acids | 208,902 (99.8) | 0.06 (0.01) | 6,781 (99.4) | 0.07 (0.02) | mmol/l |
| Unsaturation | Degree of unsaturation | Fatty acids | 208,978 (99.9) | 1.36 (0.08) | 6,794 (99.6) | 1.35 (0.07) | degree |
| Val | Valine | Branched−chain amino acids | 208,982 (99.9) | 0.21 (0.04) | 6,793 (99.6) | 0.26 (0.05) | mmol/l |
| VLDL-C | VLDL cholesterol | Cholesterol | 209,144 (100) | 0.73 (0.25) | 6,820 (100) | 0.92 (0.30) | mmol/l |

| Metabolic biomarkers | Full Name | Class | UK Biobank n (%) | UK Biobank mean (SD) | ESTHER n (%) | ESTHER mean (SD) | Unit |
|---|---|---|---|---|---|---|---|
| VLDL-CE | Cholesteryl esters in VLDL | Cholesteryl esters | 209,144 (100) | 0.44 (0.15) | 6,820 (100) | 0.56 (0.17) | mmol/l |
| VLDL-FC | Free cholesterol in VLDL | Free cholesterol | 209,144 (100) | 0.30 (0.11) | 6,820 (100) | 0.36 (0.12) | mmol/l |
| VLDL-L | Total lipids in VLDL | Total lipids | 209,144 (100) | 2.16 (0.87) | 6,820 (100) | 2.44 (0.96) | mmol/l |
| VLDL-P | Concentration of VLDL particles | Lipoprotein particle concentrations | 209,144 (100) | 0.01 (0.01) | 6,820 (100) | 0.01 (0.01) | mmol/l |
| VLDL-PL | Phospholipids in VLDL | Phospholipids | 209,144 (100) | 0.48 (0.18) | 6,820 (100) | 0.58 (0.21) | mmol/l |
| VLDL-size | Average diameter for VLDL particles | Lipoprotein particle sizes | 209,144 (100) | 38.66 (1.24) | 6,820 (100) | 38.14 (1.04) | nm |
| VLDL-TG | Triglycerides in VLDL | Triglycerides | 209,144 (100) | 0.95 (0.49) | 6,820 (100) | 0.94 (0.52) | mmol/l |
| XL-HDL-C | Cholesterol in very large HDL | Very large HDL (average diameter 14.3 nm) | 209,144 (100) | 0.08 (0.04) | 6,820 (100) | 0.09 (0.03) | mmol/l |
| XL-HDL-C-pct | Cholesterol to total lipids ratio in very large HDL | Very large HDL ratios | 209,054 (99.9) | 50.07 (4.96) | 6,820 (100) | 50.60 (4.92) | % |
| XL-HDL-CE | Cholesteryl esters in very large HDL | Very large HDL (average diameter 14.3 nm) | 209,144 (100) | 0.06 (0.03) | 6,820 (100) | 0.07 (0.02) | mmol/l |
| XL-HDL-CE-pct | Cholesteryl esters to total lipids ratio in very large HDL | Very large HDL ratios | 209,054 (99.9) | 34.90 (3.54) | 6,820 (100) | 35.78 (3.87) | % |
| XL-HDL-FC | Free cholesterol in very large HDL | Very large HDL (average diameter 14.3 nm) | 209,144 (100) | 0.02 (0.01) | 6,820 (100) | 0.03 (0.01) | mmol/l |
| XL-HDL-FC-pct | Free cholesterol to total lipids ratio in very large HDL | Very large HDL ratios | 209,054 (99.9) | 15.17 (3.68) | 6,820 (100) | 14.71 (2.82) | % |
| XL-HDL-L | Total lipids in very large HDL | Very large HDL (average diameter 14.3 nm) | 209,144 (100) | 0.17 (0.08) | 6,820 (100) | 0.18 (0.07) | mmol/l |
| XL-HDL-P | Concentration of very large HDL particles | Very large HDL (average diameter 14.3 nm) | 209,144 (100) | 0.01 (0.01) | 6,820 (100) | 0.01 (0.01) | mmol/l |
| XL-HDL-PL | Phospholipids in very large HDL | Very large HDL (average diameter 14.3 nm) | 209,144 (100) | 0.08 (0.05) | 6,820 (100) | 0.08 (0.04) | mmol/l |
| XL-HDL-PL-pct | Phospholipids to total lipids ratio in very large HDL | Very large HDL ratios | 209,054 (99.9) | 44.51 (6.75) | 6,820 (100) | 43.17 (6.25) | % |
| XL-HDL-TG | Triglycerides in very large HDL | Very large HDL (average diameter 14.3 nm) | 209,144 (100) | 0.01 (0.00) | 6,820 (100) | 0.01 (0.01) | mmol/l |
| XL-HDL-TG-pct | Triglycerides to total lipids ratio in very large HDL | Very large HDL ratios | 209,054 (99.9) | 5.41 (4.00) | 6,820 (100) | 3.93 (4.32) | % |
| XL-VLDL-C | Cholesterol in very large VLDL | Very large VLDL (average diameter 64 nm) | 209,144 (100) | 0.05 (0.03) | 6,820 (100) | 0.06 (0.03) | mmol/l |
| XL-VLDL-C-pct | Cholesterol to total lipids ratio in very large VLDL | Very large VLDL ratios | 207,296 (99.1) | 29.26 (8.08) | 6,820 (100) | 34.73 (11.19) | % |
| XL-VLDL-CE | Cholesteryl esters in very large VLDL | Very large VLDL (average diameter 64 nm) | 209,144 (100) | 0.03 (0.01) | 6,820 (100) | 0.04 (0.02) | mmol/l |
| XL-VLDL-CE-pct | Cholesteryl esters to total lipids ratio in very large VLDL | Very large VLDL ratios | 207,296 (99.1) | 17.28 (6.78) | 6,820 (100) | 22.03 (9.86) | % |
| XL-VLDL-FC | Free cholesterol in very large VLDL | Very large VLDL (average diameter 64 nm) | 209,144 (100) | 0.02 (0.01) | 6,820 (100) | 0.02 (0.02) | mmol/l |

| Metabolic biomarkers | Full Name | Class | UK Biobank n (%) | UK Biobank mean (SD) | ESTHER n (%) | ESTHER mean (SD) | Unit |
|---|---|---|---|---|---|---|---|
| XL-VLDL-FC-pct | Free cholesterol to total lipids ratio in very large VLDL | Very large VLDL ratios | 207,296 (99.1) | 11.98 (1.73) | 6,820 (100) | 10.65 (4.62) | % |
| XL-VLDL-L | Total lipids in very large VLDL | Very large VLDL (average diameter 64 nm) | 209,144 (100) | 0.21 (0.14) | 6,820 (100) | 0.21 (0.14) | mmol/l |
| XL-VLDL-P | Concentration of very large VLDL particles | Very large VLDL (average diameter 64 nm) | 209,144 (100) | 0.01 (0.01) | 6,820 (100) | 0.01 (0.01) | mmol/l |
| XL-VLDL-PL | Phospholipids in very large VLDL | Very large VLDL (average diameter 64 nm) | 209,144 (100) | 0.04 (0.03) | 6,820 (100) | 0.04 (0.03) | mmol/l |
| XL-VLDL-PL-pct | Phospholipids to total lipids ratio in very large VLDL | Very large VLDL ratios | 207,296 (99.1) | 18.33 (2.67) | 6,820 (100) | 17.56 (5.96) | % |
| XL-VLDL-TG | Triglycerides in very large VLDL | Very large VLDL (average diameter 64 nm) | 209,144 (100) | 0.12 (0.08) | 6,820 (100) | 0.11 (0.09) | mmol/l |
| XL-VLDL-TG-pct | Triglycerides to total lipids ratio in very large VLDL | Very large VLDL ratios | 207,296 (99.1) | 52.42 (8.15) | 6,820 (100) | 45.18 (13.12) | % |
| XS-VLDL-C | Cholesterol in very small VLDL | Very small VLDL (average diameter 31.3 nm) | 209,144 (100) | 0.19 (0.05) | 6,820 (100) | 0.24 (0.07) | mmol/l |
| XS-VLDL-C-pct | Cholesterol to total lipids ratio in very small VLDL | Very small VLDL ratios | 209,144 (100) | 50.96 (4.81) | 6,820 (100) | 51.68 (4.73) | % |
| XS-VLDL-CE | Cholesteryl esters in very small VLDL | Very small VLDL (average diameter 31.3 nm) | 209,144 (100) | 0.13 (0.04) | 6,820 (100) | 0.17 (0.05) | mmol/l |
| XS-VLDL-CE-pct | Cholesteryl esters to total lipids ratio in very small VLDL | Very small VLDL ratios | 209,144 (100) | 34.93 (4.34) | 6,820 (100) | 35.73 (4.36) | % |
| XS-VLDL-FC | Free cholesterol in very small VLDL | Very small VLDL (average diameter 31.3 nm) | 209,144 (100) | 0.06 (0.02) | 6,820 (100) | 0.08 (0.02) | mmol/l |
| XS-VLDL-FC-pct | Free cholesterol to total lipids ratio in very small VLDL | Very small VLDL ratios | 209,144 (100) | 16.02 (0.62) | 6,820 (100) | 15.95 (0.57) | % |
| XS-VLDL-L | Total lipids in very small VLDL | Very small VLDL (average diameter 31.3 nm) | 209,144 (100) | 0.37 (0.09) | 6,820 (100) | 0.48 (0.13) | mmol/l |
| XS-VLDL-P | Concentration of very small VLDL particles | Very small VLDL (average diameter 31.3 nm) | 209,144 (100) | 0.01 (0.01) | 6,820 (100) | 0.01 (0.01) | mmol/l |
| XS-VLDL-PL | Phospholipids in very small VLDL | Very small VLDL (average diameter 31.3 nm) | 209,144 (100) | 0.11 (0.03) | 6,820 (100) | 0.14 (0.04) | mmol/l |
| XS-VLDL-PL-pct | Phospholipids to total lipids ratio in very small VLDL | Very small VLDL ratios | 209,144 (100) | 29.25 (0.92) | 6,820 (100) | 29.28 (1.13) | % |
| XS-VLDL-TG | Triglycerides in very small VLDL | Very small VLDL (average diameter 31.3 nm) | 209,144 (100) | 0.07 (0.02) | 6,820 (100) | 0.09 (0.03) | mmol/l |
| XS-VLDL-TG-pct | Triglycerides to total lipids ratio in very small VLDL | Very small VLDL ratios | 209,144 (100) | 19.79 (4.28) | 6,820 (100) | 19.04 (4.13) | % |

