

<https://doi.org/10.1038/s41746-025-02075-2>

Metabolomic characterization of frailty identifies subtype-specific management strategies

Check for updates

Lushan Xiao^{1,2,5}, Qijie Deng^{3,5}, Jiaren Wang^{1,2,5}, Shengxing Liang^{3,5}, Ruining Li^{1,2,5}, Hao Cui^{1,2}, Yan Li^{1,2}, Pu Jiang^{1,2}, Rongfeng Zhang³, Lin Zeng^{1,2}✉, Chang Hong^{1,2}✉, Weinan Lai⁴✉ & Li Liu^{1,2}✉

The clinical application of the frailty phenotype and frailty index still has some limitations, and whether the classification of frailty based on metabolites is beneficial to the management of the frailty population remains unclear. This study analyzed 160,407 UK Biobank participants to define frailty subtypes using metabolic profiles. Based on 251 biomarkers, machine learning identified 11 key metabolites, leading to four novel frailty subtypes. Subtypes III and IV, characterized by adverse metabolic features such as high GlycA and low LAFA, were designated as high-risk groups. These subtypes showed significantly increased risks for 13 chronic diseases and all-cause mortality compared to lower-risk subtypes. Adherence to a healthy diet was associated with risk reduction in the high-risk groups. These findings demonstrate the heterogeneity of frailty and suggest that metabolite-based subtyping could improve prognostic precision and guide targeted dietary interventions in clinical practice.

Frailty affects approximately 24% of the global population aged 50 years and older^{1,2}. It is a state of decreased physiological reserve stemming from dysregulated multisystem processes. Core mechanisms include chronic inflammation (inflammaging), characterized by elevated cytokines like IL-6, which drive muscle catabolism and endocrine dysfunction. This is compounded by mitochondrial dysfunction, cellular senescence, and disrupted anabolic signaling (e.g., GH/IGF-1 axis), culminating in sarcopenia and impaired stress resistance^{3,4}. Frailty is a risk factor for chronic diseases, such as cardiovascular diseases (CVDs), cerebrovascular diseases, chronic obstructive pulmonary disease (COPD), kidney diseases, liver cancer, and colorectal cancer^{5,6}. Therefore, many clinical guidelines advocate for routine monitoring of frailty^{7,8}. There are two classic models for defining frailty: the frailty phenotype and frailty index (FI). The frailty phenotype is characterized by unintentional weight loss, reduced grip strength, slow walking speed, poor endurance, low energy consumption, and reduced physical activity. In contrast, the FI reflects cumulative health deficits associated with age. Studies comparing these models suggest that the FI more accurately predicts the risk of adverse events, such as falls and hospitalization, than the frailty phenotype^{9,10}.

However, the measures used to calculate the FI are not easily obtained. Therefore, more accurate and convenient methods are needed to measure frailty and identify individuals at a greater risk of frailty-related diseases.

Frailty is linked to disturbances in amino acid and fatty acid metabolism¹¹. Although studies have shown that metabolic disorders in patients with frailty impact the prognosis of chronic diseases, such as type 2 diabetes, CVDs, and metabolic-associated steatotic liver disease (MASLD)^{12–14}, there is no consensus on the management and application of blood metabolites in this population. Consequently, it is difficult for clinicians to optimize patient management by controlling metabolites at an early stage. Diet plays a crucial role in metabolism, and high-fat, high-sugar diets significantly increase the risk of metabolic-related diseases, including diabetes, MASLD, and CVDs^{15–18}. Diet quality also affects the progression of frailty^{19–21}. A healthy diet can improve the prognosis of various chronic diseases, such as upper gastrointestinal cancer, kidney disease, COPD, and colorectal cancer^{22–25}. However, it remains unclear whether a healthy diet can modulate metabolic status and improve the prognosis of frail populations at higher risk of chronic diseases.

¹Department of Health Management, Nanfang Hospital, Southern Medical University, Guangzhou, China. ²Guangdong Provincial Key Laboratory for Prevention and Control of Major Liver Diseases, Department of Infectious Diseases, Nanfang Hospital, Southern Medical University, Guangzhou, China. ³Nanfang Hospital, Southern Medical University, Guangzhou, China. ⁴Department of Rheumatology and Immunology, Nanfang Hospital, Southern Medical University, Guangzhou, China. ⁵These authors contributed equally: Lushan Xiao, Qijie Deng, Jiaren Wang, Shengxing Liang, Ruining Li. ✉e-mail: lin_zeng1126@163.com; shanoicy@163.com; Laiwn123@smu.edu.cn; liuli@i.smu.edu.cn

This study aimed to explore novel frailty subtypes based on metabolite profiles, identify frail groups at higher risk of chronic diseases, and investigate whether a healthy diet can improve the prognosis of these groups.

Results

Baseline characteristics of the study population

From the UK Biobank cohort of 502,401 individuals, 160,407 individuals were included in the analysis after applying exclusion criteria. Of the included participants, 28,196 individuals were classified as frail.

The demographic and clinical characteristics of the participants are summarized in Table 1. Frail participants were predominantly older, more likely to be women, and less often White. They also had higher rates of smoking, lower alcohol consumption, and elevated average systolic pressure (SBP), average diastolic pressure (DBP), body mass index (BMI), alanine transaminase (ALT), aspartate transaminase (AST), Alkaline Phosphatase (ALP), triglyceride, and HbA1c levels.

Identification of key frailty features and novel subtypes

To identify metabolic features associated with frailty, standard Z-score normalization was applied to 251 metabolic data points. The CatBoost algorithm was used to analyze these features. The SHAP explainer was applied to the best-trained model, calculating SHAP values to rank the top 20 most influential features (Fig. 1a, Supplementary Fig. 2). GlycA (glycoprotein acetylation) was the most important metabolite, followed by LA/FA (percentage of linoleic acid to total fatty acids). The top 11 metabolites with the highest SHAP values were selected for further analysis using the XGBoost model.

Following hyperparameter tuning and five-fold cross-validation, the XGBoost model achieved an AUC of 0.785 in the train set and 0.757 in the test set, indicating strong predictive performance (Fig. 1b). PCA was performed on the 11 selected metabolic features, with PC1 and PC2 accounting for 44.54% and 16.63% of the variance, respectively.

Using PC1 and PC2, K-means clustering analysis was performed. The optimal number of clusters was determined to be four based on elbow and silhouette coefficient plots (Fig. 1c, Supplementary Fig. 1). The resulting NMI of 0.864 and ARI of 0.881 demonstrate that the clustering is highly stable. Frail participants were subsequently stratified into four novel subtypes for further analysis.

Baseline characteristics and metabolic profiles of novel frailty subtypes

Table 2 summarizes the demographic, anthropometric, and clinical characteristics of the four identified frailty subtypes. Participants in subtypes III and IV were older, predominantly men, and more likely to be White. These subtypes were also associated with lower SBP and higher BMI levels. Distinct patterns among the subtypes emerged based on standardized cluster variables (Fig. 2, Supplementary Fig. 5, Supplementary Table 11).

1. Subtype I: Comprising 8409 participants, this group exhibited extremely low levels of GlycA and Val (valine), the highest ratios of LA/FA, and elevated proportions of docosahexaenoic acid to total fatty acids (DHA/FA). They also had higher ratios of polyunsaturated to mono-unsaturated fatty acids (PUFA/MUFA). Subtype I was named low GlycA and Val-related frailty (LGVF).
2. Subtype II: This group, with 6971 participants, was characterized by the highest concentrations of Alb (albumin) and LA, and was termed as high Alb and LA-related frailty (HALF).
3. Subtype III: Including 7477 participants, this subtype demonstrated the lowest levels of Alb and LA and was named low Alb and LA-related frailty (LALF).
4. Subtype IV: With 5339 participants, this group displayed high levels of GlycA, elevated Val concentrations, higher proportions of mono-unsaturated fatty acids to total fatty acids (MUFA/FA), and notably low LA/FA ratios. It was named as high GlycA and Val-related frailty (HGVF).

The metabolic profile of the non-frail participants closely resembled that of subtype I.

