npj Aging

https://doi.org/10.1038/s41514-026-00400-x

Article in Press

Circulating levels of insulin-like growth factor binding protein 7 are associated with risks of chronic diseases and death

Received: 19 December 2025

Accepted: 29 April 2026

Cite this article as: Li, Z., Tao, C., Zhou, Z. *et al.* Circulating levels of insulin-like growth factor binding protein 7 are associated with risks of chronic diseases and death. *npj Aging* (2026). https://doi.org/10.1038/s41514-026-00400-x

Zhi Li, Chengzhe Tao, Ziyi Zhou, Michael N. Pollak, Edward L. Giovannucci & Mingyang Song

Circulating levels of insulin-like growth factor binding protein 7 are associated with risks of chronic diseases and death

Zhi Li [a, b, 1], Chengzhe Tao [c, d, 1], Ziyi Zhou [d, e, 1], Michael N Pollak [f, *], Edward L Giovannucci [b, e, *], Mingyang Song [b, e, *]

[a] China CDC Key Laboratory of Environment and Population Health, National Institute of Environmental Health, Chinese Center for Disease Control and Prevention, Beijing, China.
[b] Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA, USA.
[c] State Key Laboratory of Reproductive Medicine and Offspring Health, Center for Global Health, School of Public Health, Nanjing Medical University, Nanjing 211166, China.
[d] School of Health and Wellbeing, University of Glasgow, United Kingdom.
[e] Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA.
[f] Departments of Oncology and Medicine, McGill University, Montreal, Quebec, Canada.

[1] Zhi Li, and Chengzhe Tao, Ziyi Zhou contribute equally to this study and should be regarded as joint first authors.

[*] Joint last authors, to whom correspondence should be addressed at:
Dr. Mingyang Song
Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA.
E-mail: msong@hsph.harvard.edu

Dr. Edward L Giovannucci
Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA, USA.
E-mail: egiovann@hsph.harvard.edu

Dr. Michael N Pollak
McGill University, Montreal, Quebec, Canada.
E-mail: michael.pollak@mcgill.ca

Abstract

Relationships between the concentration of circulating IGFBP-7 and risk of disease and mortality have been suggested by small-scale investigations. In this prospective study, we investigated these relationships among 53,003 UK Biobank participants. Higher IGFBP-7 level was significantly associated with increased risk for liver cancer, all-cause mortality, diabetes, and other diseases. Associations were robust

across sex and age groups and persisted over long follow-up. IGFBP-7 polygenic risk scores also predicted cancer and mortality risk. IGFBP-7 level was strongly correlated with levels of previously identified aging-related proteins, but after adjustment for these proteins, remained associated with risk of bladder cancer, liver cancer, multiple myeloma, all-cause mortality, liver-related mortality, and diabetes. Our findings indicate IGFBP-7 as a novel biomarker of mortality and disease risk.

Introduction

Insulin-like growth factor-binding protein-7 (IGFBP-7) is a member of the insulin-like growth factor binding protein (IGFBP) family but differs from other IGFBPs in its lower affinity for IGFs and higher affinity for insulin[1-3]. IGFBP-7 has been reported to be a key senescence-inducing factor, promoting senescence and apoptosis in mesenchymal stem cells (MSCs) and other systems [4-6]. While its biological functions are not fully understood, there is evidence that IGFBP-7 plays roles in developmental processes, tumor growth, and angiogenesis[7-9]. Certain functions of IGFBP-7 may be unrelated to its IGF or insulin binding properties but rather due to its interactions with other proteins, including the unexpected binding to the IGF-I receptor and the well characterized binding to CD93, a key protein mainly expressed on endothelial cells and involved in intercellular adhesion, suggesting the relevance of IGFBP-7 to endothelial biology [10,11].

Several previous studies confined to specific subpopulations have provided early evidence that IGFBP-7 level is related to risk of metabolic, neoplastic, and

cardiovascular diseases[12-14]. For cardiovascular disease, IGFBP-7 appears to mediate inflammation- and fibrosis-related pathways influencing cardiac outcomes, whereas for breast cancer, IGFBP-7 modulates IGF-I receptor–related signaling that affects prognosis and treatment response[15-17]. However, these prior studies were limited by modest sample sizes and a narrow scope of outcomes. Large-scale epidemiological studies are therefore needed to generate more precise and robust risk estimates, evaluate less common disease endpoints, and enable complementary longitudinal and genetic analyses of circulating IGFBP-7 in relation to disease and mortality.

In a prospective cohort study within the UK Biobank (UKB), we comprehensively characterized the associations of IGFBP-7 with risk of all-cause and cause-specific mortality, overall cancer, 21 site-specific cancers, CVD, and diabetes. In addition to circulating levels, we constructed a polygenic risk score (PRS) for IGFBP-7, based on the genome-wide association data, to evaluate the genetically predicted IGFBP-7 in relation to the outcomes[18].

Results

Population characteristics

The study design and main analytical approaches are summarized schematically in **Figure 1**. The characteristics of participants in this study are summarized in **Table 1**. A total of 53,003 participants with available measurement of IGFBP-7 level were included. The mean age of participants was 56.8 years, with 46% being male. The average BMI was 27.5 kg/m^2. Participants in the highest IGFBP-7 quartile were older,

had moderately higher BMI and less favorable lifestyle, including lower levels of physical activity and healthy eating, and were somewhat more likely to smoke and consume alcohol daily. These results are of interest in the context of a recent study of environmental and genetic contributions to aging and mortality which reported major environmental effects on aging biomarkers, although not specifically examining IGFBP-7[19].

Distribution and key predictors of IGFBP-7

IGFBP-7 level was estimated by the Olink platform and expressed as Normalized Protein eXpression (NPX) values, which represent arbitrary units relative to an internal standard on a log2 scale. The distribution of IGFBP-7 is shown in **Figure 2A.** **Figure 2B** presents the proportion of variation in circulating IGFBP-7 explained by 21 selected demographic, genetic, lifestyle, biochemical, and dietary factors. For each variable, R^2 was estimated from a separate linear regression model, while the overall R^2 was obtained from a model including all variables simultaneously. The overall model accounts for 20% of the variation in IGFBP-7. Among individual predictors, age accounted for the largest proportion of variability (7.8%), followed by rs10013945 (4.4%), rs1718860 (4.2%), sex (3.3%), and BMI (2.4%). IGFBP-7 increased with age ($\beta = 0.012$ per year, $P < 0.001$; $r = 0.28$) in a roughly linear manner across 39–70 years, in both sexes, with men showing higher overall levels but a slightly weaker age-related slope (P for interaction < 0.001). We evaluated the intra-individual variability of IGFBP-7 across three time points among 1,002 individuals

with repeated measurements at baseline (instance 0, 2006-2010), first follow-up (instance 2, 2014+), and second follow-up (instance 3, 2019+). The Spearman correlation coefficients between time points ranged from 0.46 to 0.47, and the intra-class correlation coefficient (ICC) across the three measurements was 0.46 (95% CI: 0.40–0.48), indicating modest within-person stability of IGFBP-7 levels over time.

Correlation of IGFBP-7 with blood biochemistry markers and age-associated proteins

We examined the Spearman correlations between IGFBP-7 and 30 blood biochemistry markers available in the UK Biobank (**Figure 3A**). These analyses were exploratory and intended to describe the broader biomarker profile associated with IGFBP-7, rather than to infer underlying mechanisms. The Spearman correlation coefficients ranged from -0.32 to 0.53. IGFBP-7 showed the strongest positive correlation with cystatin C (r = 0.53), followed by testosterone, urate, gamma glutamyltransferase, and creatinine. Notably, inverse correlations were observed with oestradiol (r = -0.32), HDL cholesterol, IGF-1, apolipoprotein A, and SHBG. The corresponding *P* values and FDR-adjusted *P* values are provided in **Table S1**.