| Metabolic biomarkers | Full Name | Class | UK Biobank n (%) | UK Biobank mean (SD) | ESTHER n (%) | ESTHER mean (SD) | Unit |
|---|---|---|---|---|---|---|---|
| XXL-VLDL-C | Cholesterol in chylomicrons and extremely large VLDL | Chylomicrons and extremely large VLDL (particle diameters from 75 nm upwards) | 209,144 (100) | 0.06 (0.04) | 6,820 (100) | 0.05 (0.05) | mmol/l |
| XXL-VLDL-C-pct | Cholesterol to total lipids ratio in chylomicrons and extremely large VLDL | Chylomicrons and extremely large VLDL ratios | 200,675 (95.9) | 29.75 (12.66) | 6,820 (100) | 35.97 (26.27) | % |
| XXL-VLDL-CE | Cholesteryl esters in chylomicrons and extremely large VLDL | Chylomicrons and extremely large VLDL (particle diameters from 75 nm upwards) | 209,144 (100) | 0.03 (0.02) | 6,820 (100) | 0.03 (0.03) | mmol/l |
| XXL-VLDL-CE-pct | Cholesteryl esters to total lipids ratio in chylomicrons and extremely large VLDL | Chylomicrons and extremely large VLDL ratios | 200,675 (95.9) | 16.88 (9.32) | 6,820 (100) | 22.14 (20.09) | % |
| XXL-VLDL-FC | Free cholesterol in chylomicrons and extremely large VLDL | Chylomicrons and extremely large VLDL (particle diameters from 75 nm upwards) | 209,144 (100) | 0.02 (0.02) | 6,820 (100) | 0.02 (0.02) | mmol/l |
| XXL-VLDL-FC-pct | Free cholesterol to total lipids ratio in chylomicrons and extremely large VLDL | Chylomicrons and extremely large VLDL ratios | 200,675 (95.9) | 12.87 (4.53) | 6,820 (100) | 7.96 (6.61) | % |
| XXL-VLDL-L | Total lipids in chylomicrons and extremely large VLDL | Chylomicrons and extremely large VLDL (particle diameters from 75 nm upwards) | 209,144 (100) | 0.23 (0.20) | 6,820 (100) | 0.18 (0.21) | mmol/l |
| XXL-VLDL-P | Concentration of chylomicrons and extremely large VLDL particles | Chylomicrons and extremely large VLDL (particle diameters from 75 nm upwards) | 209,144 (100) | 0.01 (0.01) | 6,820 (100) | 0.01 (0.01) | mmol/l |
| XXL-VLDL-PL | Phospholipids in chylomicrons and extremely large VLDL | Chylomicrons and extremely large VLDL (particle diameters from 75 nm upwards) | 209,144 (100) | 0.04 (0.03) | 6,820 (100) | 0.03 (0.04) | mmol/l |
| XXL-VLDL-PL-pct | Phospholipids to total lipids ratio in chylomicrons and extremely large VLDL | Chylomicrons and extremely large VLDL ratios | 200,675 (95.9) | 15.39 (4.24) | 6,820 (100) | 11.04 (8.67) | % |
| XXL-VLDL-TG | Triglycerides in chylomicrons and extremely large VLDL | Chylomicrons and extremely large VLDL (particle diameters from 75 nm upwards) | 209,144 (100) | 0.14 (0.13) | 6,820 (100) | 0.10 (0.13) | mmol/l |
| XXL-VLDL-TG-pct | Triglycerides to total lipids ratio in chylomicrons and extremely large VLDL | Chylomicrons and extremely large VLDL ratios | 200,675 (95.9) | 54.86 (13.84) | 6,820 (100) | 49.07 (16.50) | % |

Abbreviations: HDL, high-density lipoprotein; IDL, intermediate-density lipoprotein; LDL, low-density lipoprotein; VLDL, very-low-density lipoprotein; SD, standard deviation.

**Table A3. Associations of LASSO-selected proteomic and metabolomic biomarkers with prevalent frailty at baseline in the UK Biobank and ESTHER cohorts**

| Omics biomarkers | UK Biobank | | ESTHER study | |
|---|---|---|---|---|
| | OR (95% CI) per 1-SD increase [a] | FDR [b] | OR (95% CI) per 1-SD increase [a] | FDR [b] |
| *HGF* | 1.77 (1.68–1.87) | <0.001 | 1.67 (1.44–1.91) | 0.001 |
| *TGFA* | 1.59 (1.51–1.67) | <0.001 | 1.46 (1.31–1.62) | 0.012 |
| *TNFRSF9* | 1.57 (1.50–1.64) | <0.001 | 1.36 (1.14–1.60) | 0.001 |
| *FGF23* | 1.56 (1.49–1.63) | <0.001 | 1.36 (1.05–1.68) | <0.001 |
| *TGFB1* | 1.51 (1.43–1.59) | <0.001 | 1.42 (1.21–1.64) | 0.013 |
| *CDCP1* | 1.51 (1.44–1.58) | <0.001 | 1.42 (1.28–1.57) | 0.002 |
| *TNFRSF11B* | 1.49 (1.42–1.57) | <0.001 | 1.55 (1.44–1.67) | <0.001 |
| *CD5* | 1.48 (1.41–1.56) | <0.001 | 1.45 (1.23–1.68) | 0.020 |
| *VEGFA* | 1.47 (1.40–1.55) | <0.001 | 1.37 (1.19–1.56) | 0.002 |
| *FGF21* | 1.47 (1.40–1.55) | <0.001 | 1.41 (1.32–1.50) | 0.012 |
| *CXCL10* | 1.45 (1.38–1.52) | <0.001 | 1.45 (1.32–1.59) | 0.037 |
| *IL6* | 1.45 (1.38–1.51) | <0.001 | 1.41 (1.36–1.47) | <0.001 |
| *CCL20* | 1.40 (1.34–1.47) | <0.001 | 1.43 (1.37–1.50) | <0.001 |
| *CCL11* | 1.31 (1.25–1.37) | <0.001 | 1.22 (1.04–1.41) | 0.044 |
| *IL18* | 1.29 (1.23–1.36) | <0.001 | 1.28 (1.12–1.45) | 0.025 |
| GlycA | 1.24 (1.19–1.29) | <0.001 | 1.24 (1.12–1.37) | <0.001 |
| *CXCL8* | 1.22 (1.17–1.28) | <0.001 | 1.30 (1.26–1.34) | 0.009 |
| Glucose | 1.20 (1.15–1.24) | <0.001 | 1.36 (1.24–1.48) | <0.001 |
| Lactate | 1.02 (0.97–1.07) | 0.467 | 0.99 (0.90–1.09) | 0.370 |
| His | 1.00 (0.96–1.05) | 0.873 | 1.03 (0.94–1.13) | 0.204 |
| Gln | 0.99 (0.94–1.04) | 0.571 | 1.06 (0.96–1.17) | 0.157 |
| ApoB/ApoA1 | 0.96 (0.91–1.01) | 0.119 | 0.89 (0.80–0.99) | 0.041 |
| LA-pct | 0.95 (0.90–0.99) | 0.043 | 0.85 (0.77–0.94) | 0.001 |
| XL-HDL-FC | 0.87 (0.82–0.92) | <0.001 | 0.82 (0.72–0.93) | 0.023 |
| *DNER* | 0.81 (0.77–0.85) | <0.001 | 0.91 (0.76–1.07) | 0.065 |
| Albumin | 0.77 (0.73–0.81) | <0.001 | 0.75 (0.65–0.85) | <0.001 |

Abbreviations: OR, odds ratio; CI, confidence interval; SD, standard deviation; FDR, false discovery rate; the abbreviations for proteins and metabolites are listed in **Appendix, Table A5**.

[a] Odds ratios (ORs) and 95% confidence intervals (CIs) were estimated using multivariable logistic regression models for prevalent frailty at baseline. Models were adjusted for age, sex, education level, smoking status, physical activity, alcohol consumption, BMI, dyslipidemia, hypertension, diabetes, cardiovascular disease, and history of cancer.

[b] *P* values are two-sided and were adjusted for multiple testing using the Benjamini-Hochberg false discovery rate (FDR) procedure (Glickman et al. 2014).

**Table A4. Associations of LASSO-selected proteomic and metabolomic biomarkers with incident frailty during follow-up in the UK Biobank and ESTHER cohorts**

| Omics biomarkers | UK Biobank | | ESTHER study | |
|---|---|---|---|---|
| | sHR (95% CI) per 1-SD increase [a] | FDR [b] | sHR (95% CI) per 1-SD increase [a] | FDR [b] |
| *TGFA* | 1.42 (1.25–1.59) | <0.001 | 1.26 (1.09–1.43) | 0.002 |
| *IL6* | 1.41 (1.25–1.58) | <0.001 | 1.22 (1.08–1.36) | 0.001 |
| *FGF23* | 1.37 (1.21–1.53) | <0.001 | 1.23 (1.10–1.36) | 0.009 |
| *TNFRSF9* | 1.34 (1.19–1.52) | <0.001 | 1.12 (0.97–1.28) | 0.123 |
| *HGF* | 1.33 (1.15–1.51) | <0.001 | 1.27 (1.10–1.45) | 0.001 |
| *CD5* | 1.32 (1.15–1.50) | <0.001 | 1.13 (0.98–1.28) | 0.088 |
| *CCL20* | 1.27 (1.11–1.43) | <0.001 | 1.15 (1.01–1.29) | 0.030 |
| *CDCP1* | 1.26 (1.10–1.43) | 0.001 | 1.30 (1.13–1.47) | <0.001 |
| *VEGFA* | 1.25 (1.09–1.42) | 0.001 | 1.10 (0.96–1.25) | 0.180 |
| *IL18* | 1.23 (1.07–1.40) | 0.004 | 1.12 (0.97–1.27) | 0.141 |
| *CXCL10* | 1.20 (1.05–1.36) | 0.007 | 1.14 (0.99–1.30) | 0.052 |
| *FGF21* | 1.20 (1.04–1.37) | 0.014 | 0.98 (0.85–1.12) | 0.195 |