Associations between novel frailty subtypes and chronic diseases

With a median follow-up time of 13.8 years, Kaplan–Meier curves illustrated differences in the cumulative incidence rates of 13 chronic diseases and all-cause mortality among the four novel frailty subtypes and non-frail participants (Supplementary Fig. 3). The highest cumulative incidence rates were observed for coronary artery disease,

Table 1 | Baseline characteristics of the study population

Characteristic	Overall N = 160,407	Non-frail N = 132,211	Frail N = 28,196	P-value
Age (years)	56.10 (8.09)	55.69 (8.12)	58.06 (7.69)	<0.001
Male	74,996.00 (46.75%)	63,116.00 (47.74%)	11,880.00 (42.13%)	<0.001
White	152,376.00 (94.99%)	126,032.00 (95.33%)	26,344.00 (93.43%)	<0.001
Current smoking	16,012.00 (9.98%)	11,022.00 (8.34%)	4990.00 (17.70%)	<0.001
Current drinking	130,937.00 (81.63%)	112,570.00 (85.14%)	18,367.00 (65.14%)	<0.001
BMI (Kg/m ²)	27.14 (4.70)	26.50 (4.13)	30.11 (5.91)	<0.001
SBP (mmHg)	137.46 (18.58)	137.25 (18.59)	138.43 (18.47)	<0.001
DBP (mmHg)	82.00 (10.12)	81.93 (10.05)	82.37 (10.43)	<0.001
ALT (U/L)	23.23 (13.96)	22.67 (13.27)	25.86 (16.61)	<0.001
AST (U/L)	26.09 (10.57)	25.78 (9.90)	27.54 (13.18)	<0.001
ALP (U/L)	82.70 (25.60)	80.96 (24.21)	90.87 (29.97)	<0.001
Cholesterol (mmol/L)	5.73 (1.13)	5.79 (1.09)	5.41 (1.24)	<0.001
Direct bilirubin (μmol/L)	1.74 (0.82)	1.76 (0.80)	1.69 (0.88)	<0.001
Total bilirubin (μmol/L)	9.19 (4.47)	9.36 (4.53)	8.39 (4.13)	<0.001
Triglycerides (mmol/L)	1.72 (1.02)	1.65 (0.97)	2.04 (1.17)	<0.001
HbA1c (mmol/mol)	35.80 (6.56)	35.08 (5.17)	39.20 (10.27)	<0.001

Means (SD) are reported for continuous variables, while count (%) is provided for categorical variables. Continuous variables were compared using the Wilcoxon rank-sum test, and categorical variables were assessed using Pearson's chi-squared test. Statistical significance was determined using P-values when comparing non-frail and frail populations.

BMI body mass index, SBP average systolic pressure, DBP average diastolic pressure, ALT alanine transaminase, AST aspartate transaminase, ALP alkaline phosphatase

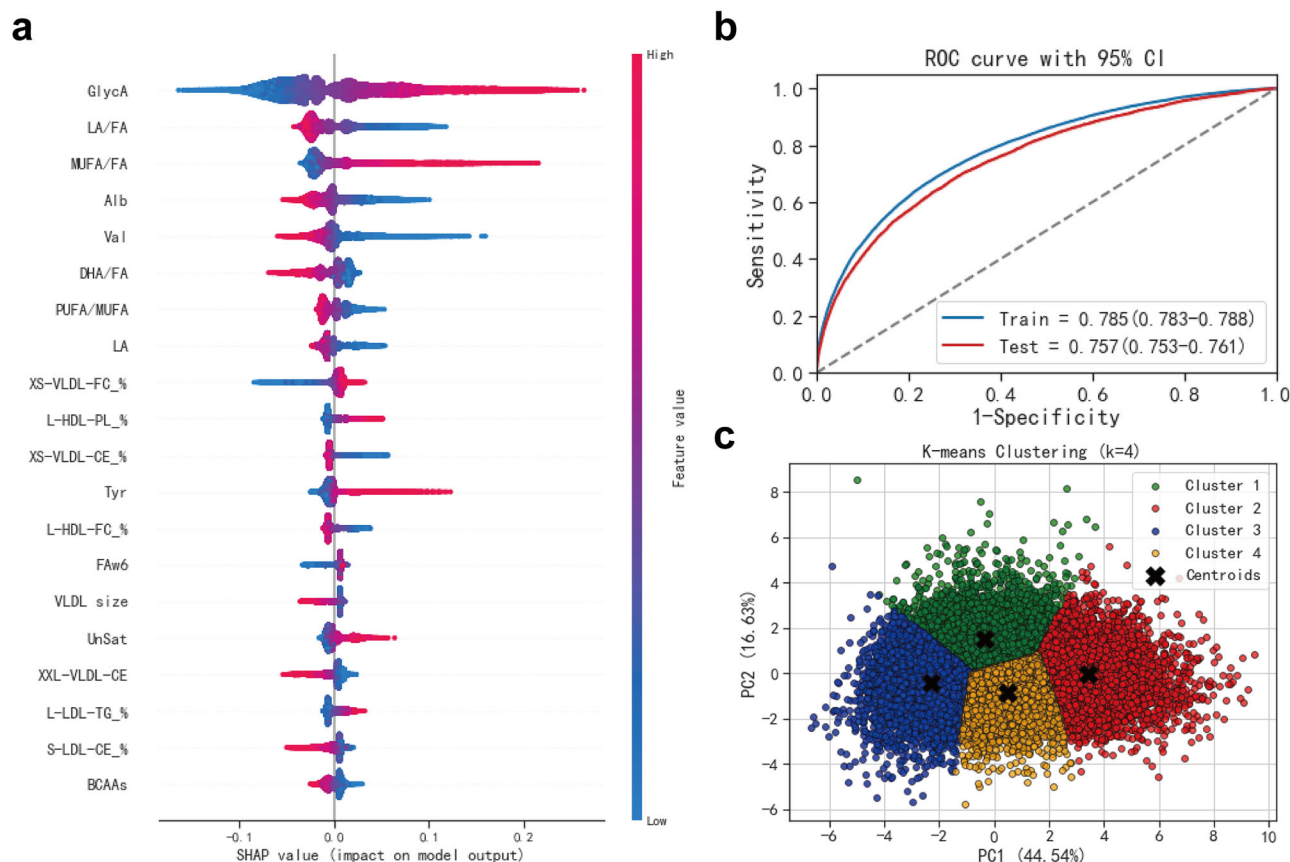


Fig. 1 | Key frailty features and novel subtypes identified using machine learning. **a** SHAP values indicate the impact of each feature in the CatBoost model. The horizontal axis represents SHAP values, reflecting feature influence; in contrast, the vertical axis ranks features by importance. **b** ROC curve of the XGBoost model in the train and test set, demonstrating predictive performance based on the top 11 key

features. The AUC value reflects prediction accuracy. **c** PCA visualization of frail participants, with clustering results from K-means analysis plotted on a two-dimensional plane. PC1 and PC2 represent the first and second principal components, respectively. ROC receiver operating characteristic, AUC area under the curve, PCA principal component analysis, PC principal component.

type 2 diabetes, COPD, and all-cause mortality. Subtypes I and II displayed cumulative incidence trajectories similar to each other, as did subtypes III and IV. Non-frail participants exhibited the lowest cumulative incidence rates across all outcomes; in contrast, subtypes III and IV had the highest cumulative risks of chronic diseases.

A multivariate Cox proportional hazards model was employed to examine the association between novel frailty subtypes and the outcomes. After adjusting for age, sex, ethnicity, current smoking and drinking, SBP, and DBP, most of the results remained significant (Fig. 3, Supplementary Tables 5, 6). Across all outcomes, the non-frail group consistently had the lowest risk.