Given the modest positive correlation between IGFBP-7 and age, we explored correlations between IGFBP-7 and the 20 proteins included in the ProtAge20 model (**Figure 3B**), which were selected from a larger set of 204 proteins through machine learning in a multi-ethnic cohort study based on their predictions for chronological age[20]. Among the 20 proteins, IGFBP-7 was most strongly correlated with COL6A3 (r

= 0.47), GDF15 (r= 0.44), and LTBP2 (r= 0.44), showing positive associations with nearly all proteins except for IGDCC4 (r = -0.03). The corresponding *P* values and FDR-adjusted *P* values are provided in **Table S2**.

Associations between IGFBP-7 serum levels and polymorphisms with disease risk and mortality

In Model 1 with age as the time scale and adjusted for sex, each 1-standard deviation (SD) increase in IGFBP-7 levels was associated with increased risks of multiple cancers and chronic diseases, and mortality (**Figure S1**). These associations were generally attenuated but remained directionally consistent in Model 2 with multivariable adjustment for fasting status, BMI, income, Townsend index, alcohol intake, smoking status, physical activity, and healthy eating score (**Figure 4A**). Statistically significant associations persisted for overall cancer (HR = 1.08; 95% CI: 1.04-1.11), liver cancer (HR = 1.57; 95% CI: 1.45-1.70), lung cancer (HR = 1.20; 95% CI: 1.10-1.31), multiple myeloma cancer (HR = 1.42; 95% CI: 1.29-1.57), all-cause mortality (HR = 1.30; 95% CI: 1.27-1.33), liver-related mortality (HR = 1.83; 95% CI: 1.68-1.99), kidney-related mortality (HR = 1.71; 95% CI: 1.44-2.04), cardiovascular disease (HR = 1.14; 95% CI: 1.10-1.18), and diabetes (HR = 1.24; 95% CI: 1.20-1.28). The corresponding FDR-adjusted *P* values for Model 1 and Model 2 are provided in **Tables S3** and **S4**, respectively.

We further evaluated the genetic predisposition to IGFBP-7 levels using a polygenic risk score (PRS) constructed from six genome-wide significant SNPs (*P*

value < 5×10^{-8}) identified in a large-scale proteomic GWAS[18]. The PRS showed a modest but statistically significant correlation with IGFBP-7 levels (Pearson r = 0.239). In **Figure 4B**, the IGFBP-7 PRS showed fewer and weaker associations compared with measured IGFBP-7 levels, with statistical significance observed only for a subset of outcomes, including overall cancer (HR = 1.07; 95% CI: 1.00-1.14), bladder cancer (HR = 1.39; 95% CI: 1.09-1.77), and gallbladder cancer (HR = 1.92; 95% CI: 1.03-3.60), as well as all-cause mortality (HR = 1.08; 95% CI: 1.01-1.15) and cancer-specific mortality (HR = 1.30; 95% CI: 1.02-1.66). The corresponding FDR-adjusted P values for Model 1 and Model 2 are provided in **Tables S5** and **S6**, respectively. Results based on six individual SNPs are provided in **Tables S7-S12**.

To further examine whether the observed associations between IGFBP-7 and mortality were driven by chronological age, we generated an age-adjusted IGFBP-7 residual representing the deviation of each participant's IGFBP-7 level from the expected value based on age, sex, BMI, and race. We then evaluated the residual in relation to all-cause mortality. Both the original IGFBP-7 and the age-adjusted residual were strongly associated with mortality, and their predictive performance was nearly identical (C-index = 0.7795 vs 0.7782).

Stratified analysis

Sex-stratified analyses demonstrated that most associations between IGFBP-7 levels and disease outcomes observed in the overall population remained directionally consistent in both males and females (**Table S13**). Notably, the association with liver

cancer was strong in both sexes, with the HR of 1.57 (95% CI: 1.41-1.75) in males and 1.54 (95% CI: 1.29-1.83) in females under Model 2. Similarly, IGFBP-7 levels were strongly associated with all-cause mortality in both sexes (male HR = 1.30; 95% CI: 1.26-1.35; female HR = 1.29; 95% CI: 1.24-1.35). On the other hand, some sex-specific patterns were noted for other outcomes. The association between IGFBP-7 and bladder cancer was significant in females (HR = 1.23; 95% CI: 1.03-1.49) but not in males. Conversely, non-Hodgkin lymphoma showed a significant positive association with IGFBP-7 in males (HR = 1.47; 95% CI: 1.29-1.67) but not females. However, none of the differences reached statistical significance in the interaction analysis with sex (**Table S13**). No significant associations were observed for melanoma, ovarian cancer, or pancreatic cancer in either sex after multivariable adjustment.

As shown in **Table S14**, the associations between IGFBP-7 and major outcomes, including liver cancer, all-cause mortality, and diabetes, were observed in both age groups (<65 and ≥65 years). HRs for most outcomes were slightly higher in the younger than older group.

Tables S15 and **S16** show that the associations between IGFBP-7 and liver cancer, all-cause mortality, diabetes, and other outcomes were generally stronger with earlier follow-up (<5 or <10 years) but remained statistically significant for events occurring later (≥5 or ≥10 years).

Sensitivity analysis

In sensitivity analyses, we additionally adjusted for IGFBP-1L and IGFBP-4 (analyses 1) to examine whether the associations of IGFBP-7 with outcomes were confounded by other IGFBP family members. We also adjusted for the available individual proteins from the ProtAge20 panel (analyses 2). Since the exact algorithm used to calculate the composite ProtAge20 score is not publicly available and cannot be directly applied to our data, we used these individual proteins in a sensitivity analysis to examine whether the associations of IGFBP-7 with outcomes could be explained by previously reported age-related proteomic signals (**Table 2**).

The associations between IGFBP-7 and major outcomes, particularly all-cause mortality and liver cancer, remained robust after these adjustments, although the effect estimates for certain outcomes (e.g., kidney-cause mortality) were attenuated. Although IGFBP-7 was correlated with several proteins in the ProtAge20 panel, its associations with disease and mortality remained significant after mutual adjustment. The persistence of several associations after adjustment suggests that the observed associations of IGFBP-7 with these outcomes were not fully explained by the available proteins examined from the ProtAge20 panel. For all-cause mortality, the hazard ratios were 1.18 (95% CI: 1.14-1.22) after additional adjustment for IGFBP-1L and IGFBP-4 (sensitivity analysis 1) and 1.08 (95% CI: 1.03-1.13) after adjustment for the 20 available proteins from ProtAge20 (sensitivity analysis 2), respectively. For liver cancer, the associations were also consistently significant, with HRs of 1.54 (95% CI: 1.39-1.70) and 1.36 (95% CI: 1.09-1.69) for sensitivity analysis 1 and 2,

respectively. The association between IGFBP-7 and multiple myeloma, cancer, CVD, diabetes remained significant in the two sensitivity analyses. In contrast, while the association with overall cancer remained positive, its statistical significance was attenuated after adjusting for the broader proteomic panel in sensitivity analysis 2. We also separately analyzed the associations between the 20 available proteins in ProtAge20 and diseases (**Table S17-S19**). The corresponding FDR-adjusted *P* values are provided in **Tables S20-S22**. For all-cause mortality, the strongest associations were observed for GDF15 (HR = 1.59; 95% CI: 1.51-1.63), FSHB (HR = 1.32; 95% CI: 1.31-1.34), and KLK3 (HR = 1.31; 95% CI: 1.28-1.34). For liver cause mortality, the most strongly associated proteins were GDF15 (HR = 1.72; 95% CI: 1.56-1.90), FSHB (HR = 1.46; 95% CI: 1.36-1.56), and ENG (HR = 1.44; 95% CI: 1.29-1.60). The effect sizes for IGFBP-7 were modest compared with GDF15 but similar in magnitude to other proteins associated with mortality in ProtAge20.