| Omics biomarkers | UK Biobank | | ESTHER study | |
|---|---|---|---|---|
| | sHR (95% CI) per 1-SD increase [a] | FDR [b] | sHR (95% CI) per 1-SD increase [a] | FDR [b] |
| *TGFB1* | 1.16 (1.00–1.32) | 0.045 | 1.01 (0.87–1.15) | 0.329 |
| Glucose | 1.15 (1.02–1.28) | 0.025 | 1.17 (1.03–1.31) | 0.016 |
| *TNFRSF11B* | 1.10 (0.95–1.26) | 0.198 | 1.18 (1.02–1.34) | 0.031 |
| *CCL11* | 1.06 (0.91–1.20) | 0.518 | 1.06 (0.92–1.21) | 0.441 |
| GlycA | 1.06 (0.92–1.20) | 0.516 | 1.19 (1.03–1.35) | 0.015 |
| His | 1.02 (0.89–1.16) | 0.339 | 0.98 (0.84–1.13) | 0.488 |
| ApoB/ApoA1 | 0.99 (0.86–1.14) | 0.477 | 1.04 (0.90–1.18) | 0.362 |
| Lactate | 0.95 (0.82–1.09) | 0.424 | 1.04 (0.89–1.20) | 0.119 |
| *CXCL8* | 0.93 (0.80–1.07) | 0.291 | 0.96 (0.81–1.11) | 0.271 |
| LA-pct | 0.93 (0.81–1.06) | 0.255 | 0.84 (0.73–0.95) | 0.008 |
| XL-HDL-FC | 0.93 (0.79–1.07) | 0.265 | 0.94 (0.79–1.10) | 0.068 |
| Gln | 0.87 (0.75–0.99) | 0.037 | 0.80 (0.69–0.91) | 0.001 |
| *DNER* | 0.80 (0.69–0.92) | 0.002 | 0.99 (0.86–1.12) | 0.370 |
| Albumin | 0.79 (0.68–0.91) | 0.001 | 0.82 (0.68–0.95) | 0.009 |

Abbreviations: sHR, subdistribution hazard ratio; CI, confidence interval; SD, standard deviation; FDR, false discovery rate; the abbreviations for proteins and metabolites are listed in **Appendix, Table A5**.

[a] Subdistribution hazard ratios (sHRs) and 95% confidence intervals (CIs) were estimated using multivariable interval-censored competing risks regression models for incident frailty during follow-up, with death treated as a competing event. Median follow-up was 7.8 years in the UK Biobank and 7.7 years in the ESTHER study. All LASSO-selected proteomic and metabolomic biomarkers were entered simultaneously into the same model and mutually adjusted for each other, in addition to adjustment for age, sex, education level, smoking status, physical activity, alcohol consumption, BMI, dyslipidemia, hypertension, diabetes, cardiovascular disease, and history of cancer.

[b] *P* values are two-sided and were adjusted for multiple testing using the Benjamini-Hochberg false discovery rate (FDR) procedure (Glickman et al. 2014).

**Table A5. Overview of LASSO-selected proteomic and metabolomic biomarkers, molecular identifiers, and putative biological functions**

| Short Names | ID | Full Name | Biological Domain |
|---|---|---|---|
| **Proteomic biomarkers** | **UniProt ID** | **Protein Full Name** | **Protein's Biological Function** |
| *CCL3* | P10147 | C-C motif chemokine 3 / MIP-1α | Inflammation and leukocyte recruitment |
| *CCL11* | P51671 | C-C motif chemokine 11 / Eotaxin | Immune signalling and eosinophil chemotaxis |
| *CCL20* | P78556 | C-C motif chemokine ligand 20 | Immune cell trafficking and inflammation |
| *CD5* | P06127 | T-cell surface glycoprotein CD5 | Adaptive immune regulation |
| *CDCP1* | Q9H5V8 | CUB domain-containing protein 1 | Tissue remodelling and cellular stress signalling |
| *CXCL10* | P02778 | C-X-C motif chemokine 10 | Immune activation and leukocyte recruitment |
| *CXCL8* | P10145 | C-X-C motif chemokine ligand 8 (IL-8) | Neutrophil recruitment and inflammation |
| *DNER* | Q8NFT8 | Delta and Notch-like epidermal growth factor-related receptor | Cell differentiation and Notch signalling |
| *FGF21* | Q9NSA1 | Fibroblast growth factor 21 | Energy metabolism and metabolic stress response |
| *FGF23* | Q9GZV9 | Fibroblast growth factor 23 | Endocrine regulation and mineral metabolism |
| *HGF* | P14210 | Hepatocyte growth factor | Tissue repair and regeneration |
| *IL12B* | P29460 | Interleukin-12 subunit beta | Immune activation and inflammatory signalling |
| *IL18* | Q14116 | Interleukin-18 | Innate immune signalling |
| *IL6* | P05231 | Interleukin-6 | Chronic inflammation and immune activation |
| *LIFR* | P42702 | Leukemia inhibitory factor receptor | Cytokine signalling and tissue homeostasis |
| *TGFA* | P01135 | Protransforming growth factor alpha | Growth factor signalling and tissue repair |
| *TGFB1* | P01137 | Transforming growth factor beta-1 | Tissue remodelling and immune regulation |
| *TNFRSF11B* | O00300 | Tumor necrosis factor receptor superfamily member 11B (Osteoprotegerin, OPG) | Bone metabolism and inflammatory signalling |
| *TNFRSF9* | Q07011 | Tumor necrosis factor receptor superfamily member 9 (CD137/4-1BB) | T-cell activation and immune regulation |
| *VEGFA* | P15692 | Vascular endothelial growth factor A | Angiogenesis and tissue repair |
| **Metabolomic biomarkers** | **HMDB ID** | **Metabolite Full Name** | **Metabolic Biological Function** |
| Albumin | HMDB0000497 | Albumin | Nutritional and metabolic status |
| ApoB-by-ApoA1 | / | Ratio of apolipoprotein B to apolipoprotein A1 | Lipoprotein metabolism and cardiovascular risk |
| Gln | HMDB0000641 | Glutamine | Amino acid metabolism and energy homeostasis |
| Glucose | HMDB0000122 | D-Glucose | Glucose metabolism |
| GlycA | / | Glycoprotein acetyls | Chronic systemic inflammation |
| His | HMDB0000177 | Histidine | Amino acid metabolism |
| LA-pct | / | Ratio of linoleic acid to total fatty acids | Fatty acid metabolism |
| Lactate | HMDB0000190 | Lactic acid / Lactate | Glycolysis and bioenergetic stress |
| S-HDL-PL | / | Phospholipids in small HDL | HDL metabolism |
| XL-HDL-FC | / | Free cholesterol in very large HDL | HDL cholesterol transport |

Abbreviations: LASSO, Least Absolute Shrinkage and Selection Operator; UniProt, Universal Protein Resource; HMDB, Human Metabolome Database.

Notes: Protein annotations are cross-referenced using UniProt identifiers (https://www.uniprot.org/). Metabolite annotations are cross-referenced using HMDB identifiers (https://hmdb.ca/metabolites/).

**Table A6. β-coefficients of the Traditional Risk factor, Inflammation, and Metabolism (TRIM) based models for prevalent frailty identification (TRIM-FI) and incident frailty prediction (TRIM-FP)**

| Characteristics | TRIM-FI | TRIM-FP |
|---|---|---|
| **Age (years)** | 0.087 | 0.069 |
| **Sex** | | |
| Women | Ref. | Ref. |
| Men | -0.096 | -1.020 |
| **Education** | | |
| Medium or low | Ref. | Ref. |
| High | -0.064 | -0.125 |
| **Smoking status** | | |
| Never smoker | Ref. | Ref. |
| Former smoker | 0.145 | 0.201 |
| Current smoker | 0.098 | 0.117 |
| **Physical activity** | | |
| Low | Ref. | Ref. |
| Moderate | -0.032 | -0.084 |
| High | -0.118 | -0.155 |
| **Alcohol consumption** | | |
| Abstainer | Ref. | Ref. |
| Low | 0.146 | 0.191 |
| Medium or high | 0.054 | 0.079 |
| **BMI** ($kg/m^2$) | | |
| < 21.5 | 0.121 | 0.160 |
| ≥ 21.5 – < 25 | Ref. | Ref. |
| ≥ 25 – < 30 | 0.386 | 0.401 |
| ≥ 30 – < 35 | 0.672 | 0.746 |
| ≥ 35 | 1.027 | 1.053 |
| **Dyslipidemia** | | |
| No | Ref. | Ref. |
| Yes | 0.818 | 0.866 |
| **Hypertension** | | |
| No | Ref. | Ref. |
| Yes | 0.821 | 0.580 |
| **Diabetes** | | |
| No | Ref. | Ref. |
| Yes | 0.598 | 0.462 |
| **Cardiovascular disease** | | |
| No | Ref. | Ref. |
| Yes | 0.563 | 0.340 |
| **History of cancer** | | |
| No | Ref. | Ref. |
| Yes | 0.845 | 0.568 |
| **Proteomic biomarkers** | | |
| *CCL11* | 0.135 | 0.059 |
| *CCL20* | 0.028 | 0.106 |
| *CD5* | 0.042 | 0.079 |
| *CDCP1* | 0.163 | 0.139 |
| *CXCL10* | 0.136 | 0.007 |
| *CXCL8* | 0.090 | 0.174 |
| *DNER* | -0.170 | -0.194 |
| *FGF21* | 0.122 | 0.091 |
| *FGF23* | 0.207 | 0.229 |
| *HGF* | 0.114 | 0.044 |
| *IL18* | 0.110 | 0.034 |
| *IL6* | 0.114 | 0.208 |
| *TGFA* | 0.149 | 0.187 |
| *TGFB1* | 0.058 | 0.024 |
| *TNFRSF11B* | 0.129 | 0.008 |

| Characteristics | TRIM-FI | TRIM-FP |
|---|---|---|
| *TNFRSF9* | 0.181 | 0.117 |
| *VEGFA* | 0.109 | 0.123 |
| **Metabolomic biomarkers** | | |
| Albumin | -0.171 | -0.219 |
| ApoB-by-ApoA1 | -0.081 | -0.029 |
| Glucose | 0.076 | 0.127 |
| Gln | -0.019 | -0.139 |
| GlycA | 0.064 | 0.022 |
| His | -0.005 | -0.067 |
| LA-pct | -0.004 | -0.093 |
| Lactate | 0.022 | -0.099 |
| XL-HDL-FC | -0.086 | -0.030 |

Abbreviations: Ref., reference category for categorical variables; TRIM-FI, Traditional Risk factor, Inflammation, and Metabolism based Frailty Identification model; TRIM-FP, Traditional Risk factor, Inflammation, and Metabolism based Frailty Prediction model; the abbreviations for proteins and metabolites are listed in **Appendix, Table A5**.