Compared to participants in Subtype I, individuals in subtypes III and IV had considerably elevated risk for all outcomes, including:

1. Coronary artery disease (subtype III: HR [95% CI] 1.13 [1.04, 1.24]; subtype IV: HR [95% CI] 1.15 [1.04, 1.26])
2. Heart failure (subtype III: HR [95% CI] 1.22 [1.09, 1.36]; subtype IV: HR [95% CI] 1.18 [1.05, 1.33])
3. MACE (subtype III: HR [95% CI] 1.19 [1.04, 1.35]; subtype IV: HR [95% CI] 1.29 [1.12, 1.48])
4. MI (subtype III: HR [95% CI] 1.35 [1.21, 1.51]; subtype IV: HR [95% CI] 1.33 [1.18, 1.50])
5. Type 2 diabetes (subtype III: HR [95% CI] 2.24 [2.01, 2.49]; subtype IV: 3.00 [2.69, 3.34])
6. MASLD (subtype III: HR [95% CI] 1.42 [1.20, 1.68]; subtype IV: HR [95% CI] 1.85 [1.56, 2.20])
7. COPD (subtype III: HR [95% CI] 1.11 [1.02, 1.21]; subtype IV: HR [95% CI] 1.15 [1.05, 1.26])

8. SLD (subtype III: HR [95% CI] 1.23 [0.94, 1.60]; subtype IV: HR [95% CI] 1.45 [1.10, 1.91])
9. PAD (subtype III: HR [95% CI] 1.23 [1.09, 1.39]; subtype IV: HR [95% CI] 1.21 [1.06, 1.38])
10. ESRD (subtype III: HR [95% CI] 1.83 [1.31, 2.56]; subtype IV: HR [95% CI] 2.49 [1.77, 3.49])
11. Kidney cancer (subtype III: HR [95% CI] 1.05 [0.69, 1.60]; subtype IV: HR [95% CI] 1.16 [0.74, 1.81])
12. Lung cancer (subtype III: HR [95% CI] 1.25 [1.03, 1.53]; subtype IV: HR [95% CI] 1.38 [1.11, 1.72])
13. AAA (subtype III: HR [95% CI] 1.16 [0.89, 1.50]; subtype IV: HR [95% CI] 1.29 [0.97, 1.70])
14. all-cause mortality (subtype III: HR [95% CI] 1.19 [1.10, 1.29]; subtype IV: HR [95% CI] 1.25 [1.14, 1.36])

Associations between healthy diet and chronic diseases in the high-risk frailty group

Based on the results of the previous analysis, we found that subtypes I and II had a lower risk of chronic diseases than subtypes III and IV. Therefore, we categorized subtypes I and II as the low-risk frailty group and subtypes III and IV as the high-risk frailty group. Statistical analysis revealed significant differences in prognosis between these two groups, with the high-risk frailty group exhibiting poorer outcomes (Fig. 4, Supplementary Tables 7, 8).

We then explored the association between healthy diet and prognosis in the high-risk frailty group to determine whether a healthy diet could

Table 2 | Baseline characteristics of the novel frailty subtypes

Characteristic	Overall N = 160,407	Non-frail N = 132,211	Subtype IV = 8409	Subtype IIIM = 6971	Subtype IIIIN = 7477	Subtype IVW = 5339	P-value
Age	56.10 (8.09)	55.69 (8.12)	57.57 (7.98)	56.84 (7.66)	59.62 (7.27)	58.25 (7.44)	<0.001
Male	74,996 (46.75%)	63,116 (47.74%)	2673 (31.79%)	2255 (32.35%)	3787 (50.65%)	3165 (59.28%)	<0.001
White	152,376 (94.99%)	126,032 (95.33%)	7587 (90.22%)	6503 (93.29%)	7092 (94.85%)	5162 (96.68%)	<0.001
Current smoking	16,012 (9.98%)	11,022 (8.34%)	1186 (14.10%)	1357 (19.47%)	1318 (17.63%)	1129 (21.15%)	<0.001
Current drinking	130,937 (81.63%)	112,570 (85.14%)	5639 (67.06%)	4388 (62.95%)	4941 (66.08%)	3399 (63.66%)	<0.001
BMI	27.14 (4.70)	26.50 (4.13)	27.62 (5.21)	30.15 (5.67)	31.12 (5.91)	32.56 (5.78)	<0.001
SBP	137.46 (18.58)	137.25 (18.59)	136.02 (19.02)	139.51 (18.79)	138.71 (17.90)	140.45 (17.51)	<0.001
DBP	82.00 (10.12)	81.93 (10.05)	80.86 (10.44)	84.20 (10.34)	81.72 (10.31)	83.29 (10.25)	<0.001
GlycA	0.80 (0.12)	0.79 (0.11)	0.77 (0.09)	0.93 (0.11)	0.84 (0.10)	0.96 (0.12)	<0.001
LA/FA	29.12 (3.41)	29.48 (3.23)	29.95 (2.72)	29.33 (2.68)	25.96 (2.28)	23.04 (2.60)	<0.001
MUFA/FA	23.60 (2.67)	23.31 (2.53)	21.97 (1.43)	25.27 (1.61)	25.10 (1.55)	29.00 (1.79)	<0.001
Alb	39.46 (3.37)	39.63 (3.32)	38.42 (3.32)	39.73 (3.58)	37.76 (3.46)	38.87 (3.46)	<0.001
Val	0.21 (0.04)	0.21 (0.04)	0.19 (0.04)	0.22 (0.04)	0.21 (0.04)	0.25 (0.05)	<0.001
DHA/FA	2.01 (0.68)	2.04 (0.68)	2.25 (0.68)	1.61 (0.43)	1.92 (0.57)	1.38 (0.49)	<0.001
PUFA/MUFA	1.83 (0.34)	1.87 (0.33)	2.05 (0.22)	1.62 (0.18)	1.62 (0.18)	1.22 (0.16)	<0.001
LA	3.48 (0.69)	3.51 (0.67)	3.20 (0.57)	4.09 (0.62)	2.86 (0.47)	3.32 (0.69)	<0.001
XS-VLDL-FC_%	16.05 (0.60)	16.10 (0.56)	16.26 (0.42)	16.18 (0.44)	15.71 (0.47)	14.89 (0.72)	<0.001
L-HDL-PL_%	50.41 (2.96)	50.07 (2.65)	49.41 (1.79)	50.28 (2.46)	53.25 (2.66)	56.48 (3.82)	<0.001
XS-VLDL-CE_%	35.39 (4.18)	35.91 (3.83)	37.03 (2.78)	34.85 (2.84)	31.18 (2.79)	26.61 (3.74)	<0.001

Means (SD) are reported for continuous variables; in contrast, count (%) is presented for categorical features. Continuous variables were compared using the Kruskal–Wallis rank-sum test, and categorical variables were analyzed with Pearson's chi-squared test. P-values were used to determine the statistical significance of differences between the non-frail participants and frailty subtypes. L-HDL-PL_% percentage of phospholipids in large high-density lipoprotein particles to total lipids, XS-VLDL-CE_% percentage of cholesterol esters in extremely small very low-density lipoprotein particles to total lipids, XS-VLDL-FC_% percentage of free cholesterol in extremely small very low-density lipoprotein particles to total lipids.