Discussion

Using proteomic and genomic data from UK Biobank, we provide robust evidence of relationships between circulating IGFBP-7 level and risks of various health outcomes, including all-cause and cause-specific mortality, overall cancer risk, specific cancer risk, cardiovascular disease, and diabetes. Polygenic risk scores for IGFBP-7 showed generally weaker or null associations with many of these outcomes, consistent with the evidence that non-genetic factors are important determinants of the interpersonal variability in IGFBP-7 levels.

We cannot determine the extent to which circulating IGFBP-7 acts as a marker of disease-related biological processes rather than as direct causal driver but note that in a murine model of heart failure, genetic and immunologic inactivation of IGFBP-7 improved cardiac function, consistent with a biological role at least in this context[21].

Systemic rather than organ-specific processes may underlie our observations, given the broad range of disease associations with IGFBP-7 level. For example, higher IGFBP-7 levels may reflect chronic inflammation that predisposes to cardiovascular and neoplastic disease. It is also possible that IGFBP-7 elevation indicates endothelial dysfunction that may increase the risk of diseases across organs. Although the physiologic implications are obscure, there is clear evidence that IGFBP-7 binds to CD93, which is highly expressed by endothelial cells and is involved in angiogenesis and neovascularization[11].

Our findings complement recent machine learning-based studies to identify aging-related proteins, such as those included in the ProtAge20 panel[20,22-25]. Although IGFBP-7 was not selected by these algorithms as a top-ranked predictor of biological aging, it demonstrated robust and consistent associations with multiple aging-related outcomes in our hypothesis-driven analysis. Rather, given its correlations with several proteins in the ProtAge20 panel, IGFBP-7 may reflect biological processes that overlap with the previously reported age-related proteomic signals. Schoenfeld residual analyses indicated a deviation from the proportional hazards' assumption in some main outcome models. Therefore, the hazard ratios from the Cox models should

be interpreted as the average relative risk estimates over follow-up rather than strictly time-invariant effects. A further limitation is that IGFBP-7 was measured using the Olink platform and expressed as NPX values, which are relative rather than absolute quantitative measures. We did not have independent validation using a high-precision immunoassay in this cohort, which may limit direct cross-platform comparison and immediate clinical translation.

In conclusion, the current study represents the largest investigation of the relationship of circulating IGFBP-7 to mortality and disease risk. Given that our cohort consisted primarily of European participants, replication in other populations is warranted. Moreover, although the biological basis for our observational findings remains to be elucidated, the persistence of the associations after adjustment for aging-related proteins supports an independent role of IGFBP-7 in aging. Additionally, measurement of IGFBP-7 at multiple time points in a subset of participants demonstrated moderate stability over nearly a decade, supporting its robustness as a circulating biomarker. Future studies integrating experimental and longitudinal data are warranted to elucidate its causal pathways and therapeutic potential.

Methods

Study population

The UK Biobank is a large prospective cohort study of ~500,000 individuals aged 39-70, recruited from 2006 to 2010, with extensive phenotypic data for long-term

health follow-up and genotype data[26,27]. All data were collected with informed consent from participants. The scientific protocol and operational procedures were reviewed and approved by the North West Research Ethics Committee (https://www.ukbiobank.ac.uk/media/gnkeyh2q/study-rationale.pdf) (REC Reference:21/NW/0157). Of these, 53,013 participants had proteomic data. In this study, after exclusion of participants with missing proteomic data, a total of 53,003 participants were included in this study. This study was carried out under the application #46466. The study was conducted in accordance with the Declaration of Helsinki.

Outcome assessment

Data on overall cancer, specific-site cancers (more than 100 incident cases, excluding non-melanoma skin cancer), CVD, type 2 diabetes (T2D), all-cause mortality, and specific-cause mortality were collected through hospital records and death registries linked to the UK Biobank, classified by ICD-10 codes (International classification of diseases 10th revisions). The cause and date of death were sourced from the National Health Service Central Register for Scotland and the National Health Service Information Centre for England and Wales. Specific-causes of mortality were classified based on the following codes: cancer-cause mortality (C00-C97), CVD-cause mortality (I20-I25, I60, I61, I63, I64), liver-cause mortality (K70-K76), kidney-cause mortality (N00-N28).

Cancer outcomes were defined based on the ICD10 codes from the national cancer registry. In addition to overall cancer, specific-site cancers with at least 100 incident cases were focused. Further information on the ICD10 codes of specific-site cancers was provided in the **Table S23**.

CVD were defined as any coronary heart disease (ICD-10 codes I20-I25) or stroke (ICD-10 codes I60, I61, I63, and I64), while T2D was diagnosed using ICD-10 code E11. For each incident disease analysis, participants with prevalent disease of the corresponding type at baseline were excluded, such that only incident events occurring after the baseline assessment were considered as outcomes.

Proteomics assay

The baseline blood samples of participants were obtained from assessment centers between 2007-2010. Proteomic profiling was measured by antibody-based Olink Explore™ Proximity Extension Assay, sample processing and quality control protocols were applied in the prior studies[28-30]. The level of IGFBP-7 were expressed as normalized protein expression (NPX) values, where a one-unit increase in NPX corresponds to an approximate doubling of protein concentration[31]. The protein was standardized and incorporated into the further analysis. IGFBP-7 was measured at instances 0, 2, and 3. Spearman correlations across these time points (among 1,100 individuals with repeated measurements) were used to assess within-person consistency.

Genetic determinants associated with IGFBP-7

SNPs representing the most significant association with levels of IGFBP-7 in a region (±1 Mb) were identified as genetic determinants from a large-scaled European proteomic Genome-wide association study (GWAS)[18]. The selection criterion for SNPs was *P* value < 5×10^{-8}. Detailed information on selected SNPs was shown in **Table S24**. All the SNPs required for our analysis were available in the imputed genotype dataset, which obtained through UK Biobank Axiom™ Array and UK BiLEVE Axiom™ Array[27]. More information on imputation of genotype data was described elsewhere[27]. We constructed PRS and individual SNP's score according to each SNP's risk alleles and log odds ratio (OR) associated with each SNP from GWAS summary statistics.

Covariates assessment

Variables included in the models were selected based on prior literature, biological plausibility, and known or hypothesized associations with IGFBP-7 levels, including demographic factors, genetic variants, lifestyle behaviors, and metabolic traits. Data on sex (female, male), race/ethnicity (white, non-white), alcohol intake (daily or almost daily, three or four times a week, once or twice a week, only on special occasions, never), smoking status (never, former, current), and physical activity (total weekly MET minutes for all activities) were collected through questionnaires. Fasting time refers to the duration between the last consumption of food or drink and the collection of blood samples. Each participant was assigned a Townsend Deprivation Index score, reflecting the socioeconomic status linked to their residential postcode.

Body mass index (BMI) was calculated from height and weight measured during visits to the assessment center. Household income before taxes was categorized as "Less than £18,000," "£18,000 to £30,999," "£31,000 to £51,999," "£52,000 to £100,000," or "Greater than £100,000." Healthy diet score ranging from 0 to 5 was consisted of five components: consuming at least 4 servings of vegetables daily, 3 servings of fruits daily, 2 servings of fish weekly, no more than 2 servings of processed meat weekly, and no more than 2 servings of unprocessed red meat weekly[32].