Notes: $\beta$-coefficients represent the change in log-odds (for prevalent frailty) or log-hazard (for incident frailty) per 1-standard deviation increase in log-transformed biomarker levels.

**Table A7. β-coefficients of the variables included in the risk scores for 10-year all-cause mortality prediction derived in 70% of the UK Biobank and used with the same values in the internal and external validation**

| Characteristics | Men aged 50–59 years | | | Men aged 60–69 years | | | Women aged 50–59 years | | | Women aged 60–69 years | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Covariates | Biomarkers | Covariates + Biomarkers | Covariates | Biomarkers | Covariates + Biomarkers | Covariates | Biomarkers | Covariates + Biomarkers | Covariates | Biomarkers | Covariates + Biomarkers |
| **Age (per 10 years)** | 0.687 | | 0.061 | 0.943 | | 0.083 | 0.553 | | 0.055 | 0.972 | | 0.091 |
| **Sex** | | | | | | | | | | | | |
| Women | Ref. | | Ref. | Ref. | | Ref. | Ref. | | Ref. | Ref. | | Ref. |
| Men | - | | - | - | | - | - | | - | - | | - |
| **Education level** | | | | | | | | | | | | |
| Medium or low | Ref. | | Ref. | Ref. | | Ref. | Ref. | | Ref. | Ref. | | Ref. |
| High | -0.003 | | -0.042 | -0.037 | | -0.009 | -0.006 | | -0.015 | -0.094 | | -0.078 |
| **Smoking status** | | | | | | | | | | | | |
| Never smoker | Ref. | | Ref. | Ref. | | Ref. | Ref. | | Ref. | Ref. | | Ref. |
| Former smoker | 0.267 | | 0.215 | 0.309 | | 0.270 | 0.259 | | 0.218 | 0.327 | | 0.281 |
| Current smoker | 1.074 | | 0.773 | 0.889 | | 0.651 | 0.928 | | 0.666 | 1.001 | | 0.784 |
| **Physical activity** | | | | | | | | | | | | |
| Low | Ref. | | Ref. | Ref. | | Ref. | Ref. | | Ref. | Ref. | | Ref. |
| Moderate | -0.149 | | -0.075 | -0.213 | | -0.156 | -0.140 | | -0.094 | -0.240 | | -0.201 |
| High | -0.193 | | -0.107 | -0.248 | | -0.179 | -0.281 | | -0.220 | -0.327 | | -0.271 |
| **Alcohol consumption** | | | | | | | | | | | | |
| Abstainer | Ref. | | Ref. | Ref. | | Ref. | Ref. | | Ref. | Ref. | | Ref. |
| Low | 0.198 | | 0.237 | 0.191 | | 0.138 | 0.237 | | 0.141 | 0.229 | | 0.132 |
| Medium or high | 0.318 | | 0.243 | 0.293 | | 0.217 | 0.354 | | 0.275 | 0.241 | | 0.165 |
| **Body mass index ($kg/m^2$)** | | | | | | | | | | | | |
| < 21.5 | 0.649 | | 0.376 | 0.542 | | 0.336 | 0.266 | | 0.210 | 0.227 | | 0.168 |
| ≥ 21.5 – < 25 | Ref. | | Ref. | Ref. | | Ref. | Ref. | | Ref. | Ref. | | Ref. |
| ≥ 25 – < 30 | -0.131 | | 0.009 | -0.025 | | 0.063 | -0.034 | | -0.136 | -0.039 | | -0.078 |
| ≥ 30 – < 35 | -0.054 | | 0.038 | -0.071 | | 0.132 | -0.026 | | -0.249 | -0.108 | | -0.221 |
| ≥ 35 | 0.350 | | 0.336 | 0.286 | | 0.218 | 0.393 | | 0.115 | 0.227 | | 0.122 |
| **Treated dyslipidemia** | | | | | | | | | | | | |
| No | Ref. | | Ref. | Ref. | | Ref. | Ref. | | Ref. | Ref. | | Ref. |
| Yes | -0.058 | | -0.032 | -0.111 | | -0.078 | -0.272 | | -0.147 | -0.064 | | -0.022 |
| **Hypertension** | | | | | | | | | | | | |
| No | Ref. | | Ref. | Ref. | | Ref. | Ref. | | Ref. | Ref. | | Ref. |
| Yes | 0.341 | | 0.190 | 0.189 | | 0.130 | 0.211 | | 0.125 | 0.123 | | 0.056 |
| **Diabetes** | | | | | | | | | | | | |
| No | Ref. | | Ref. | Ref. | | Ref. | Ref. | | Ref. | Ref. | | Ref. |

| Characteristics | Men aged 50–59 years | | | Men aged 60–69 years | | | Women aged 50–59 years | | | Women aged 60–69 years | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Covariates | Biomarkers | Covariates + Biomarkers | Covariates | Biomarkers | Covariates + Biomarkers | Covariates | Biomarkers | Covariates + Biomarkers | Covariates | Biomarkers | Covariates + Biomarkers |
| Yes | 0.610 | | 0.336 | 0.401 | | 0.250 | 0.295 | | 0.125 | 0.360 | | 0.199 |
| **Cardiovascular disease** | | | | | | | | | | | | |
| No | Ref. | | Ref. | Ref. | | Ref. | Ref. | | Ref. | Ref. | | Ref. |
| Yes | 0.636 | | 0.509 | 0.553 | | 0.441 | 0.465 | | 0.363 | 0.534 | | 0.459 |
| **Cancer** | | | | | | | | | | | | |
| No | Ref. | | Ref. | Ref. | | Ref. | Ref. | | Ref. | Ref. | | Ref. |
| Yes | 1.186 | | 1.010 | 0.712 | | 0.666 | 1.402 | | 1.322 | 0.754 | | 0.732 |
| **Biomarkers** | | | | | | | | | | | | |
| Acetate | | 0.031 | 0.029 | | 0.043 | 0.040 | | - | - | | - | - |
| Acetoacetate | | - | - | | - | - | | 0.046 | 0.034 | | 0.056 | 0.047 |
| Acetone | | 0.066 | 0.085 | | 0.074 | 0.083 | | - | - | | - | - |
| Albumin | | -0.191 | -0.164 | | -0.075 | -0.060 | | -0.204 | -0.186 | | -0.063 | -0.040 |
| bOHbutyrate | | 0.012 | 0.039 | | 0.015 | 0.005 | | - | - | | - | - |
| Citrate | | 0.068 | 0.071 | | 0.071 | 0.053 | | - | - | | - | - |
| Creatinine | | 0.044 | 0.055 | | 0.043 | 0.030 | | 0.052 | 0.045 | | 0.065 | 0.069 |
| Glucose | | 0.187 | 0.122 | | 0.127 | 0.088 | | 0.038 | 0.021 | | 0.085 | 0.065 |
| GlycA | | 0.351 | 0.273 | | 0.305 | 0.252 | | 0.376 | 0.290 | | 0.335 | 0.281 |
| HDL size | | 0.051 | 0.073 | | - | - | | - | - | | - | - |
| His | | -0.066 | -0.046 | | -0.046 | -0.035 | | -0.016 | -0.041 | | -0.037 | -0.029 |
| IDL-CE-pct | | -0.134 | -0.071 | | -0.039 | -0.022 | | - | - | | - | - |
| L-LDL-CE-pct | | -0.004 | -0.006 | | | | | - | - | | - | - |
| LA-pct | | -0.277 | -0.167 | | -0.233 | -0.162 | | -0.310 | -0.241 | | -0.257 | -0.204 |
| Lactate | | 0.095 | 0.055 | | - | - | | - | - | | - | - |
| Leu | | - | - | | - | - | | -0.100 | -0.100 | | - | - |
| M-LDL-TG-pct | | - | - | | | | | 0.184 | 0.140 | | 0.114 | 0.055 |
| Omega-6/Omega-3 | | 0.246 | 0.190 | | 0.223 | 0.187 | | 0.266 | 0.218 | | 0.251 | 0.221 |
| PUFA | | - | - | | - | - | | - | - | | -0.050 | -0.011 |
| S-HDL-CE | | - | - | | -0.172 | -0.144 | | - | - | | -0.100 | -0.111 |
| S-LDL-CE | | - | - | | -0.085 | -0.018 | | - | - | | - | - |
| Tyr | | 0.165 | 0.142 | | 0.146 | 0.142 | | - | - | | - | - |
| Val | | -0.331 | -0.287 | | -0.281 | -0.259 | | -0.095 | -0.044 | | -0.176 | -0.134 |
| VLDL-size | | -0.290 | -0.175 | | -0.201 | -0.177 | | -0.229 | -0.204 | | -0.215 | -0.190 |
| XL-HDL-FC | | 0.225 | 0.201 | | 0.041 | 0.070 | | - | - | | - | - |
| XXL-VLDL-PL-pct | | 0.226 | 0.211 | | - | - | | - | - | | - | - |

Notes: The method used to calculate the risk score was based on previously published literature. The names of the metabolic biomarkers are indicated by representative abbreviations, and the full names of which are detailed in **Appendix, Table A2**.