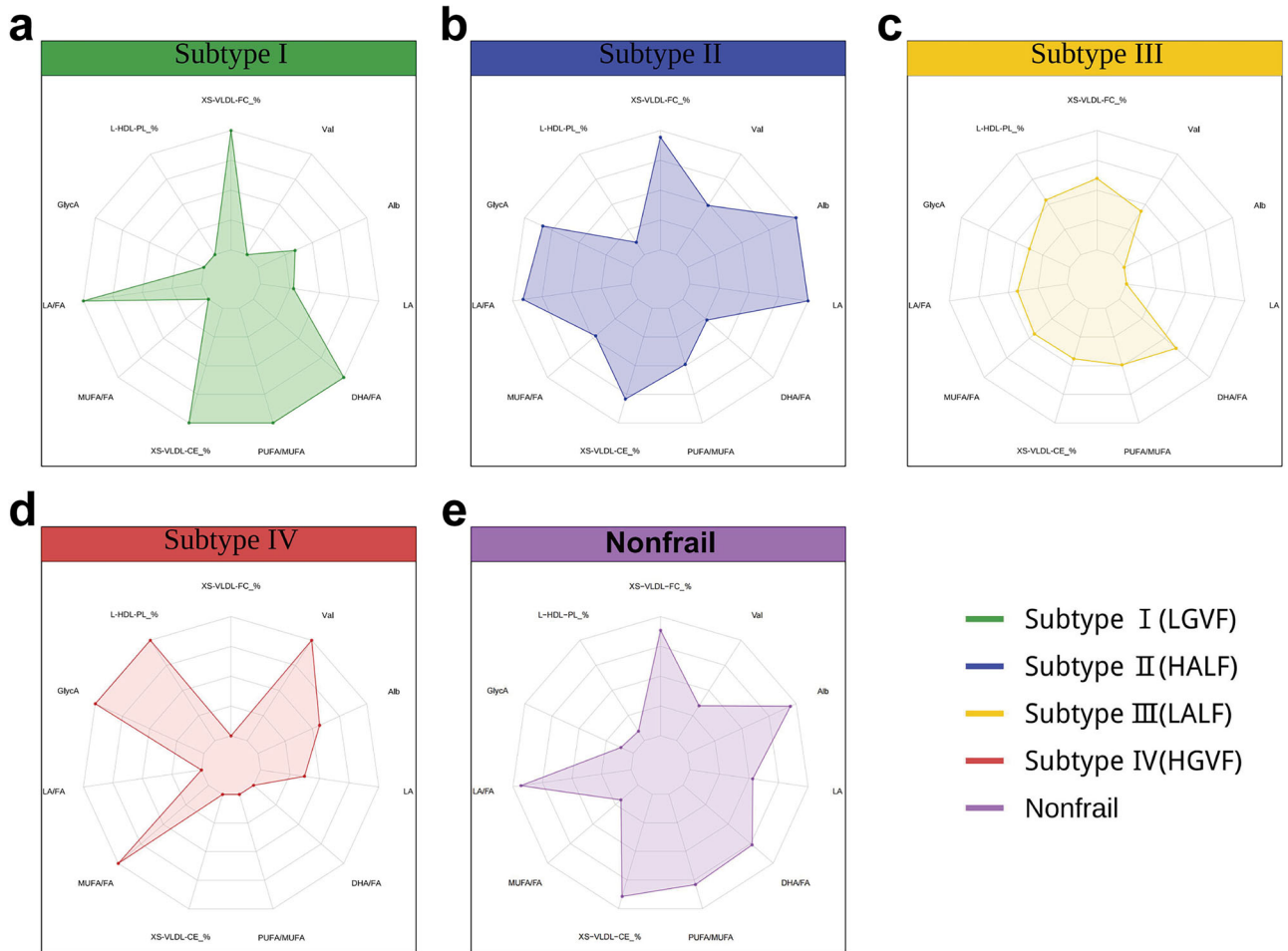
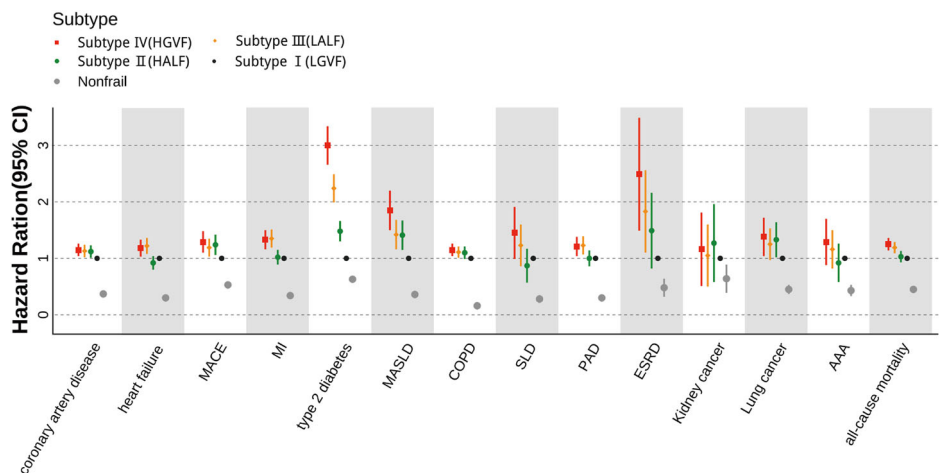


Fig. 2 | Radar chart illustrating the characteristics of the novel frailty subtypes. Cluster means for each variable were standardized by calculating cohort means and standard deviations and converting the cluster means to z-scores. This standardization eliminated numerical differences between variables, facilitating comparisons.

Radar charts display z-scores for each subtype, highlighting distinctive patterns across variables. **a** Subtype I, low GlycA and Val related frailty (LGVF); **b** Subtype II, high Alb and LA related frailty (HALF); **c** Subtype III, low Alb and LA related frailty (LALF); **d** Subtype IV, high GlycA and Val related frailty (HGVF); **e** Non-frail.

Fig. 3 | Comparisons of disease risks by novel frailty subtypes using multivariate Cox proportional hazards models, adjusted for age, sex, ethnicity, current smoking, current drinking, SBP, and DBP. Significance level: $P < 0.05$.



improve their outcomes. In the high-risk frailty group, after adjusting for confounding factors, participants with characteristics of a healthy diet had reduced risks of outcomes compared with those with an unhealthy diet (Fig. 5, Supplementary Table 9). HRs and 95% CIs are as follows:

1. Coronary artery disease (HR [95% CI]: 0.83 [0.73, 0.94])

2. Heart failure (HR [95% CI]: 0.86 [0.74, 0.99])

3. MACE (HR [95% CI]: 0.98 [0.83, 1.15])

4. MI (HR [95% CI]: 0.98 [0.86, 1.12])

5. MASLD (HR [95% CI]: 0.93 [0.77, 1.13])

6. COPD (HR [95% CI]: 0.81 [0.72, 0.91])

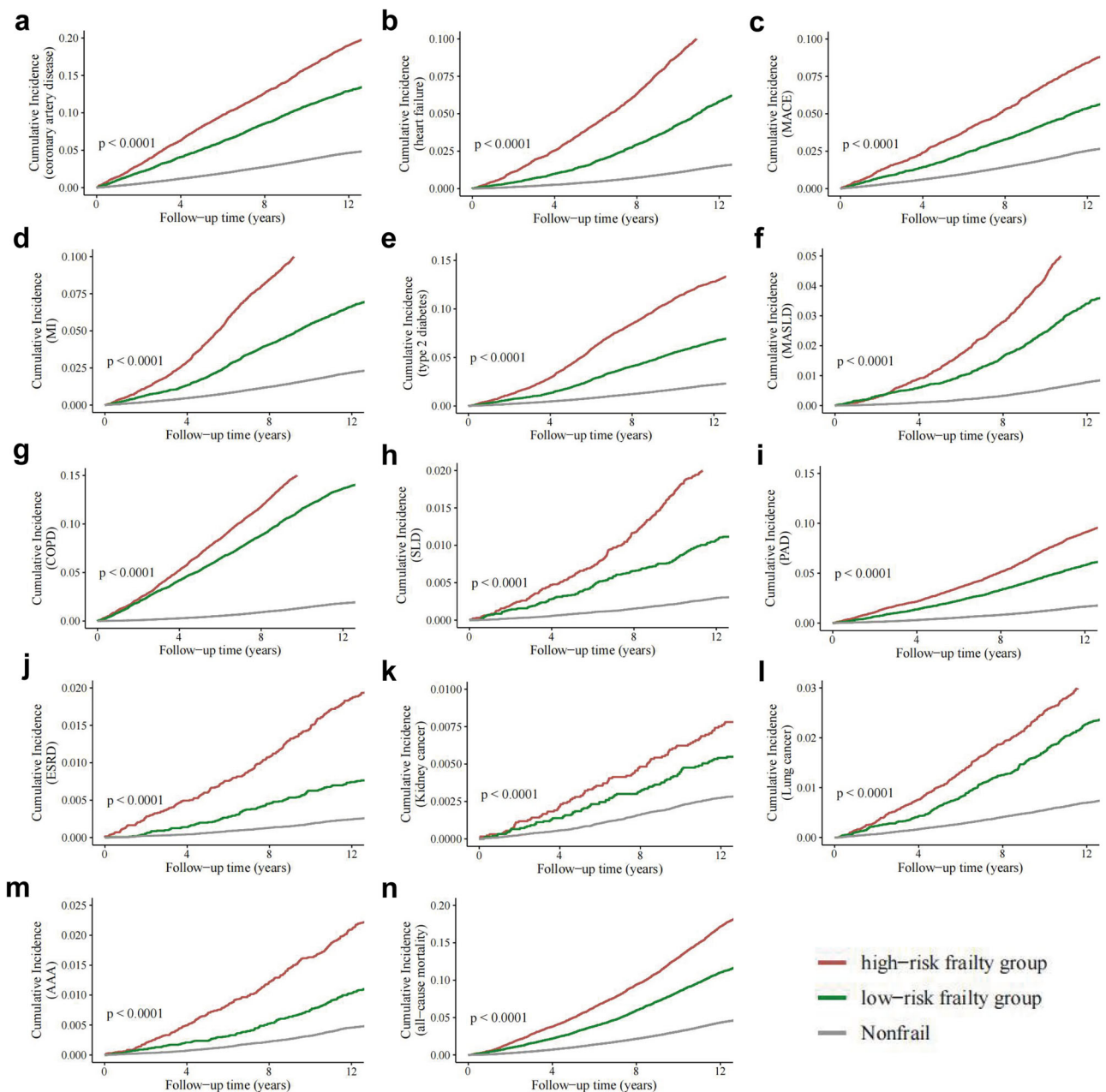


Fig. 4 | Cumulative incidence rates of outcomes stratified by low-risk frailty group and high-risk frailty group. The Kaplan–Meier survival curve was employed to estimate the cumulative density of events at specific time points. The Log-rank test was used to evaluate statistical differences between survival curves. **a** Coronary artery disease. **b** Heart failure. **c** MACE. **d** MI. **e** Type 2 diabetes. **f** MASLD **g** COPD. **h** SLD. **i** PAD. **j** ESRD. **k** Kidney cancer. **l** Lung Cancer. **m** AAA. **n** All-cause mortality.

Significance level: $P < 0.05$. MACE major adverse cardiovascular events, MI myocardial infarction, MASLD metabolic dysfunction-associated steatotic liver disease, COPD chronic obstructive pulmonary disease, SLD severe liver disease, PAD peripheral artery disease, ESRD end-stage renal disease, AAA abdominal aortic aneurysm.

7. SLD (HR [95% CI]: 0.82 [0.58, 1.14])
8. PAD (HR [95% CI]: 0.83 [0.71, 0.98])
9. ESRD (HR [95% CI]: 0.61 [0.41, 0.91])
10. Kidney cancer (HR [95% CI]: 0.96[0.56, 1.65])
11. Lung cancer (HR [95% CI]: 0.71[0.53, 0.96])
12. AAA (HR [95% CI]: 0.84 [0.60, 1.18])
13. All-cause mortality (HR [95% CI]: 0.87 [0.78, 0.97])

Healthy diet was associated with an increased risk of type 2 diabetes in the high-risk frailty group, but this association was not significant ($P > 0.05$).

Subgroup analysis

Furthermore, subgroup analyses were performed based on age and sex, adjusting for factors such as age, sex, ethnicity, current smoking status, current alcohol consumption status, SBP, and DBP. The associations between subtypes III and IV and the risks of coronary artery disease, heart failure, MACE, MI, type 2 diabetes, MASLD, COPD, SLD, and all-cause mortality were more pronounced in women under 60 years of age than in other populations. In contrast, the associations between subtypes III and IV and the risk of PAD were more pronounced in individuals aged 60 and older. The association between subtypes III and IV and the risk of lung

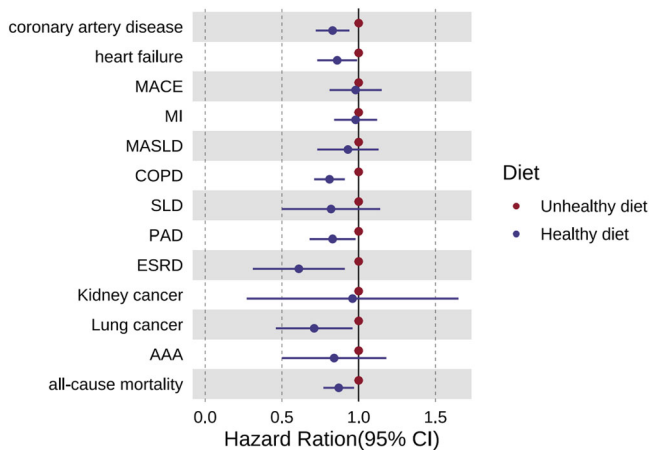


Fig. 5 | The impact of a healthy diet on the risk of chronic diseases and all-cause mortality in the high-risk frailty group. HR values were obtained from multivariate Cox proportional hazards regression, adjusted for age, sex, ethnicity, current smoking, current drinking, SBP, and DBP. Significance level: $P < 0.05$.

cancer was also more pronounced in individuals under 60 years of age. No interactions were observed for kidney cancer or AAA (Supplementary Table 10).

Discussion

This study identified 11 metabolic features associated with frailty from approximately 160,000 individuals in the UK Biobank using 251 NMR biomarkers and classified four novel frailty subtypes through cluster analysis. Each subtype displayed distinct metabolic and clinical characteristics. Subtypes I and II had a lower risk of chronic diseases compared to subtypes III and IV. Consequently, subtypes I and II were grouped into the low-risk frailty group, while subtypes III and IV were grouped into the high-risk frailty group. Furthermore, our findings demonstrated that high adherence to a healthy diet significantly reduced the risk of chronic diseases in the high-risk frailty group, offering insights to inform personalized clinical decision-making.

Previous studies have applied various methods to categorize frail populations. For instance, research involving approximately 6000 participants identified clustering patterns of multidimensional health issues in older adults, resulting in four subtypes of geriatric frailty²⁶. Another study, conducted in an Asian cohort, explored the relationship between frailty and intrinsic capacity, identifying subgroups with distinct outcomes over a year²⁷. Linzy et al. used a data-driven approach to identify three frailty subtypes—NCF, MTF, and RTF—each exhibiting varying degrees and rates of neurocognitive decline, with MTF showing the steepest trajectory²⁸. Similarly, Okoye et al. identified four clusters of patients with heart failure based on frailty, comorbidities, and B-type natriuretic peptide levels²⁹. In contrast to these studies, our research is the first to conduct a cluster analysis on a large UK-based cohort spanning all ages. We identified subtypes with unique metabolic profiles and examined their associations with 13 chronic diseases and all-cause mortality, contributing to the advancement of personalized medicine.

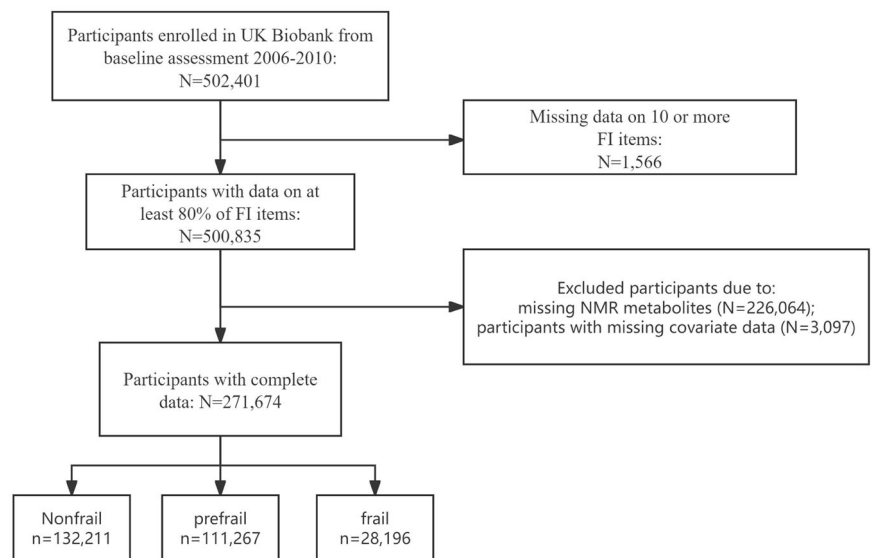
Among the 11 differential metabolites, GlycA emerged as the most significant. The association between GlycA and frailty can be mechanistically explained through its link to inflammatory pathways, particularly the interleukin-6 (IL-6) signaling axis, which is a known driver of muscle atrophy—a hallmark of frailty. GlycA, a composite biomarker of acute-phase glycoproteins, is associated with various inflammatory markers, including IL-6³⁰. The IL-6 pathway is critical in regulating muscle metabolism, where chronic elevation can drive muscle wasting. In experimental models, such as in mice with colon cancer, elevated IL-6 levels have been observed, and IL-6 inhibition was shown to prevent cancer-induced muscle mass loss³¹. IL-6 exerts its

effects by binding to the glycoprotein 130 (GP130) receptor, activating Janus kinases (JAKs) and the signal transducer and activator of transcription 3 (STAT3) pathway. This signaling cascade has been linked to muscle atrophy in numerous contexts. Prolonged IL-6 elevation in cultured myotubes and skeletal muscle cells can lead to increased expression of mitochondrial fission proteins (DRP-1 and FIS-1), which are implicated in muscle atrophy and cellular stress responses³¹. Thus, elevated GlycA levels in our high-risk subtypes (particularly HGVF) may reflect activation of this IL-6-mediated proteolytic pathway, contributing to the frailty phenotype and its associated adverse outcomes.