Statistical analysis

Cox proportional hazards models were initially applied to estimate hazard ratios (HRs) for the associations between IGFBP-7 level and each outcome, using age as the underlying timescale. The time scale began at the participant's recruitment age, with the exit point being the age at which the first of the following events occurred: study outcomes, death, or the end of the follow-up period (31 August 2022 for Scotland, 31 October 2022 for England, and 31 May 2022 for Wales). In UKB, the age at recruitment (field ID 21022) was recorded only as an integer. To obtain a more precise age estimate, we used participants' month of birth (field ID 52) and year of birth (field ID 34) to approximate their date of birth, setting it as the first day of their birth month and year. We then calculated a decimal-form recruitment age by determining the number of days between the approximate date of birth and the recruitment date (field ID 53), divided by 365.25. IGFBP-7 was analyzed as a

continuous variable, with each 1-standard deviation (SD) increase in its levels. For each outcome Cox modeling, two models were conducted with increasing numbers of covariates. Model 1 included adjustment for sex. Model 2 was further adjusted for fasting status, BMI, income, Townsend index, alcohol intake, smoking status, physical activity and health diet score. The proportional hazards assumption was evaluated based on Schoenfeld residuals. Because formal diagnostics for the fully adjusted age-scale models were numerically unstable in some analyses, reduced Cox models including IGFBP-7 and sex were additionally used for PH assessment.

We also conducted stratified analyses by age (<65 years, ≥65 years), sex (male, female) and follow up time (<5 years, ≥5 years; <10 years, ≥10 years), the interaction between stratified factor and IGFBP-7 was analyzed by including interaction terms. Several sensitivity analyses to estimate the robustness of associations between IGFBP-7 and diseases were conducted: (i) additional adjustment for IGFBP-4 and IGFBP-1L in the Cox models, as these showed the strongest correlations with IGFBP-7 among the available IGFBP family members in our dataset; (ii) adjusting for 20 available proteins from the ProtAge20 panel recently identified in a diverse-population cohort study[20].

The associations of IGFBP-7 PRS and individual SNP scores with overall cancer, specific-site cancers, CVD, T2D, and mortality were also analyzed based on Cox models while follow-up time was used as the timescale. Model 1 adjusted for age and sex, while Model 2 further adjusted for the top 10 genetic principal components. To

explore the measured correlates of circulating IGFBP-7, we examined the proportion of inter-individual variation in IGFBP-7 explained by selected demographic, genetic, lifestyle, biochemical, and dietary factors, including age, sex, BMI, income, smoking status, alcohol intake, Townsend index, race, summed MET, processed meat intake, red meat intake, vegetable intake, fish intake, fruit intake, and six individual SNP scores. For each variable, the explained variance (R^2) was derived from a separate linear regression model with IGFBP-7 as the dependent variable, and the overall R^2 was estimated from a model including all variables simultaneously. Moreover, the associations between IGFBP-7 and 30 blood biochemical markers, 20 aging-related biomarkers (ProtAge20) were also examined using spearman correlation. We calculated false discovery rate using the Benjamini–Hochberg method for exploratory analyses involving multiple comparisons, including correlations of IGFBP-7 with biochemical markers and ProtAge20 proteins, outcome-wide analyses of circulating IGFBP-7 and IGFBP-7 PRS, single-SNP analyses, and analyses of ProtAge20 proteins across outcomes. All statistical tests were two-sided, with a significance threshold set at $P < 0.05$. All analyses were conducted using R software.

Data availability

UKB data are available through a procedure described at https://www.ukbiobank.ac.uk/enable-your-research.

Code availability

All code employed in this study can be accessed by academic researchers upon request to the corresponding author.

Acknowledgements

We acknowledge the UKB participants for their dedication to participating in ongoing research and electronic health record linkage. All UKB data were accessed under UKB application no. 46466. This study was supported in part by grants to M.N.P. from the Terry Fox Foundation New Frontiers Program (Canada) and the Cancer Research Society (Montreal), and to M.S. from the National Institutes of Health (R01CA285851).

Author information

These authors contributed equally: Zhi Li, Chengzhe Tao and Ziyi Zhou.

Authors and Affiliations

Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA, USA.

Zhi Li, Edward L Giovannucci, Mingyang Song

State Key Laboratory of Reproductive Medicine, Center for Global Health, School of Public Health, Nanjing Medical University, Nanjing 211166, China.

Chengzhe Tao

School of Health and Wellbeing, University of Glasgow, United Kingdom.

Chengzhe Tao, Ziyi Zhou

McGill University, Montreal, Quebec, Canada.

Michael N Pollak

Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA.

Ziyi Zhou, Edward L Giovannucci, Mingyang Song

China CDC Key Laboratory of Environment and Population Health, National Institute of Environmental Health, Chinese Center for Disease Control and Prevention, Beijing, China.

Zhi Li

Contributions

Z.L., C.T., and Z.Z. contributed equally to this work and performed the main statistical analyses. Z.L. interpreted the data and drafted the initial manuscript. M.N.P., E.L.G., and M.S. obtained funding, contributed to study design, and supervised the overall project. Z.L., C.T., Z.Z., E.L.G., and M.S. revised the manuscript. M.S. supervised the overall study.

Corresponding author

Correspondence to Michael N Pollak, Edward L Giovannucci, Mingyang Song.

Ethics declarations

Competing interests

The authors declare no competing interests.

Reference:

1. Lit, K.K., Zhirenova, Z. & Blocki, A. Insulin-like growth factor-binding protein 7 (IGFBP7): A microenvironment-dependent regulator of angiogenesis and vascular remodeling. *Front Cell Dev Biol* **12**, 1421438 (2024).