**Table A8. Equations used to calculate the metabolomics-based mortality risk prediction (MetaboMR) clocks 1 and 2**

| Subgroup | Intercept and *β*-coefficients |
|---|---|
| **Step 1: Elastic net models** [a] | |
| **To derive MetaboMR clock1** | |
| Men aged 50–59 years | Raw metabolomic age1 = 55.123 + 0.017×Acetate + 0.024×Acetone - 0.259×Albumin + 0.040×bOHbutyrate + 0.224×Citrate + 0.001×Creatinine + 0.077×Glucose + 0.028×GlycA + 0.166×HDL-size - 0.062×His - 0.035×IDL-CE-pct - 0.014×L-LDL-CE-pct - 0.014×LA-pct + 0.012×Lactate + 0.257×Omega-6/Omega-3 + 0.200×Tyr - 0.203×Val - 0.093×VLDL-size + 0.135×XL-HDL-FC + 0.054×XXL-VLDL-PL-pct |
| Men aged 60–69 years | Raw metabolomic age1 = 64.116 + 0.011×Acetate + 0.105×Acetone - 0.212×Albumin + 0.186×bOHbutyrate + 0.251×Citrate + 0.120×Creatinine + 0.064×Glucose + 0.060×GlycA - 0.058×His - 0.002×IDL-CE-pct - 0.038×LA-pct + 0.222×Omega-6/Omega-3 - 0.010×S-HDL-CE - 0.322×S-LDL-CE + 0.135×Tyr - 0.180×Val - 0.021×VLDL-size + 0.026×XL-HDL-FC |
| Women aged 50–59 years | Raw metabolomic age1 = 54.834 + 0.102×Acetoacetate - 0.047×Albumin + 0.070×Creatinine + 0.129×Glucose + 0.129×GlycA - 0.078×His + 0.034×LA-pct - 0.181×Leu + 0.063×M-LDL-TG-pct + 0.359×Omega-6/Omega-3 - 0.138×Val - 0.085×VLDL-size |
| Women aged 60–69 years | Raw metabolomic age1 = 63.907 + 0.079×Acetoacetate - 0.232×Albumin + 0.179×Creatinine + 0.099×Glucose + 0.081×GlycA - 0.071×His - 0.015×LA-pct + 0.188×M-LDL-TG-pct + 0.169×Omega-6/Omega-3 - 0.023×PUFA - 0.060×S-HDL-CE - 0.143×Val - 0.124×VLDL-size |
| **To derive MetaboMR clock2** | |
| Men aged 50–59 years | Raw metabolomic age2 = 54.532 - 0.001×(Education=High) + 0.549×(Smoking=Former) + 0.063×(Smoking=Current) - 0.063×(Physical activity=Moderate) - 0.119×(Physical activity=High) + 0.146×(Alcohol consumption=Low) + 0.109×(Alcohol consumption=Medium or High) + 0.100×(BMI< 21.5 kg/m2) + 0.042×(BMI ≥ 25 – < 30 kg/m2) + 0.241×(BMI ≥ 30 – < 35 kg/m2) + 0.435×(BMI ≥ 35 kg/m2) - 0.673×(Treated dyslipidemia) + 0.510×Hypertension + 0.007×Diabetes + 0.967×(Cardiovascular disease) + 0.881×Cancer + 0.013×Acetate + 0.002×Acetone - 0.278×Albumin + 0.043×bOHbutyrate + 0.252×Citrate + 0.003×Creatinine + 0.016×Glucose + 0.021×GlycA + 0.108×HDL-size - 0.040×His - 0.025×IDL-CE-pct - 0.012×L-LDL-CE-pct - 0.047×LA-pct + 0.026×Lactate + 0.214×Omega-6/Omega-3 + 0.165×Tyr - 0.157×Val - 0.010×VLDL-size + 0.067×XL-HDL-FC + 0.050×XXL-VLDL-PL-pct |
| Men aged 60–69 years | Raw metabolomic age2 = 63.589 - 0.116×(Education=High) + 0.347×(Smoking=Former) + 0.145×(Smoking=Current) - 0.040×(Physical activity=Moderate) - 0.124×(Physical activity=High) + 0.047×(Alcohol consumption=Low) + 0.257×(Alcohol consumption=Medium or High) + 0.010×(BMI< 21.5 kg/m2) + 0.001×(BMI ≥ 25 – < 30 kg/m2) + 0.162×(BMI ≥ 30 – < 35 kg/m2) + 0.531×(BMI ≥ 35 kg/m2) - 0.389×(Treated dyslipidemia) + 0.457×Hypertension + 0.130×Diabetes + 0.913×(Cardiovascular disease) + 0.884×Cancer + 0.003×Acetate + 0.109×Acetone - 0.236×Albumin + 0.191×bOHbutyrate + 0.239×Citrate + 0.089×Creatinine + 0.026×Glucose + 0.037×GlycA - 0.042×His - 0.040×IDL-CE-pct - 0.036×LA-pct + 0.181×Omega-6/Omega-3 - |

| Subgroup | Intercept and *β*-coefficients |
|---|---|
| **Step 1: Elastic net models** [a] | |
| **To derive MetaboMR clock1** | |
| | 0.024×S-HDL-CE - 0.187×S-LDL-CE + 0.143×Tyr - 0.163×Val - 0.003×VLDL-size + 0.125×XL-HDL-FC |
| Women aged 50–59 years | Raw metabolomic age2 = 54.480 - 0.006×(Education=High) + 0.363×(Smoking=Former) + 0.154×(Smoking=Current) - 0.070×(Physical activity=Moderate) - 0.004×(Physical activity=High) + 0.058×(Alcohol consumption=Low) + 0.055×(Alcohol consumption=Medium or High) + 0.008×(BMI< 21.5 kg/m2) - 0.107×(BMI ≥ 25 – < 30 kg/m2) - 0.311×(BMI ≥ 30 – < 35 kg/m2) + 0.535×(BMI ≥ 35 kg/m2) - 0.787×(Treated dyslipidemia) + 0.582×Hypertension + 0.058×Diabetes + 0.874×(Cardiovascular disease) + 0.731×Cancer + 0.098×Acetoacetate - 0.079×Albumin + 0.065×Creatinine + 0.083×Glucose + 0.127×GlycA - 0.063×His + 0.090×LA-pct - 0.195×Leu + 0.118×M-LDL-TG-pct + 0.341×Omega-6/Omega-3 - 0.176×Val - 0.102×VLDL-size |
| Women aged 60–69 years | Raw metabolomic age2 = 63.620 - 0.190×(Education=High) + 0.067×(Smoking=Former) + 0.465×(Smoking=Current) - 0.085×(Physical activity=Moderate) - 0.087×(Physical activity=High) + 0.164×(Alcohol consumption=Low) + 0.356×(Alcohol consumption=Medium or High) + 0.037×(BMI< 21.5 kg/m2) - 0.045×(BMI ≥ 25 – < 30 kg/m2) - 0.132×(BMI ≥ 30 – < 35 kg/m2) + 0.639×(BMI ≥ 35 kg/m2) - 0.787×(Treated dyslipidemia) + 0.671×Hypertension + 0.083×Diabetes + 0.848×(Cardiovascular disease) + 0.732×Cancer + 0.080×Acetoacetate - 0.274×Albumin + 0.157×Creatinine + 0.056×Glucose + 0.029×GlycA - 0.047×His - 0.062×LA-pct + 0.073×M-LDL-TG-pct + 0.088×Omega-6/Omega-3 - 0.146×PUFA - 0.009×S-HDL-CE - 0.122×Val - 0.151×VLDL-size |
| **Step 2: Linear regression models** [b] | |
| **To derive MetaboMR clock1** | |
| Men aged 50–59 years | MetaboMR clock1 = -1.128 + 1.021 × Raw metabolomic age1 |
| Men aged 60–69 years | MetaboMR clock1 = -0.819 + 1.013 × Raw metabolomic age1 |
| Women aged 50–59 years | MetaboMR clock1 = -0.752 + 1.003 × Raw metabolomic age1 |
| Women aged 60–69 years | MetaboMR clock1 = -0.559 + 1.009 × Raw metabolomic age1 |
| **To derive MetaboMR clock2** | |
| Men aged 50–59 years | MetaboMR clock2 = -1.034 + 1.019 × Raw metabolomic age2 |
| Men aged 60–69 years | MetaboMR clock2 = -0.941 + 1.015 × Raw metabolomic age2 |
| Women aged 50–59 years | MetaboMR clock2 = -0.803 + 1.017 × Raw metabolomic age2 |
| Women aged 60–69 years | MetaboMR clock2 = -0.513 + 1.005 × Raw metabolomic age2 |

Abbreviations: MetaboMR clock1, metabolomics-based mortality risk prediction clock 1; MetaboMR clock2, metabolomics-based mortality risk prediction clock 2.

Notes: The names of the metabolic biomarkers are indicated by representative abbreviations, and the full names of which are detailed in **Appendix, Table A2**.

[a] The intercept and coefficients stem from an elastic net model, in which chronological age was the dependent variable. Raw metabolomic age1 was derived solely from the selected metabolomic biomarkers in each stratum. Raw metabolomic age2 was derived from both the selected metabolomic biomarkers and traditional mortality risk factors in each stratum. The estimates stem from the training set (70% of the UK Biobank).

[b] The intercept and coefficients stem from a linear regression model with chronological age as the dependent variable and the raw metabolomic age 1 or 2 as the independent variable. The estimates stem from the training set (70% of the UK Biobank).