Linoleic acid (LA), a polyunsaturated fatty acid (PUFA), showed high SHAP values. Mendelian randomization studies suggest that elevated PUFA levels may prevent frailty³², possibly owing to LA's antioxidant and anti-inflammatory properties^{33,34}. LA's influence on metabolic syndrome and related diseases may partly explain the improved prognosis observed in frail individuals with high LA concentrations. Subtype II, characterized by high Alb levels, aligns with research highlighting Alb's mediating role in frailty and in-hospital mortality among patients with COPD, possibly via inflammatory mechanisms^{35,36}. Conversely, subtype III, with low Alb levels, exhibited a poorer prognosis. This finding reinforces the association between frailty and hypoalbuminaemia, as demonstrated in older adults as well as surgical and hospitalized patients³⁶. Subtype IV was distinguished by elevated Val levels, which may negatively affect frailty by inhibiting muscle synthesis—a hallmark of frailty linked to poor outcomes^{37–39}. Restricting Val intake has shown the potential to improve frailty conditions in preclinical models³⁸. It is worth noting that the metabolic profiles defining our frailty subtypes are not formed in isolation but are profoundly influenced by a constellation of behavioral, social, and clinical factors. Socioeconomic status (SES) is a fundamental determinant of health, shaping dietary patterns, access to nutrient-rich foods, and exposure to chronic stress, all of which can directly modulate systemic inflammation (e.g., GlycA levels) and fatty acid metabolism^{40,41}. Consequently, the adverse metabolic signatures observed in our high-risk subtypes (III and IV) may be partially driven by socioeconomic disparities. Furthermore, medication use represents a critical, often necessary, confounder in metabolic studies. For instance, statins drastically alter cholesterol and lipoprotein metabolism⁴², while metformin and anti-inflammatory drugs can influence insulin sensitivity and inflammatory pathways⁴³. The distinct metabolite levels we observed (e.g., in LA/FA, GlycA) could therefore reflect both the underlying pathophysiology of frailty and the metabolic effects of treatments for its associated comorbidities. Lastly, physical activity is a powerful modulator of the metabolome, influencing energy substrate utilization, insulin sensitivity, and inflammation⁴⁴. Sedentary behavior, often more prevalent in frail individuals, can lead to ectopic fat accumulation and reduced mitochondrial function, thereby contributing to the pro-inflammatory characteristic of our high-risk subtypes. In summary, while we identified distinct metabolite-driven frailty subtypes, their manifestation is likely orchestrated by a complex interplay between biology, behavior (diet and exercise), social determinants, and clinical management.

Given the limitations of the frailty phenotype and FI⁴⁵, integrating novel frailty subtypes based on metabolomics and clinical data into routine practice could enhance dynamic monitoring and stratified management of frailty. For the high-risk frailty group, promoting adherence to healthy dietary programs could actively reduce the risk of chronic conditions, such as coronary artery disease, heart failure, MI, MACE, MASLD, SLD, ESRD, COPD, PAD, AAA, lung cancer, kidney cancer, and all-cause mortality. This comprehensive strategy may alleviate the burden of frailty and improve clinical outcomes. In detail, first, for individuals identified as high-risk frailty (Subtypes III and IV), clinicians should prioritize them for intervention. The metabolic characteristics of these patients, such as systemic inflammation (high GlycA), amino acid metabolism dysregulation (high Val), and fatty acid composition imbalance (low LA/FA, high MUFA/FA), provide clear therapeutic targets for intervention. Second, personalized

Fig. 6 | Flowchart of participant enrollment.



management plans should be developed for this high-risk group. Our research suggests that active lifestyle interventions, particularly nutritional therapies, are most beneficial for them. Based on the results of this study, we propose the following specific dietary recommendations: For Subtype III (LALF: low albumin, low linoleic acid): Encourage the intake of high-quality protein (to boost serum albumin levels) and foods rich in Omega-6 PUFA (such as linoleic acid) (e.g., soybean oil, sunflower oil, nuts) to improve the fatty acid profile. For Subtype IV (HG VF: high GlycA, high valine): The intervention should focus on anti-inflammatory diets (e.g., consuming more Omega-3 rich fish, reducing processed foods and saturated fats) and consider limiting foods high in branched-chain amino acids (BCAAs) (e.g., certain red meats and dairy products) to reduce inflammation and regulate amino acid metabolism.

The limitations of this study should be considered when interpreting the results. First, the UK Biobank cohort predominantly consists of individuals of European descent, with approximately 95% of participants identified as White, as detailed in Table 1. This European bias limits the generalizability of the findings to other racial or ethnic groups. Therefore, caution is warranted when applying these results to more diverse populations, and further validation in Asian and African cohorts is essential to assess the robustness and applicability of these findings across different genetic and environmental contexts. Second, although the NMR metabolomics platform by Nightingale Health offers a comprehensive and standardized metabolite assessment, it does not capture the full blood metabolome and includes a limited range of metabolites. Additionally, some biomarkers lack disease specificity, necessitating further exploration of their links to frailty. Third, randomized controlled trials are essential to validate the clinical utility and impact of these newly defined frailty subtypes. Future research should focus on evaluating their effectiveness and feasibility across different medical settings and populations, considering factors such as cost-effectiveness, resource allocation, and clinical acceptance. Finally, we acknowledge that the absence of spatial transcriptomics and single-cell RNA sequencing data limits our mechanistic understanding of frailty subtypes and chronic diseases. These techniques are crucial for exploring cellular heterogeneity and intercellular interactions. Future research should focus on validating mechanisms such as liver-muscle crosstalk, using these advanced methods to provide deeper insights and identify potential therapeutic targets.

In summary, our study elucidates metabolites associated with frailty, highlights the potential of novel frailty subtypes in managing frail

individuals with unhealthy metabolism, and underscores the role of a healthy diet in mitigating the risk of chronic diseases and promoting health.

Methods

Study population

The UK Biobank was approved by the North West Research Ethics Committee (REC reference: 21/NW/0157) and all participants signed an informed consent. This community-based cohort study included about 500,000 volunteers from England, Scotland, and Wales. Baseline socio-demographic, lifestyle, health-related data, and blood samples were collected between March 2006 and October 2010.

Participants with missing data on 10 or more frailty-related items, missing metabolite information, prefrailty, or covariates with >20% missing values were excluded. Finally, 160,407 participants were included, with a median follow-up time of 13.8 years (Fig. 6).

Measurement of frailty

Frailty was assessed using the FI, which reflects the accumulation of health deficits. This index is based on multiple indicators across various physiological and psychological domains, including symptoms, diagnosed diseases, and disabilities¹⁶. From the UK Biobank, 49 items were selected to construct an FI (Supplementary Table 1)¹⁷. Each participant's health deficits were assessed, and the total number of health deficits was divided by 49 to calculate the FI, which ranged from 0 to 1. Participants with an FI ≤ 0.10 were classified as non-frail, those with an FI of 0.10–0.21 as pre-frail, and those with an FI > 0.21 as frail¹⁸.

Metabolite analysis

The UK Biobank's NMR metabolomic data included 251 metabolomic biomarkers (170 absolute concentrations and 81 derived ratios) from the EDTA plasma samples of approximately 280,000 participants. These biomarkers include clinically validated indicators, such as cholesterol, fatty acids, amino acids, and inflammation markers, as well as emerging biomarkers, such as lipoprotein subclasses. For this study, all 251 metabolic biomarkers were analyzed (For more details, refer to Class 220 of the UK Biobank and Supplementary Table 2).