2. Oh, Y., *et al.* Synthesis and characterization of insulin-like growth factor-binding protein (IGFBP)-7. Recombinant human mac25 protein specifically binds IGF-I and -II. *J Biol Chem* **271**, 30322-30325 (1996).
3. Pollak, M. The insulin and insulin-like growth factor receptor family in neoplasia: an update. *Nat Rev Cancer* **12**, 159-169 (2012).
4. Severino, V., *et al.* Insulin-like growth factor binding proteins 4 and 7 released by senescent cells promote premature senescence in mesenchymal stem cells. *Cell Death Dis* **4**, e911 (2013).
5. Chugh, S., *et al.* Pilot study identifying myosin heavy chain 7, desmin, insulin-like growth factor 7, and annexin A2 as circulating biomarkers of human heart failure. *Proteomics* **13**, 2324-2334 (2013).
6. Swisshelm, K., Ryan, K., Tsuchiya, K. & Sager, R. Enhanced expression of an insulin growth factor-like binding protein (mac25) in senescent human mammary epithelial cells and induced expression with retinoic acid. *Proc Natl Acad Sci U S A* **92**, 4472-4476 (1995).
7. Slater, T., Haywood, N.J., Matthews, C., Cheema, H. & Wheatcroft, S.B. Insulin-like growth factor binding proteins and angiogenesis: from cancer to cardiovascular disease. *Cytokine Growth Factor Rev* **46**, 28-35 (2019).
8. Burger, A.M., *et al.* Down-regulation of T1A12/mac25, a novel insulin-like growth factor binding protein related gene, is associated with disease progression in breast carcinomas. *Oncogene* **16**, 2459-2467 (1998).
9. Sprenger, C.C., Damon, S.E., Hwa, V., Rosenfeld, R.G. & Plymate, S.R. Insulin-like growth factor binding protein-related protein 1 (IGFBP-rP1) is a potential tumor suppressor protein for prostate cancer. *Cancer Res* **59**, 2370-2375 (1999).
10. Evdokimova, V., *et al.* IGFBP7 binds to the IGF-1 receptor and blocks its activation by insulin-like growth factors. *Sci Signal* **5**, ra92 (2012).
11. Xu, Y., Sun, Y., Zhu, Y. & Song, G. Structural insight into CD93 recognition by IGFBP7. *Structure* **32**, 282-291.e284 (2024).
12. Orenduff, M.C., *et al.* Plasma Insulin-like Growth Factor-Binding Protein-7 Is Positively Associated with Age, Obesity, Mortality, and Cancer in Postmenopausal Women. *Cancer Epidemiol Biomarkers Prev* **34**, 922-932 (2025).
13. Liu, Y., *et al.* Serum IGFBP7 levels associate with insulin resistance and the risk of metabolic syndrome in a Chinese population. *Sci Rep* **5**, 10227 (2015).
14. Zhang, L., *et al.* Insulin-like growth factor-binding protein-7 (IGFBP7) links senescence to heart failure. *Nat Cardiovasc Res* **1**, 1195-1214 (2022).
15. Godina, C., Pollak, M.N. & Jernström, H. Targeting IGF-IR improves neoadjuvant chemotherapy efficacy in breast cancers with low IGFBP7 expression. *NPJ Precis Oncol* **8**, 212 (2024).
16. Godina, C., *et al.* Genetic determinants and clinical significance of circulating and tumor-specific levels of insulin-like growth factor binding protein 7 (IGFBP7) in a Swedish breast cancer cohort. *Carcinogenesis* **46**(2025).
17. Katoh, M., *et al.* Vaccine Therapy for Heart Failure Targeting the Inflammatory Cytokine Igfbp7. *Circulation* **150**, 374-389 (2024).
18. Ferkingstad, E., *et al.* Large-scale integration of the plasma proteome with genetics and disease. *Nat Genet* **53**, 1712-1721 (2021).

19. Argentieri, M.A., *et al.* Integrating the environmental and genetic architectures of aging and mortality. *Nat Med* **31**, 1016-1025 (2025).
20. Argentieri, M.A., *et al.* Proteomic aging clock predicts mortality and risk of common age-related diseases in diverse populations. *Nat Med* **30**, 2450-2460 (2024).
21. Ko, T., *et al.* Cardiac fibroblasts regulate the development of heart failure via Htra3-TGF-β-IGFBP7 axis. *Nat Commun* **13**, 3275 (2022).
22. Zhao, R., *et al.* Plasma proteomics-based organ-specific aging for all-cause mortality and cause-specific mortality: a prospective cohort study. *Geroscience* **47**, 1411-1423 (2025).
23. Kivimäki, M., *et al.* Proteomic organ-specific ageing signatures and 20-year risk of age-related diseases: the Whitehall II observational cohort study. *Lancet Digit Health* **7**, e195-e204 (2025).
24. Robinson, O., *et al.* Associations of proteomic age with mortality and incident chronic diseases in the European Prospective Investigation into Cancer and Nutrition (EPIC). *Res Sq* (2025).
25. Goeminne, L.J.E., *et al.* Plasma protein-based organ-specific aging and mortality models unveil diseases as accelerated aging of organismal systems. *Cell Metab* **37**, 205-222.e206 (2025).
26. Sudlow, C., *et al.* UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* **12**, e1001779 (2015).
27. Bycroft, C., *et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203-209 (2018).
28. Sun, B.B., *et al.* Plasma proteomic associations with genetics and health in the UK Biobank. *Nature* **622**, 329-338 (2023).
29. Elliott, P. & Peakman, T.C. The UK Biobank sample handling and storage protocol for the collection, processing and archiving of human blood and urine. *Int J Epidemiol* **37**, 234-244 (2008).
30. Dhindsa, R.S., *et al.* Rare variant associations with plasma protein levels in the UK Biobank. *Nature* **622**, 339-347 (2023).
31. Titova, O.E., *et al.* Plasma proteome and incident myocardial infarction: sex-specific differences. *Eur Heart J* **45**, 4647-4657 (2024).
32. Peng, H., *et al.* Lifestyle Factors, Genetic Risk, and Cardiovascular Disease Risk among Breast Cancer Survivors: A Prospective Cohort Study in UK Biobank. *Nutrients* **15**(2023).

Table 1. Baseline characteristics of UK Biobank participants by Category of IGFBP-7 levels.

Characteristic	IGFBP-7 levels			
	Q1 (N=13,035)	**Q2 (N=13,320)**	**Q3 (N=12,751)**	**Q4 (N=13,033)**
Age (yrs.), mean (SD)	53.34 (8.27)	56.30 (8.13)	57.93 (7.75)	59.70 (7.31)
Sex, %				
Male	4,183 (32.1%)	5,970 (44.8%)	6,458 (50.6%)	7,412 (56.9%)
Female	8,852 (67.9%)	7,350 (55.2%)	6,293 (49.4%)	5,621 (43.1%)
Smoking status, %				
Never	7,469 (57.4%)	7,454 (56.0%)	6,825 (53.6%)	6,446 (49.5%)
Previous	4,214 (32.4%)	4,480 (33.7%)	4,514 (35.4%)	4,983 (38.3%)
Current	1,298 (10.0%)	1,324 (9.9%)	1,346 (10.6%)	1,532 (11.8%)
Do not know	42 (0.3%)	49 (0.4%)	49 (0.4%)	54 (0.4%)
BMI (mean (SD))	26.74 (4.18)	27.14 (4.40)	27.46 (4.68)	28.56 (5.63)
Household income, %				
Less than 18,000	2,161 (18.4%)	2,543 (21.3%)	2,602 (23.0%)	3,356 (29.4%)
18,000 to 30,999	2,593 (22.1%)	2,897 (24.3%)	2,942 (26.0%)	3,049 (26.7%)
31,000 to 51,999	3,131 (26.6%)	3,000 (25.2%)	2,650 (23.4%)	2,340 (20.5%)
52,000 to 100,000	2,599 (22.1%)	2,319 (19.4%)	1,952 (17.3%)	1,611 (14.1%)
Greater than 100,000	749 (6.4%)	573 (4.8%)	553 (4.9%)	408 (3.6%)
Missing	516 (4.4%)	593 (5.0%)	604 (5.3%)	663 (5.8%)
Townsend deprivation index, mean (SD)	-1.24 (3.15)	-1.26 (3.16)	-1.21 (3.17)	-1.00 (3.25)
Alcohol intake, %				
Daily or almost daily	2,532 (19.4%)	2,743 (20.6%)	2,554 (20.1%)	2,662 (20.5%)
Three or four times a week	3,219 (24.7%)	3,052 (22.9%)	2,915 (22.9%)	2,538 (19.5%)
Once or twice a week	3,522 (27.0%)	3,555 (26.7%)	3,239 (25.4%)	3,186 (24.5%)

One to three times a month	1,462 (11.2%)	1,396 (10.5%)	1,448 (11.4%)	1,383 (10.6%)
Special occasions only	1,342 (10.3%)	1,521 (11.4%)	1,483 (11.6%)	1,781 (13.7%)
Never	928 (7.1%)	1,030 (7.7%)	1,084 (8.5%)	1,442 (11.1%)
Do not know	18 (0.1%)	10 (0.1%)	11 (0.1%)	23 (0.2%)
Healthy eating index score, %				
0	218 (1.7%)	254 (2.0%)	272 (2.2%)	319 (2.6%)
1	1,257 (10.0%)	1,311 (10.3%)	1,342 (11.0%)	1,529 (12.4%)
2	2,964 (23.5%)	3,028 (23.8%)	2,861 (23.4%)	3,024 (24.5%)
3	3,774 (29.9%)	3,756 (29.5%)	3,516 (28.8%)	3,610 (29.2%)
4	3,117 (24.7%)	3,076 (24.1%)	2,925 (23.9%)	2,722 (22.0%)
5	1,290 (10.2%)	1,320 (10.4%)	1,307 (10.7%)	1,164 (9.4%)
Summed MET minutes per week for all activity, mean (SD)	42.60 (42.01)	44.26 (43.72)	45.54 (45.24)	42.48 (44.10)

Table 2. Sensitivity analyses for the association of IGFBP-7 with outcomes.