**Table A9. LASSO-selected proteins identified from the Olink Target Inflammation and Olink Explore Proteomic Panels, including gene symbols, UniProt identifiers, and full names**

| Subgroups | Selected Gene Symbol | UniProt ID | Full Name |
|---|---|---|---|
| **Olink Target Inflammation Panel** | | | |
| Men | *CCL20* | *P78556* | C-C motif chemokine 20 |
| Men | *CSF1* | *P09603* | Macrophage colony-stimulating factor 1 |
| Men | *CXCL9* | *Q07325* | C-X-C motif chemokine 9 |
| Men | *DNER* | *Q8NFT8* | Delta and Notch-like epidermal growth factor-related receptor |
| Women | *CXCL8* | *P10145* | Interleukin-8 |
| Women | *FGF21* | *Q9NSA1* | Fibroblast growth factor 21 |
| Both | *CDCP1* | *Q9H5V8* | CUB domain-containing protein 1 |
| Both | *FGF23* | *Q9GZV9* | Fibroblast growth factor 23 |
| Both | *HGF* | *P14210* | Hepatocyte growth factor |
| Both | *IL6* | *P05231* | Interleukin-6 |
| Both | *KITLG* | *P21583* | Kit ligand |
| Both | *TNFRSF11B* | *O00300* | Tumor necrosis factor receptor superfamily member 11B |
| Both | *TNFRSF9* | *Q07011* | Tumor necrosis factor receptor superfamily member 9 |
| **Olink Explore Proteomic Panels** | | | |
| Men | *AGER* | *Q15109* | Advanced glycosylation end product-specific receptor |
| Men | *C7* | *P10643* | Complement component C7 |
| Men | *EGFR* | *P00533* | Epidermal growth factor receptor |
| Men | *IGFBP2* | *P18065* | Insulin-like growth factor-binding protein 2 |
| Men | *MMP12* | *P39900* | Macrophage metalloelastase |
| Men | *MSTN* | *O14793* | Growth/differentiation factor 8 |
| Men | *PIGR* | *P01833* | Polymeric immunoglobulin receptor |
| Men | *PLAUR* | *Q03405* | Urokinase plasminogen activator surface receptor |
| Men | *RET* | *P07949* | Proto-oncogene tyrosine-protein kinase receptor Ret |
| Men | *SERPINA3* | *P01011* | Alpha-1-antichymotrypsin |
| Men | *TNFRSF10B* | *O14763* | Tumor necrosis factor receptor superfamily member 10B |
| Men | *TNFSF13B* | *Q9Y275* | Tumor necrosis factor ligand superfamily member 13B |
| Women | *ANGPT2* | *O15123* | Angiopoietin-2 |
| Women | *CALCA* | *P01258* | Calcitonin |
| Women | *GAL* | *P22466* | Galanin peptides |
| Women | *HAVCR1* | *Q96D42* | Hepatitis A virus cellular receptor 1 |
| Women | *IGFBP7* | *Q16270* | Insulin-like growth factor-binding protein 7 |
| Women | *IL19* | *Q9UHD0* | Interleukin-19 |
| Women | *ITGAV* | *P06756* | Integrin alpha-V |
| Women | *TNR* | *Q92752* | Tenascin-R |
| Both | *AREG* | *P15514* | Amphiregulin |
| Both | *CEACAM5* | *P06731* | Carcinoembryonic antigen-related cell adhesion molecule 5 |
| Both | *GDF15* | *Q99988* | Growth/differentiation factor 15 |
| Both | *IL6* | *P05231* | Interleukin-6 |
| Both | *NEFL* | *P07196* | Neurofilament light polypeptide |
| Both | *NTproBNP* | *P16860* | N-terminal prohormone of brain natriuretic peptide |
| Both | *WFDC2* | *Q14508* | WAP four-disulfide core domain protein 2 |

Notes: Detailed functional information, including primary sequences and biological characterization for each protein, is cross-referenced using the UniProt database identifiers (https://www.uniprot.org/).

**Table A10. β-coefficients of the variables included in the risk scores for 10-year all-cause mortality prediction derived in 70% of the UK Biobank and used with the same values in the internal and external validation sets**

| Characteristics | Olink Target Inflammation Panel | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Men | | | Women | | | | |
| | Covariates | Biomarkers | Covariates + Biomarkers | Covariates | Biomarkers | Covariates + Biomarkers | | |
| **Age (per 10 years)** | 0.852 | | 0.622 | 0.869 | | 0.739 | | |
| **Sex** | | | | | | | | |
| Women | Ref. | | Ref. | Ref. | | Ref. | | |
| Men | - | | - | - | | - | | |
| **Education** | | | | | | | | |
| Medium or low | Ref. | | Ref. | Ref. | | Ref. | | |
| High | -0.128 | | -0.085 | -0.069 | | -0.086 | | |
| **Smoking status** | | | | | | | | |
| Never smoker | Ref. | | Ref. | Ref. | | Ref. | | |
| Former smoker | 0.372 | | 0.258 | 0.274 | | 0.213 | | |
| Current smoker | 1.120 | | 0.821 | 1.095 | | 0.772 | | |
| **Physical activity** | | | | | | | | |
| Low | Ref. | | Ref. | Ref. | | Ref. | | |
| Moderate | -0.016 | | -0.109 | -0.126 | | -0.102 | | |
| High | -0.036 | | -0.177 | -0.459 | | -0.381 | | |
| **Alcohol consumption** | | | | | | | | |
| Abstainer | Ref. | | Ref. | Ref. | | Ref. | | |
| Low | 0.070 | | 0.029 | 0.032 | | 0.075 | | |
| Medium or high | 0.130 | | 0.065 | 0.160 | | 0.114 | | |
| **BMI** ($kg/m^2$) | | | | | | | | |
| < 21.5 | 0.554 | | 0.457 | 0.151 | | 0.116 | | |
| ≥ 21.5 – < 25 | Ref. | | Ref. | Ref. | | Ref. | | |
| ≥ 25 – < 30 | -0.049 | | -0.152 | -0.179 | | -0.282 | | |
| ≥ 30 – < 35 | -0.105 | | -0.172 | -0.247 | | -0.528 | | |
| ≥ 35 | 0.254 | | 0.280 | 0.122 | | 0.322 | | |
| **Dyslipidemia** | | | | | | | | |
| No | Ref. | | Ref. | Ref. | | Ref. | | |
| Yes | 0.021 | | 0.044 | 0.306 | | 0.252 | | |
| **Hypertension** | | | | | | | | |
| No | Ref. | | Ref. | Ref. | | Ref. | | |
| Yes | 0.151 | | 0.069 | 0.409 | | 0.321 | | |
| **Diabetes** | | | | | | | | |
| No | Ref. | | Ref. | Ref. | | Ref. | | |
| Yes | 0.559 | | 0.356 | 0.226 | | 0.064 | | |
| **Cardiovascular disease** | | | | | | | | |
| No | Ref. | | Ref. | Ref. | | Ref. | | |
| Yes | 0.580 | | 0.475 | 0.323 | | 0.235 | | |
| **History of cancer** | | | | | | | | |
| No | Ref. | | Ref. | Ref. | | Ref. | | |
| Yes | 0.890 | | 0.769 | 0.912 | | 0.869 | | |
| **Biomarkers** | | | | | | | | |

| Characteristics | Olink Explore Proteomic Panel | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Men | | | Women | | | | |
| | Covariates | Biomarkers | Covariates + Biomarkers | Covariates | Biomarkers | Covariates + Biomarkers | | |
| **Age (per 10 years)** | 0.852 | | 0.250 | 0.869 | | 0.517 | | |
| **Sex** | | | | | | | | |
| Women | Ref. | | Ref. | Ref. | | Ref. | | |
| Men | - | | - | - | | - | | |
| **Education** | | | | | | | | |
| Medium or low | Ref. | | Ref. | Ref. | | Ref. | | |
| High | -0.128 | | -0.081 | -0.069 | | -0.099 | | |
| **Smoking status** | | | | | | | | |
| Never smoker | Ref. | | Ref. | Ref. | | Ref. | | |
| Former smoker | 0.372 | | 0.144 | 0.274 | | 0.138 | | |
| Current smoker | 1.120 | | 0.222 | 1.095 | | 0.207 | | |
| **Physical activity** | | | | | | | | |
| Low | Ref. | | Ref. | Ref. | | Ref. | | |
| Moderate | -0.016 | | -0.130 | -0.126 | | -0.105 | | |
| High | -0.036 | | -0.138 | -0.459 | | -0.373 | | |
| **Alcohol consumption** | | | | | | | | |
| Abstainer | Ref. | | Ref. | Ref. | | Ref. | | |
| Low | 0.070 | | 0.018 | 0.032 | | 0.012 | | |
| Medium or high | 0.130 | | 0.087 | 0.160 | | 0.042 | | |
| **BMI** ($kg/m^2$) | | | | | | | | |
| < 21.5 | 0.554 | | 0.170 | 0.151 | | 0.131 | | |
| ≥ 21.5 – < 25 | Ref. | | Ref. | Ref. | | Ref. | | |
| ≥ 25 – < 30 | -0.049 | | -0.107 | -0.179 | | -0.189 | | |
| ≥ 30 – < 35 | -0.105 | | -0.234 | -0.247 | | -0.388 | | |
| ≥ 35 | 0.254 | | 0.211 | 0.122 | | 0.246 | | |
| **Dyslipidemia** | | | | | | | | |
| No | Ref. | | Ref. | Ref. | | Ref. | | |
| Yes | 0.021 | | 0.011 | 0.220 | | 0.164 | | |
| **Hypertension** | | | | | | | | |
| No | Ref. | | Ref. | Ref. | | Ref. | | |
| Yes | 0.151 | | 0.017 | 0.409 | | 0.251 | | |
| **Diabetes** | | | | | | | | |
| No | Ref. | | Ref. | Ref. | | Ref. | | |
| Yes | 0.559 | | 0.217 | 0.226 | | 0.013 | | |
| **Cardiovascular disease** | | | | | | | | |
| No | Ref. | | Ref. | Ref. | | Ref. | | |
| Yes | 0.580 | | 0.221 | 0.323 | | 0.036 | | |
| **History of cancer** | | | | | | | | |
| No | Ref. | | Ref. | Ref. | | Ref. | | |
| Yes | 0.890 | | 0.710 | 0.912 | | 0.791 | | |
| **Biomarkers** | | | | | | | | |