Selection of differential metabolites

The CatBoost algorithm was used for feature selection, with frail and non-frail populations as binary outcomes. The dataset was allocated into training and validation sets in a 7:3 ratio, and optimal model parameters were

determined through five-fold cross-validation. SHAP values were then calculated to rank features according to their importance, and 11 differential metabolites were selected for further analysis based on these SHAP values. SHAP values were applied to interpret the model outputs and quantify feature importance. This method is grounded in cooperative game theory, which fairly allocates the contribution of each feature to the final prediction by considering all possible combinations of features. The mean absolute SHAP value was used to rank the features, where a higher value indicates a greater overall influence on the model's prediction. The direction and magnitude of each feature's effect were also interpreted: a positive SHAP value indicates that the feature increases the predicted risk of frailty, while a negative value suggests a protective effect. The dispersion of SHAP values reflects the consistency and strength of each feature's influence across the population. XGBoost modeling was used for hyperparameter tuning and five-fold cross-validation to ensure the model's best performance and reliability. The accuracy of the 11 differential metabolites in identifying frailty status was assessed using the AUC as the performance standard.

Determination of cluster number

PCA was performed on the differential metabolites for dimensionality reduction, and the optimal number of clusters was determined using an elbow plot and silhouette coefficient. The elbow plot calculates the sum of squared errors for different cluster numbers and identifies the "elbow" position, indicating the optimal number of clusters. The silhouette coefficient combines the cohesion and separation of clusters, with values ranging from -1 to 1 ; a higher value indicates better clustering performance. K-means clustering analysis was then performed. We generated 100 bootstrap samples and applied K-means clustering to each sample. The stability of the clusters was assessed using the Adjusted Rand Index (ARI) and Normalized Mutual Information (NMI), with values greater than 0.8 indicating strong stability.

Definition of outcome events

Thirteen chronic diseases and all-cause mortality were considered, including coronary artery disease, heart failure, MACE, MI, type 2 diabetes, MASLD, COPD, SLD, PAD, ESRD, kidney cancer, lung cancer, and AAA. Outcomes were determined using the International Classification of Diseases, 10th Edition codes, excluding individuals with these outcomes prior to baseline. Prognoses between different subtypes were assessed using a Cox proportional hazards model and Kaplan–Meier survival curves.

Dietary intake assessment

A touchscreen questionnaire completed by participants at baseline was used to collect data on the consumption frequency of 12 foods over the past year: beef, lamb, pork, processed meat, oily fish, non-oily fish, fresh fruit, dried fruit, raw vegetables, cooked vegetables, grains, and bread. New data fields were created for the intake of (1) red meat, (2) total fish, (3) total vegetable, (4) total fruit, (5) whole grain, and (6) refined grain. Red meat intake was derived from the sum of beef, lamb, and pork; total fish intake was derived from the sum of oily and non-oily fish. Total vegetable intake combined cooked and salad/raw vegetables; total fruit intake considered fresh and dried fruit. Whole and refined grains were categorized based on bread and other grain types consumed. Participants' food intake was categorized into seven groups: red meat, processed meat, total fish, total fruit, total vegetables, whole grains, and refined grains. Portion sizes were defined for each food item, and weekly consumption data for bread and grain were converted into daily consumption data²². For more information on the diet, please refer to Supplementary Table 4.

Healthy diet score estimation

We used seven dietary factors and critical value based on recommendations for cardiometabolic health: increasing fruits, vegetables, whole grains, and fish intake and reducing red meat, processed meat, and refined grain intake^{49,50}. A healthy diet score was calculated using these components: total fruit ≥ 3 servings/day; total vegetables ≥ 3 servings/day; total fish ≥ 2 servings/

week; processed meat ≤ 1 serving/week; red meat ≤ 1.5 servings/week; whole grains ≥ 3 servings/day; and refined grains ≤ 1.5 servings/day. Each favorable dietary factor received 1 point, with scores ranging from 0 to 1. Participants were classified into unhealthy (score < 4) and healthy (score ≥ 4) diet categories (Supplementary Table 3).

Covariate assessment

Covariate data were collected through self-completed touchscreen questionnaires or oral interviews. These included age, sex, ethnicity (White or other), current smoking and drinking status, BMI, SBP, DBP, and ALT, AST, ALP, cholesterol, direct bilirubin, total bilirubin, triglycerides, and HbA1c levels. Missing covariate data (less than 20%) were imputed using the MICE package in R.

All analyses were conducted using R (v4.1.3) and Python (v3.8.19). Statistically, significance was defined as $P < 0.05$. For multiple comparisons across the outcomes, we applied False Discovery Rate (FDR) correction using the Benjamini-Hochberg method to control for false positives.

Data availability

Data obtained from the UK Biobank are available on application at [<https://www.ukbiobank.ac.uk/register-apply>] (<https://www.ukbiobank.ac.uk/register-apply>) (92668).

Received: 22 July 2025; Accepted: 8 October 2025;

Published online: 05 December 2025

References

- Hewitt, J. et al. The prevalence of frailty and its association with clinical outcomes in general surgery: a systematic review and meta-analysis. *Age Ageing* **47**, 793–800 (2018).
- O'Caomh, R. et al. Prevalence of frailty in 62 countries across the world: a systematic review and meta-analysis of population-level studies. *Age Ageing* **50**, 96–104 (2021).
- Ferrucci, L. & Fabbri, C. Inflammageing: chronic inflammation in ageing, cardiovascular disease, and frailty. *Nat. Rev. Cardiol.* **15**, 155–171 (2018).
- Clegg, A., Young, J., Iliffe, S., Rikkert, M. O. & Rockwood, K. Frailty in elderly people. *Lancet* **381**, 752–762 (2013).
- Chen, S. et al. Frailty and long-term survival of patients with colorectal cancer: a meta-analysis. *Ageing Clin. Exp. Res.* **34**, 1485–1494 (2022).
- Gilardi, F. et al. Frailty screening and assessment tools: a review of characteristics and use in public health. *Ann. Ig.* **30**, 128–129 (2018).
- Jiang, R. et al. Associations of physical frailty with health outcomes and brain structure in 483 033 middle-aged and older adults: a population-based study from the UK Biobank. *Lancet Digit. Health* **5**, e350–e359 (2023).
- Chen, Y. et al. Genetic and phenotypic associations of frailty with cardiovascular indicators and behavioral characteristics. *J. Adv. Res.* **71**, 263–277 (2025).
- Li, J. J. et al. Comparison of three frailty scales for prediction of adverse outcomes among older adults: a prospective cohort study. *J. Nutr. Health Aging* **25**, 419–424 (2021).
- Oviedo-Briones, M. et al. The ability of eight frailty instruments to identify adverse outcomes across different settings: the FRAILTOOLS project. *J. Cachexia Sarcopenia Muscle* **13**, 1487–1501 (2022).
- Mishra, M., Wu, J., Kane, A. E. & Howlett, S. E. The intersection of frailty and metabolism. *Cell Metab.* **36**, 893–911 (2024).
- Cao, X. et al. Associations between frailty and the increased risk of adverse outcomes among 38,950 UK Biobank participants with prediabetes: prospective cohort study. *JMIR Public Health Surveill.* **9**, e45502 (2023).
- Xiao, L. et al. Polygenic risk score of metabolic dysfunction-associated steatotic liver disease amplifies the health impact on severe liver disease and metabolism-related outcomes. *J. Transl. Med.* **22**, 650 (2024).