Outcome	Sensitivity analysis 1			Sensitivity analysis 2		
	HR (95%CI)	***P* value**	**FDR-adjusted *P* value**	**HR (95%CI)**	***P* value**	**FDR-adjusted *P* value**
Bladder cancer	1.13 (0.98-1.29)	8.15E-02	1.63E-01	1.21 (1.04-1.42)	1.51E-02	4.03E-02
Breast cancer	0.96 (0.88-1.05)	4.02E-01	5.49E-01	1.07 (0.94-1.21)	3.09E-01	5.24E-01
Chronic lymphocytic leukemia cancer	1.22 (0.99-1.51)	6.00E-02	1.29E-01	0.96 (0.64-1.42)	8.26E-01	8.93E-01
CNS cancer	0.80 (0.61-1.06)	1.27E-01	2.38E-01	0.94 (0.65-1.35)	7.34E-01	8.59E-01
Colorectal cancer	0.99 (0.88-1.12)	9.23E-01	9.57E-01	0.92 (0.77-1.09)	3.22E-01	5.31E-01
Esophagus cancer	1.12 (0.93-1.36)	2.39E-01	3.72E-01	0.98 (0.75-1.29)	8.85E-01	9.18E-01
Gallbladder cancer	0.96 (0.61-1.50)	8.48E-01	9.31E-01	1.08 (0.62-1.91)	7.78E-01	8.81E-01
Head and neck cancer	1.01 (0.80-1.27)	9.18E-01	9.57E-01	1.07 (0.81-1.42)	6.34E-01	8.20E-01

Kidney cancer	1.01 (0.82-1.25)	8.97E-01	9.57E-01	1.06 (0.82-1.38)	6.44E-01	8.20E-01
Liver cancer	1.54 (1.39-1.70)	1.04E-16	7.27E-16	1.36 (1.09-1.69)	6.07E-03	1.79E-02
Lung cancer	1.09 (0.98-1.23)	1.18E-01	2.28E-01	1.03 (0.88-1.20)	7.32E-01	8.59E-01
Non-Hodgkin lymphoma cancer	1.10 (0.95-1.28)	2.06E-01	3.50E-01	0.84 (0.65-1.08)	1.70E-01	3.37E-01
Melanoma cancer	0.93 (0.79-1.09)	3.90E-01	5.47E-01	0.91 (0.73-1.14)	4.24E-01	6.24E-01
Multiple myeloma cancer	1.40 (1.26-1.56)	3.72E-10	1.73E-09	1.36 (1.14-1.63)	6.85E-04	2.74E-03
Ovary cancer	0.94 (0.71-1.24)	6.41E-01	7.80E-01	0.91 (0.62-1.34)	6.29E-01	8.20E-01
Overall cancer	1.03 (1.00-1.07)	6.52E-02	1.35E-01	1.02 (0.97-1.07)	3.95E-01	5.97E-01
Pancreas cancer	0.85 (0.66-1.08)	1.85E-01	3.24E-01	0.84 (0.60-1.19)	3.32E-01	5.32E-01
Prostate cancer	1.00 (0.93-1.08)	9.73E-01	9.83E-01	1.01 (0.92-1.12)	7.87E-01	8.81E-01
Stomach cancer	1.12 (0.89-1.40)	3.22E-01	4.87E-01	1.09 (0.82-1.46)	5.50E-01	7.70E-01
Thyroid cancer	1.15 (0.84-1.58)	3.73E-01	5.40E-01	1.01 (0.59-1.73)	9.66E-01	9.83E-01
Uterus cancer	1.03 (0.82-1.28)	8.27E-01	9.28E-01	1.22 (0.91-1.63)	1.91E-01	3.57E-01
All-cause mortality	1.18 (1.14-1.22)	2.12E-23	1.70E-22	1.08 (1.03-1.13)	1.96E-03	6.86E-03
CVD cause mortality	1.15 (0.91-1.46)	2.30E-01	3.68E-01	1.22 (0.93-1.58)	1.45E-01	3.12E-01
Cancer cause mortality	1.06 (0.92-1.22)	4.32E-01	5.76E-01	1.02 (0.84-1.22)	8.66E-01	9.15E-01
Liver cause mortality	1.77 (1.58-1.97)	2.37E-24	2.21E-23	1.64 (1.35-2.00)	5.87E-07	3.29E-06
Kidney cause mortality	1.18 (0.76-1.82)	4.62E-01	6.02E-01	1.31 (0.80-2.14)	2.83E-01	4.95E-01
CVD	1.08 (1.04-1.12)	3.30E-05	1.09E-04	1.04 (0.98-1.09)	1.74E-01	3.37E-01
Diabetes	1.13 (1.09-1.18)	4.98E-09	2.14E-08	1.10 (1.04-1.17)	1.24E-03	4.61E-03

Notes: Sensitivity analysis 1 was adjusted by sex, fasting status, BMI, income, Townsend index, alcohol intake, smoking status, physical activity and health diet score, IGFBP-4, IGFBP-1L; sensitivity analysis 2 was further adjusted by sex, fasting status, BMI, income, Townsend index, alcohol intake, smoking status, physical activity and health diet score, 20 available proteins in protage20 (ACRV1, AGRP, CDCP1, COL6A3, CXCL17, EDA2R, ENG, ELN, GDF15, GFAP, FSHB, IGDCC4, KLK3, KLK7, LECT2, LTBP2, NEFL, PODXL2,PTPRR, SCARF2).

Figure 1. Overview of the study design and analytic approaches.