| Characteristics | Olink Target Inflammation Panel | | | | | | Olink Explore Proteomic Panel | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Men | | | Women | | | Men | | | Women | | |
| | Covariates | Biomarkers | Covariates + Biomarkers | Covariates | Biomarkers | Covariates + Biomarkers | Covariates | Biomarkers | Covariates + Biomarkers | Covariates | Biomarkers | Covariates + Biomarkers |
| *CCL20* | | 0.061 | 0.053 | | - | - | | - | - | | - | - |
| *CSF1* | | 0.053 | 0.292 | | - | - | | - | - | | - | - |
| *CXCL9* | | 0.182 | 0.175 | | - | - | | - | - | | - | - |
| *DNER* | | -0.651 | -0.373 | | - | - | | - | - | | - | - |
| *CDCP1* | | 0.256 | 0.131 | | 0.221 | 0.050 | | - | - | | - | - |
| *FGF23* | | 0.134 | 0.067 | | 0.144 | 0.152 | | - | - | | - | - |
| *HGF* | | 0.355 | 0.357 | | 0.231 | 0.199 | | - | - | | - | - |
| *IL6* | | 0.154 | 0.099 | | 0.144 | 0.140 | | 0.069 | 0.060 | | 0.087 | 0.089 |
| *KITLG* | | -0.174 | -0.264 | | -0.363 | -0.365 | | - | - | | - | - |
| *TNFRSF11B* | | 0.503 | 0.152 | | 0.339 | 0.154 | | - | - | | - | - |
| *TNFRSF9* | | 0.289 | 0.146 | | 0.446 | 0.386 | | - | - | | - | - |
| *CXCL8* | | - | - | | 0.104 | 0.094 | | - | - | | - | - |
| *FGF21* | | - | - | | 0.077 | 0.058 | | - | - | | - | - |
| *AGER* | | - | - | | - | - | | -0.473 | -0.441 | | - | - |
| *C7* | | - | - | | - | - | | 0.308 | 0.247 | | - | - |
| *EGFR* | | - | - | | - | - | | -0.539 | -0.301 | | - | - |
| *IGFBP2* | | - | - | | - | - | | 0.086 | 0.139 | | - | - |
| *MMP12* | | - | - | | - | - | | 0.290 | 0.236 | | - | - |
| *MSTN* | | - | - | | - | - | | -0.211 | -0.200 | | - | - |
| *PIGR* | | - | - | | - | - | | 0.144 | 0.175 | | - | - |
| *PLAUR* | | - | - | | - | - | | 0.107 | 0.126 | | - | - |
| *RET* | | - | - | | - | - | | -0.213 | -0.222 | | - | - |
| *SERPINA3* | | - | - | | - | - | | 0.318 | 0.341 | | - | - |
| *TNFRSF10B* | | - | - | | - | - | | 0.128 | 0.126 | | - | - |
| *TNFSF13B* | | - | - | | - | - | | 0.263 | 0.229 | | - | - |
| *AREG* | | - | - | | - | - | | 0.102 | 0.115 | | 0.215 | 0.255 |
| *CEACAM5* | | - | - | | - | - | | 0.183 | 0.181 | | 0.227 | 0.189 |
| *GDF15* | | - | - | | - | - | | 0.516 | 0.397 | | 0.166 | 0.014 |
| *NEFL* | | - | - | | - | - | | 0.266 | 0.240 | | 0.430 | 0.233 |
| *NTproBNP* | | - | - | | - | - | | 0.176 | 0.144 | | 0.110 | 0.071 |
| *WFDC2* | | - | - | | - | - | | -0.013 | -0.023 | | 0.273 | 0.312 |
| *ANGPT2* | | - | - | | - | - | | - | - | | 0.145 | 0.094 |
| *CALCA* | | - | - | | - | - | | - | - | | -0.099 | -0.085 |
| *GAL* | | - | - | | - | - | | - | - | | -0.257 | -0.245 |
| *HAVCR1* | | - | - | | - | - | | - | - | | 0.170 | 0.122 |
| *IGFBP7* | | - | - | | - | - | | - | - | | 0.328 | 0.381 |
| *IL19* | | - | - | | - | - | | - | - | | 0.138 | 0.146 |
| *ITGAV* | | - | - | | - | - | | - | - | | 0.613 | 0.718 |
| *TNR* | | - | - | | - | - | | - | - | | -0.373 | -0.286 |

Notes: The method used to calculate the risk score was based on previously published literature (Deelen et al. 2019; Fischer et al. 2014). Protein annotations and corresponding UniProt identifiers are provided in **Appendix, Table A9.**

**Table A11. Equations for inflammaging and proteomic ageing clocks**

| Subgroup | Intercept and *β*-coefficients |
|---|---|
| **Step 1: Elastic net models** [a] | |
| **To derive Raw inflammaging ages** | |
| Men | Raw inflammaging age1 = 59.915 - 0.569×*CCL20* - 1.759×*CSF1* + 1.140×*CXCL9* - 3.659×*DNER* + 1.867×*CDCP1* - 0.040×*FGF23* - 0.329×*HGF* + 0.365×*IL6* + 1.867×*KITLG* + 4.017×*TNFRSF11B* + 0.421×*TNFRSF9* |
| Women | Raw inflammaging age1 = 59.296 + 0.137×*CXCL8* - 0.038×*FGF21* + 2.267×*CDCP1* - 0.556×*FGF23* + 0.053×*HGF* + 0.173×*IL6* + 0.674×*KITLG* + 2.972×*TNFRSF11B* + 0.992×*TNFRSF9* |
| Men | Raw inflammaging age2 = 57.709 - 0.017×(Education=High) + 0.768×(Smoking=Former) + 0.790×(Smoking=Current) - 0.519×(Physical activity=Moderate) - 0.566×(Physical activity=High) + 0.133×(Alcohol consumption=Low) + 0.173×(Alcohol consumption=Medium or High) + 0.824×(BMI< 21.5 kg/m2) + 0.231×(BMI ≥ 25 – < 30 kg/m2) + 0.441×(BMI ≥ 30 – < 35 kg/m2) + 1.332×(BMI ≥ 35 kg/m2) + 1.422×(Dyslipidemia) + 1.158×Hypertension + 0.163×Diabetes + 1.364×(Cardiovascular disease) + 1.320×(History of cancer) - 0.516×*CCL20* - 1.318×*CSF1* + 0.904×*CXCL9* - 3.289×*DNER* + 1.691×*CDCP1* - 0.101×*FGF23* - 0.261×*HGF* + 0.369×*IL6* + 1.431×*KITLG* + 3.697×*TNFRSF11B* + 0.488×*TNFRSF9* |
| Women | Raw inflammaging age2 = 57.142 - 0.156×(Education=High) + 0.469×(Smoking=Former) + 1.502×(Smoking=Current) – 0.400×(Physical activity=Moderate) - 0.579×(Physical activity=High) + 0.101×(Alcohol consumption=Low) + 0.274×(Alcohol consumption=Medium or High) + 0.521×(BMI< 21.5 kg/m2) + 0.175×(BMI ≥ 25 – < 30 kg/m2) + 0.674×(BMI ≥ 30 – < 35 kg/m2) + 1.023×(BMI ≥ 35 kg/m2) + 1.713×(Dyslipidemia) + 2.010×Hypertension + 0.375×Diabetes + 1.652×(Cardiovascular disease) + 1.748×(History of cancer) + 0.112×*CXCL8* - 0.082×*FGF21* + 2.102×*CDCP1* - 0.547×*FGF23* - 0.018×*HGF* + 0.140×*IL6* + 0.424×*KITLG* + 2.737×*TNFRSF11B* + 0.949×*TNFRSF9* |
| **To derive Raw proteomic ages** | |
| Men | Raw proteomic age1 = 58.957 - 0.631×*AGER* + 0.749×C7 - 2.776×*EGFR* + 0.047×*IGFBP2* + 1.340×*MMP12* - 0.073×*MSTN* - 1.131×*PIGR* - 0.881×*PLAUR* - 1.562×*RET* - 1.630×*SERPINA3* - 0.183×*TNFRSF10B* - 0.738×*TNFSF13B* - 0.725×*AREG* - 0.064×*CEACAM5* + 2.188×*GDF15* + 0.150×*IL6* + 2.047×*NEFL* + 0.714×*NTproBNP* + 1.330×*WFDC2* |
| Women | Raw proteomic age1 = 58.768 - 0.419×*ANGPT2* + 0.070×*CALCA* - 0.653×*GAL* + 0.945×*HAVCR1* + 0.050×*IGFBP7* - 0.312×*IL19* + 4.232×*ITGAV* - 0.438×*TNR* - 0.899×*AREG* - 0.053×*CEACAM5* + 2.623×*GDF15* + 0.238×*IL6* + 2.882×*NEFL* + 0.584×*NTproBNP* + 0.371×*WFDC2* |
| Men | Raw proteomic age2 = 56.304 - 0.077×(Education=High) + 0.635×(Smoking=Former) + 2.462×(Smoking=Current) – 0.476×(Physical activity=Moderate) - 0.545×(Physical activity=High) + 0.265×(Alcohol consumption=Low) + 0.302×(Alcohol consumption=Medium or High) + 0.882×(BMI< 21.5 kg/m2) + 0.001×(BMI ≥ 25 – < 30 kg/m2) + 1.263×(BMI ≥ 30 – < 35 kg/m2) + 1.222×(BMI ≥ 35 kg/m2) + 1.138×(Dyslipidemia) + 1.082×Hypertension + 0.416×Diabetes + 1.950×(Cardiovascular disease) + 1.678×(History of cancer) - 0.596×*AGER* + 0.610×*C7* - 3.139×*EGFR* + 0.228×*IGFBP2* + 1.309×*MMP12* - 0.097×*MSTN* - 0.784×*PIGR* - 0.386×*PLAUR* - 1.491×*RET* - 1.227×*SERPINA3* - 0.133×*TNFRSF10B* - 0.691×*TNFSF13B* - 0.819×*AREG* + 0.094×*CEACAM5* + 2.031×*GDF15* + 0.119×*IL6* + 1.800×*NEFL* + 0.560×*NTproBNP* + 1.568×*WFDC2* |