14. Zheng, H. et al. Association between the atherogenic index of plasma trajectory and risk of heart failure among hypertensive patients: a prospective cohort study. *Cardiovasc. Diabetol.* **23**, 301 (2024).
15. DiNicolantonio, J. J., Lucan, S. C. & O'Keefe, J. H. The evidence for saturated fat and for sugar related to coronary heart disease. *Prog. Cardiovasc. Dis.* **58**, 464–472 (2016).
16. Wu, J. H. Y., Micha, R. & Mozaffarian, D. Dietary fats and cardiometabolic disease: mechanisms and effects on risk factors and outcomes. *Nat. Rev. Cardiol.* **16**, 581–601 (2019).
17. Al-Qawasmeh, R. H. & Tayyem, R. F. Dietary and lifestyle risk factors and metabolic syndrome: literature review. *Curr. Res. Nutr. Food Sci. J.* **6**, 594–608 (2018).
18. Zheng, H., Sechi, L. A., Navarese, E. P., Casu, G. & Vidili, G. Metabolic dysfunction-associated steatotic liver disease and cardiovascular risk: a comprehensive review. *Cardiovasc. Diabetol.* **23**, 346 (2024).
19. Kojima, G., Avgerinou, C., Iliffe, S. & Walters, K. Adherence to Mediterranean diet reduces incident frailty risk: systematic review and meta-analysis. *J. Am. Geriatr. Soc.* **66**, 783–788 (2018).
20. Lopez-Garcia, E., Hagan, K. A., Fung, T. T., Hu, F. B. & Rodríguez-Artalejo, F. Mediterranean diet and risk of frailty syndrome among women with type 2 diabetes. *Am. J. Clin. Nutr.* **107**, 763–771 (2018).
21. Struijk, E. A. et al. Diet quality and risk of frailty among older women in the Nurses' Health Study. *Am. J. Clin. Nutr.* **111**, 877–883 (2020).
22. Liu, W., Wang, T., Zhu, M. & Jin, G. Healthy diet, polygenic risk score, and upper gastrointestinal cancer risk: a prospective study from UK Biobank. *Nutrients* **15**, 1344 (2023).
23. Tang, R. et al. Smoking timing, healthy diet, and risk of incident CKD among smokers: findings from UK Biobank. *Am. J. Kidney Dis.* **84**, 593–600.e591 (2024).
24. Zhang, T. et al. Joint association of serum urate and healthy diet with chronic obstructive pulmonary disease incidence: results from the UK Biobank study. *Food Funct.* **15**, 4642–4651 (2024).
25. Huang, J. et al. Association of healthy diet score and adiposity with risk of colorectal cancer: findings from the UK Biobank prospective cohort study. *Eur. J. Nutr.* **63**, 2055–2069 (2024).
26. Yan, Z. et al. Development of frailty subtypes and their associated risk factors among the community-dwelling elderly population. *Aging* **12**, 1128–1140 (2020).
27. Chew, J. et al. Disentangling the relationship between frailty and intrinsic capacity in healthy community-dwelling older adults: a cluster analysis. *J. Nutr. Health Aging* **25**, 1112–1118 (2021).
28. Bohn, L., Zheng, Y., McFall, G. P. & Dixon, R. A. Portals to frailty? Data-driven analyses detect early frailty profiles. *Alzheimers Res. Ther.* **13**, 1 (2021).
29. Okoye, C. et al. Predicting mortality and re-hospitalization for heart failure: a machine-learning and cluster analysis on frailty and comorbidity. *Aging Clin. Exp. Res.* **35**, 2919–2928 (2023).
30. Dungan, K., Binkley, P. & Osei, K. GlycA is a novel marker of inflammation among non-critically ill hospitalized patients with type 2 diabetes. *Inflammation* **38**, 1357–1363 (2015).
31. Schumertl, T., Lokau, J. & Garbers, C. IL-6 signaling in immunopathology: from basic biology to selective therapeutic intervention. *Immunotargets Ther.* **14**, 681–695 (2025).
32. Tomata, Y., Wang, Y., Hägg, S. & Jylhävä, J. Fatty acids and frailty: a Mendelian randomization study. *Nutrients* **13**, 3539 (2021).
33. Spittler, G. Peroxidation of linoleic acid and its relation to aging and age dependent diseases. *Mech. Ageing Dev.* **122**, 617–657 (2001).
34. Wang, Q. et al. Dietary oleic acid supplementation and blood inflammatory markers: a systematic review and meta-analysis of randomized controlled trials. *Crit. Rev. Food Sci. Nutr.* **62**, 2508–2525 (2020).
35. Li, L. et al. Frailty and in-hospital mortality in older patients with acute exacerbation of COPD: a real-world prospective cohort study. *Respir. Med.* **228**, 107663 (2024).
36. Zhang, L. et al. Association between frailty and hypoproteinaemia in older patients: meta-analysis and systematic review. *BMC Geriatr.* **24**, 689 (2024).
37. Guo, K. et al. Effects of sarcopenia and frailty on postoperative recovery in elderly patients: a prospective cohort study. *J. Cachexia Sarcopenia Muscle* **14**, 2642–2652 (2023).
38. Richardson, N. E. et al. Lifelong restriction of dietary branched-chain amino acids has sex-specific benefits for frailty and life span in mice. *Nat. Aging* **1**, 73–86 (2021).
39. Wolfe, R. R. Branched-chain amino acids and muscle protein synthesis in humans: myth or reality? *J. Int. Soc. Sports Nutr.* **14**, 30 (2022).
40. Stringhini, S. et al. Socioeconomic status and the 25×25 risk factors as determinants of premature mortality: a multicohort study and meta-analysis of 1.7 million men and women. *Lancet* **389**, 1229–1237 (2017).
41. Hu, F. R. et al. The association of socioeconomic status with incident frailty and mortality: evidence from the UK Biobank. *J. Nutr. Health Aging* **27**, 1015–1024 (2023).
42. Stancu, C. & Sima, A. Statins: mechanism of action and effects. *J. Cell Mol. Med.* **5**, 378–387 (2001).
43. Lawler, P. R. et al. Targeting cardiovascular inflammation: next steps in clinical translation. *Eur. Heart J.* **42**, 113–131 (2021).
44. Cunningham, C. et al. Consequences of physical inactivity in older adults: a systematic review of reviews and meta-analyses. *Scand. J. Med. Sci. Sports* **30**, 816–827 (2020).
45. Marangoni, F. et al. Dietary linoleic acid and human health: focus on cardiovascular and cardiometabolic effects. *Atherosclerosis* **292**, 90–98 (2020).
46. Searle, S. D., Mitnitski, A., Gahbauer, E. A., Gill, T. M. & Rockwood, K. A standard procedure for creating a frailty index. *BMC Geriatr.* **8**, 24 (2008).
47. Williams, D. M., Jylhävä, J., Pedersen, N. L. & Hägg, S. A frailty index for UK Biobank participants. *J. Gerontol. A* **74**, 582–587 (2019).
48. Sun, Y. et al. Frailty, genetic predisposition, and incident atrial fibrillation. *Eur. Heart J.* **45**, 1281–1283 (2024).
49. Mozaffarian, D. Dietary and policy priorities for cardiovascular disease, diabetes, and obesity: a comprehensive review. *Circulation* **133**, 187–225 (2016).
50. Lourida, I. et al. Association of lifestyle and genetic risk with incidence of dementia. *JAMA* **322**, 430–437 (2019).

Acknowledgements

This research was conducted under application number 92668 from the UK Biobank resource. We extend our sincere gratitude to all participants and investigators involved in the UK Biobank study for generously sharing their data. We also acknowledge BioRender.com, R, and Python for providing the essential tools used to create the figures in this work. Additionally, we thank Editage (<https://www.editage.cn>) for their professional English-language editing services. This work was supported by the National Key R&D Program of China (No. 2023YFC2308500, 2021YFC2500805), National Nature Science Foundation of China (No. 81972897, 82172751, 82404077, 82400664), Guangdong Natural Science Foundation (No. 2022A1515110656), the Postdoctoral Fellowship Program of CPSF (Grant Number GZC20240663).

Author contributions

L. Xiao, Q. Deng, J. Wang, S. Liang, and R. Li contributed equally to this work. C. Hong, W. Lai, and L. Xiao performed conceptualization. L. Zeng and L. Liu provided resources. Q. Deng and J. Wang conducted the investigation and developed the methodology. Q. Deng, C. Hong, and L. Zeng performed data curation. L. Xiao and Q. Deng conducted metabolite analysis. L. Xiao and J. Wang performed cluster analysis. L. Xiao, Q. Deng, and J. Wang conducted the prognosis analysis. S. Liang and R. Li wrote the original draft. P. Jiang, H. Cui, Y. Li, R. Zhang, and L. Zeng reviewed and edited the manuscript. L. Xiao,

C. Hong, and L. Liu acquired funding. L. Liu, W. Lai, C. Hong, and L. Zeng provided supervision. All authors read and approved the manuscript.

Competing interests

The authors declare no competing interest.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41746-025-02075-2>.

Correspondence and requests for materials should be addressed to Lin Zeng, Chang Hong, Weinan Lai or Li Liu.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025