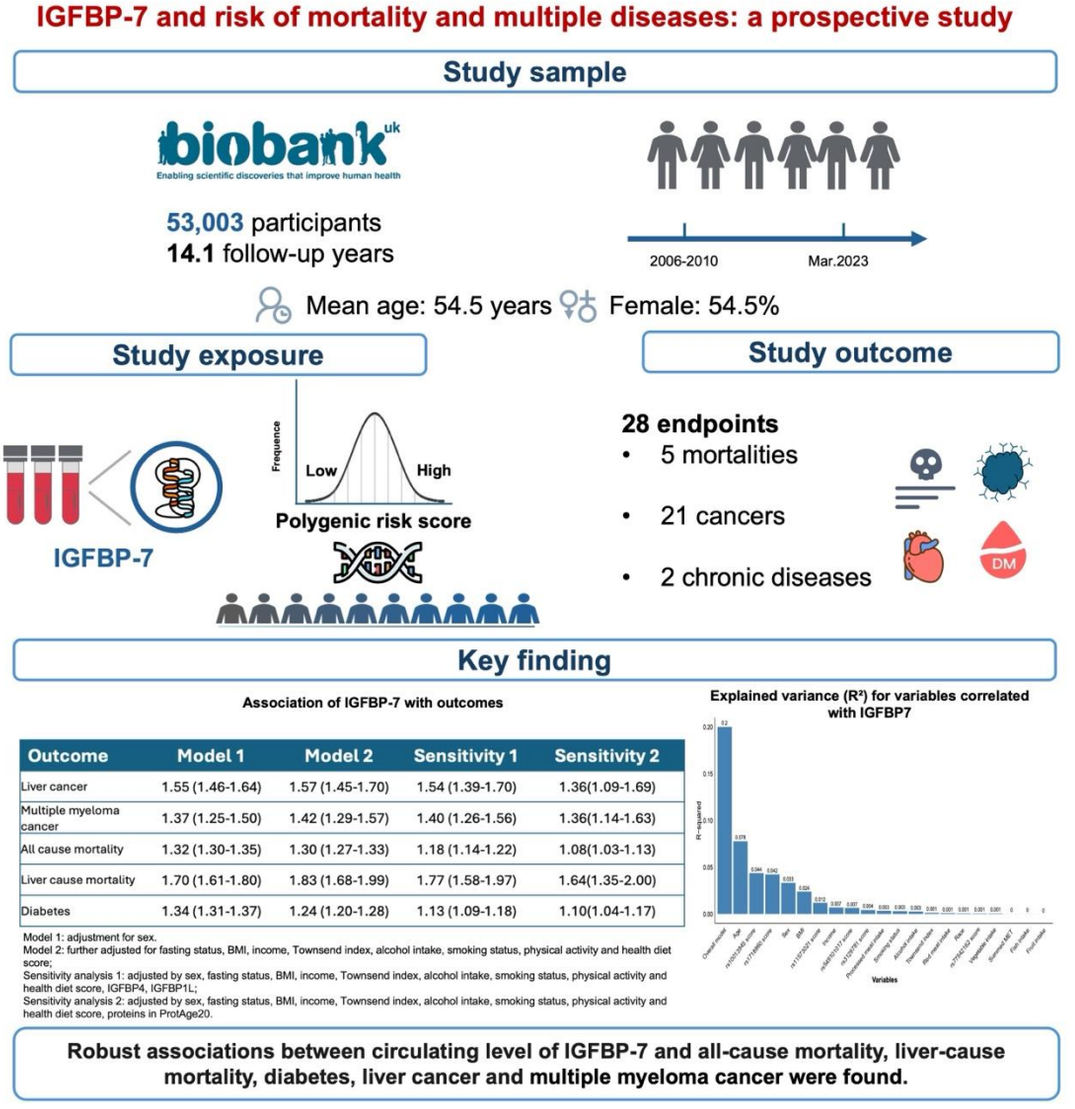
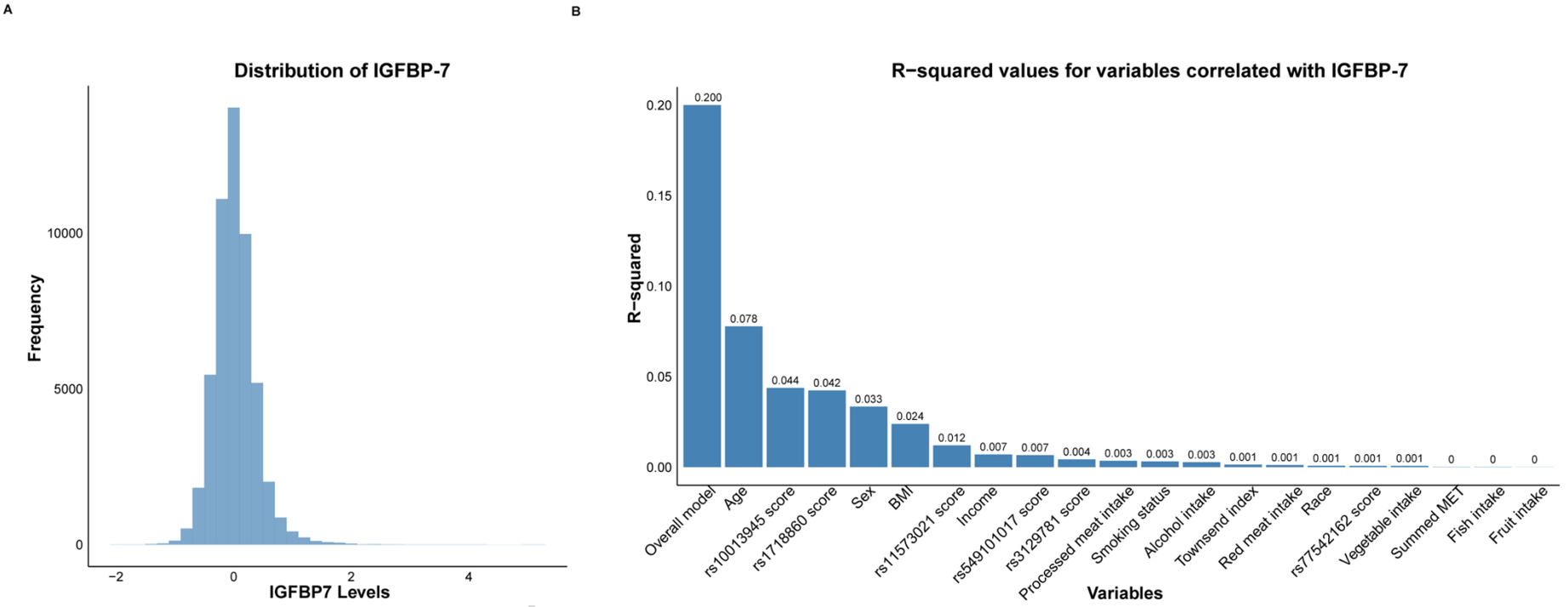
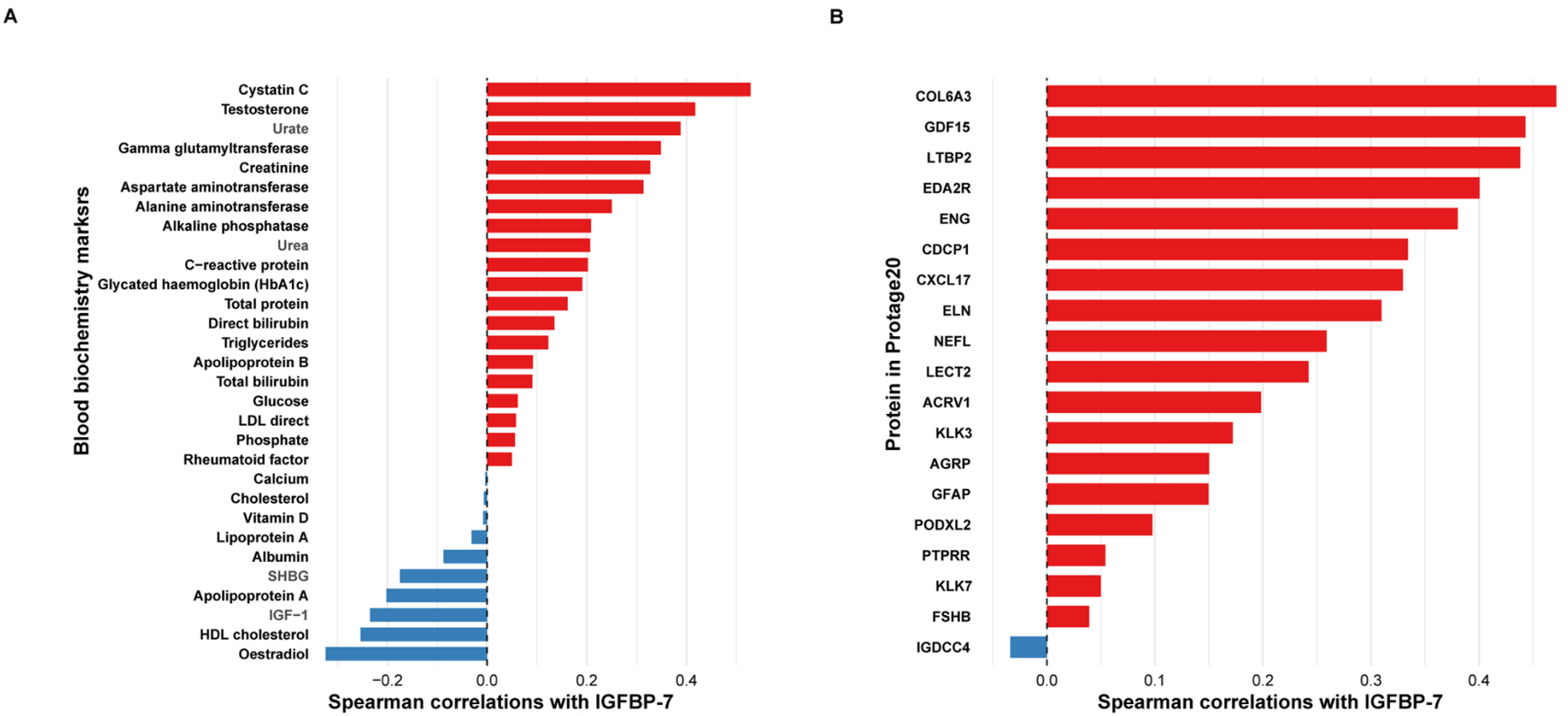