| | |
|---|---|
| Women | Raw proteomic age2 = 57.061 - 0.065×(Education=High) + 0.247×(Smoking=Former) + 2.407×(Smoking=Current) – 0.241×(Physical activity=Moderate) - 0.324×(Physical activity=High) + 0.177×(Alcohol consumption=Low) + 0.221×(Alcohol consumption=Medium or High) + 0.131×(BMI< 21.5 kg/m2) + 0.631×(BMI ≥ 25 – < 30 kg/m2) + 0.665×(BMI ≥ 30 – < 35 kg/m2) + 0.859×(BMI ≥ 35 kg/m2) + 1.286×(Dyslipidemia) + 1.627×Hypertension + 0.588×Diabetes + 1.728×(Cardiovascular disease) + 1.751×(History of cancer) - 0.197×*ANGPT2* + 0.001×*CALCA* - 0.648×*GAL* + 0.711×*HAVCR1* + 0.001×*IGFBP7* - 0.348×*IL19* + 4.086×*ITGAV* - 0.588×*TNR* - 0.930×*AREG* + 0.147×*CEACAM5* + 2.371×*GDF15* + 0.138×*IL6* + 2.629×*NEFL* + 0.496×*NTproBNP* + 0.931×*WFDC2* |
| **Step 2: Linear regression models [b]** | |
| **To derive Inflammaging clocks** | |
| Men (Clock1) | Inflammaging Clock 1 = -0.306 + 1.014 × Raw inflammaging age1 |
| Women (Clock1) | Inflammaging Clock 1 = -0.249 + 1.007 × Raw inflammaging age1 |
| Men (Clock2) | Inflammaging Clock 2 = -0.263 + 1.009 × Raw inflammaging age2 |
| Women (Clock2) | Inflammaging Clock 2 = -0.353 + 1.012 × Raw inflammaging age2 |
| **To derive Proteomic ageing clocks** | |
| Men (Clock1) | Proteomic Ageing Clock 1 = -0.271 + 1.011 × Raw proteomic age1 |
| Women (Clock1) | Proteomic Ageing Clock 1 = -0.223 + 1.004 × Raw proteomic age1 |
| Men (Clock2) | Proteomic Ageing Clock 2 = -0.310 + 1.003 × Raw proteomic age2 |
| Women (Clock2) | Proteomic Ageing Clock 2 = -0.386 + 1.008 × Raw proteomic age2 |

Notes: Protein annotations and corresponding UniProt identifiers are provided in **Appendix, Table A9**.

[a] The intercept and coefficients stem from an elastic net model, in which chronological age was the dependent variable. Raw age1 was derived solely from the selected inflammatory/proteomic biomarkers in each stratum. Raw age2 was derived from both the selected inflammatory/proteomic biomarkers and traditional mortality risk factors in each stratum. The estimates stem from the training set (70% UKB).

[b] The intercept and coefficients stem from a linear regression model with chronological age as the dependent variable and the raw inflammaging/proteomic age 1 or 2 as the independent variables. The estimates stem from the training set (70% UKB).

# CURRICULUM VITAE

## PERSONAL DETAILS

| | |
|---|---|
| Name und first name: | Lei Peng |
| Date of birth: | 25. August 1996 |
| Place of birth: | Sichuan, Volksrepublik China |
| Nationality: | Chinese |
| Marital status: | Single |
| ORCID iD: | 0000-0001-8754-020X |

## EDUCATION

| | |
|---|---|
| 09.2023 – Present | PhD Candidate in Clinical Epidemiology, Heidelberg University, Heidelberg, Germany |
| 09.2020 – 06.2023 | Master of Science in Epidemiology and Biostatistics, Peking University, Beijing, China |
| 09.2014 – 06.2019 | Bachelor of Science in Public Health and Preventive Medicine, Southwest Medical University, Luzhou, China |

## PROFESSIONAL / RESEARCH CAREER

| | |
|---|---|
| 07.2025 – Present | Doctoral Researcher, Clinical Epidemiology of Early Cancer Detection, German Cancer Research Center, Heidelberg, Germany |
| 09.2023 – 06.2025 | Doctoral Researcher, Division of Clinical Epidemiology and Aging Research, German Cancer Research Center, Heidelberg, Germany |
| 09.2023 – Present | Member of Junior Research Group of Ben Schöttker, Network Aging Research, Heidelberg University, Heidelberg, Germany |
| 09.2020 – 07.2023 | Master's Student, Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education/Beijing), Department of Cancer Epidemiology, Peking University Cancer Hospital & Institute, Beijing, China |
| 07.2019 – 08.2020 | Biostatistician, Lipu Data Analysis Studio, Luzhou, China |
| 12.2017 – 05.2018 | Public Health Intern, Mianyang Center for Disease Control and Prevention, Mianyang, China |
| 05.2017 – 11.2017 | Clinical Intern, The Affiliated Hospital of Southwest Medical University, Luzhou, China |

# ACKNOWLEDGEMENTS

The completion of this dissertation would not have been possible without the support, guidance, and encouragement of many individuals and institutions, to whom I would like to express my sincere gratitude.

First and foremost, I am deeply grateful to my supervisor, Prof. Dr. Ben Schöttker, for offering me the opportunity to pursue my doctoral studies at the German Cancer Research Center (DKFZ) and Heidelberg University. His scientific guidance, continuous support, and high academic standards have profoundly influenced my development as a researcher. I am especially thankful for the trust and academic freedom he gave me to develop my own research ideas and complete this dissertation.

I also sincerely thank Prof. Dr. Hermann Brenner for his valuable scientific advice and for creating an outstanding research environment in clinical epidemiology and ageing research. My sincere appreciation also goes to my co-authors, collaborators and colleagues, especially Dr. Sigrid V. Carlsson, Prof. Dr. Klaus Hauer, Dr. Silvia Calderazzo, Prof. Dr. Bernd Holleczek, Dr. Ruijie Xie, Dr. Tomislav Vlaski, Dr. Durgesh Wankhede, Dr. Megha Bhardwaj, and Ms. Miriam Degen, for their collaboration, insightful discussions, and valuable contributions. I would particularly like to thank Dr. Silvia Calderazzo for statistical support, Dr. Utz Benscheid for preparing and managing the ESTHER database, Dr. Romana Kimmel for laboratory assistance, and Dr. Durgesh Wankhede for his contribution to the development of the biological ageing clock web calculators.

This dissertation was made possible through the ESTHER study and the UK Biobank. I sincerely thank all study participants, general practitioners, study staff, laboratory personnel, and research coordinators whose dedication made these invaluable cohort resources available.

Financial support is gratefully acknowledged. I sincerely thank the China Scholarship Council (CSC, No. 202306010093) for supporting my doctoral studies in Germany. I also acknowledge the organizations supporting the ESTHER study and the UK Biobank, whose long-term commitment to population-based research has laid the foundation for studies such as this dissertation.

Finally, I would like to express my deepest gratitude to my family for their unconditional love, encouragement, and unwavering support throughout this journey. Their confidence in me has been my greatest source of strength.

To everyone who has supported me, directly or indirectly, I offer my heartfelt thanks.

# EIDESSTATTLICHE VERSICHERUNG

1. Bei der eingereichten Dissertation zu dem Thema
   *„Associations of Metabolomic and Proteomic Profiles with Frailty and Mortality in Population-Based Cohorts"*

   handelt es sich um meine eigenständig erbrachte Leistung.

2. Ich habe nur die angegebenen Quellen und Hilfsmittel benutzt und mich keiner unzulässigen Hilfe Dritter bedient. Insbesondere habe ich wörtlich oder sinngemäß aus anderen Werken übernommene Inhalte als solche kenntlich gemacht.

3. Die Arbeit oder Teile davon habe ich bislang nicht an einer Hochschule des In- oder Auslands als Bestandteil einer Prüfungs- oder Qualifikationsleistung vorgelegt.

4. Die Richtigkeit der vorstehenden Erklärungen bestätige ich.

5. Die Bedeutung der eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unrichtigen oder unvollständigen eidesstattlichen Versicherung sind mir bekannt. Ich versichere an Eides statt, dass ich nach bestem Wissen die reine Wahrheit erklärt und nichts verschwiegen habe.

Ort und Datum Unterschrift

# ANGABEN ZU VERWENDETEN KI-BASIERTEN ELEKTRONISCHEN HILFSMITTEL

Zur Dokumentation der verwendeten Hilfsmittel ist der schriftlichen Ausarbeitung ein besonderer Anhang hinzugefügt, der eine Liste und Beschreibung aller verwendeter KI-basierter Hilfsmittel enthält. Der besondere Anhang zur Dokumentation der verwendeten Hilfsmittel enthält folgende Angaben:

1. Ziel der Verwendung: ChatGPT (OpenAI) wurde zur sprachlichen und stilistischen Überarbeitung englischsprachiger Textpassagen sowie zur Verbesserung von Grammatik, Ausdruck und Lesbarkeit verwendet.

2. Verwendungsweise: ChatGPT wurde zur sprachlichen Überarbeitung selbstständig verfasster Textpassagen eingesetzt. Die wissenschaftlichen Inhalte, statistischen Analysen, Interpretation der Ergebnisse und Schlussfolgerungen wurden eigenständig erarbeitet. Alle überarbeiteten Texte wurden vom Verfasser geprüft und eigenverantwortlich in die Dissertation übernommen.

3. Betroffene Kapitel und Abschnitte: ChatGPT wurde zur sprachlichen und stilistischen Überarbeitung in den Abschnitten 1.1–1.8 und 4.5–4.8 verwendet.

Der Gebrauch dieser Hilfsmittel inklusive Art, Ziel und Umfang des Gebrauchs wurde mit meinem offiziellen Betreuer Prof. Dr. Ben Schöttker abgesprochen.

Mir ist bewusst, dass insbesondere der Versuch einer nicht dokumentierten Nutzung KI-basierter Hilfsmittel als Täuschungsversuch zu werten ist:

Gem. § 16 Abs. 2 der Promotionsordnung „Dr. med. / dent.“:

„Ergibt sich vor Aushändigung der Promotionsurkunde, dass der Kandidat/die Kandidatin bei einer Promotionsleistung getäuscht hat, so können einzelne oder alle Promotionsleistungen für ungültig erklärt werden. In schweren Fällen kann die Zulassung zum Promotionsverfahren zurückgenommen werden.“

Und § 16 Abs. 2 der Promotionsordnung „Dr. sc. hum.“:

„Ergibt sich vor Aushändigung der Promotionsurkunde, dass der Doktorand / Doktorandin bei einer Promotionsleistung getäuscht hat, so kann der Promotionsausschuss diese Promotionsleistung oder alle bisher erbrachten Promotionsleistungen für ungültig erklären. In besonders schweren Fällen kann der Promotionsausschuss die Annahme als Doktorand / Doktorandin endgültig widerrufen.“

Ort und Datum Unterschrift