Figure 2. Distribution and determinants of IGFBP-7 levels in the UK Biobank.



A. Histogram showing the distribution of IGFBP-7 protein levels (standardized NPX values) among 53,003 participants.
B. R-squared values for individual predictors of IGFBP-7 levels from a multivariable linear regression model. The overall model explained 20.0% of the variance. The largest contributors included age (7.8%), genetic scores (rs10013945 and rs1718860), sex and BMI. Other variables such as income, alcohol intake, physical activity, and several biochemical and dietary factors contributed modestly.

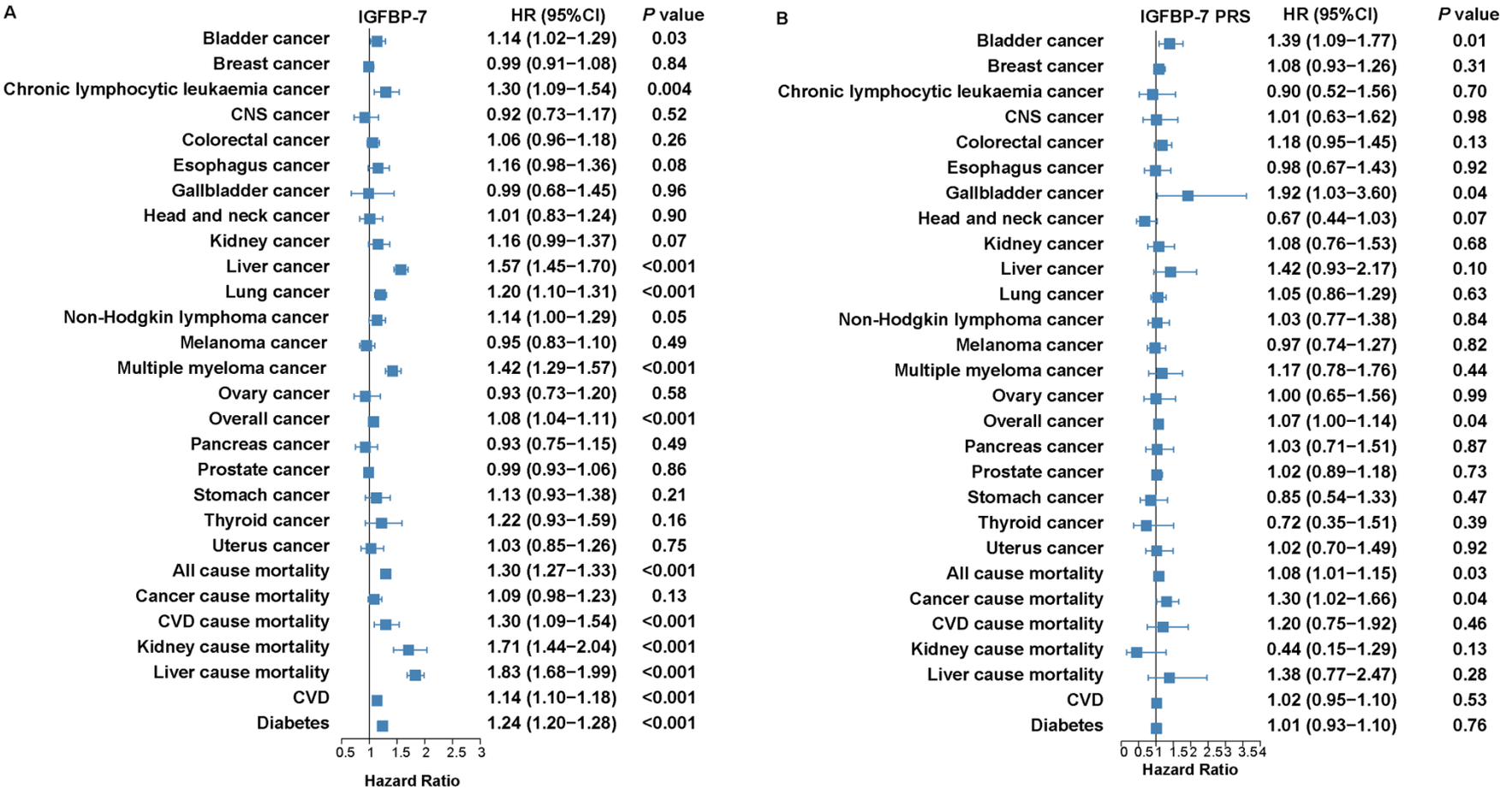
Figure 3. Correlations of IGFBP-7 with biochemistry markers and aging-related proteins.



A. Spearman correlations between IGFBP-7 and 30 circulating blood biochemistry markers in the UK Biobank. IGFBP-7 was most positively correlated with kidney function and inflammation-related markers, including cystatin C ($r^2 = 0.53$), testosterone, urate, and gamma glutamyltransferase. Negative correlations were observed with oestradiol, HDL cholesterol, IGF-1, and apolipoproteins.

B. Spearman correlations between IGFBP-7 and 20 age-associated proteins (ProtAge20). IGFBP-7 showed the strongest correlations with COL6A3, GDF15, and LTBP2. IGDCC4 was the only protein showing a weak inverse correlation.

Figure 4. Associations of IGFBP-7 levels and IGFBP-7 PRS with risk of mortality and disease outcomes.



A. Proteomic associations between standardized IGFBP-7 levels and 31 health outcomes in UK Biobank participants, including 22 site-specific cancers, overall cancer, five cause-specific mortalities, all-cause mortality, cardiovascular disease (CVD), and diabetes. Hazard ratios (HRs) and

95% confidence intervals were estimated using Cox proportional hazards models adjusted for age, sex, fasting status, BMI, income, Townsend index, alcohol intake, smoking status, physical activity, and diet quality.
B. Associations between IGFBP-7 polygenic risk score (PRS) and the same set of disease and mortality outcomes. IGFBP-7 PRS was significantly associated with overall cancer, bladder, gallbladder, and kidney cancers, as well as all-cause and cancer-specific mortality. Models were adjusted for age, sex, and the top genetic principal components.