Research Square

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

Biological Age Modeling Using Electronic Health Records

Smitha Edakalavan

sme57@pitt.edu

University of Pittsburgh

Alison Morris

Rutgers, The State University of New Jersey

Barbara Methe

University of Pittsburgh

Josiah Radder

University of Pittsburgh

Melanie Konigshoff

University of Pittsburgh

Seyed Mehdi Nouraie

University of Pittsburgh

Article

Keywords:

Posted Date: September 21st, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10464023/v1

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

Additional Declarations: No competing interests reported.

Abstract

Background

Chronological age poorly captures individual differences in physiologic health across the lifespan. Although biological age has shown promise as a more informative measure of health status, most studies have relied on curated research cohorts rather than routinely collected electronic health record (EHR) data. Among clinical biological measures, Levine PhenoAge and Klemera-Doubal Method (KDM), which estimate biological age from clinical biomarkers, are among the most widely used. However, KDM, in particular, has been rarely applied outside large biobanks or narrow disease-specific hospital contexts, and direct comparisons of KDM with PhenoAge using identical biomarkers in the same real-world population remain scarce.

Objective

To evaluate the feasibility and clinical relevance of estimating biological age from EHR data by applying Levine PhenoAge and the Klemera-Doubal Method (KDM) to a real-world, multi-disease patient population, including a head-to-head comparison of the two measures, and to evaluate their associations with mortality, morbidity, frailty, and hospital readmission.

Methods

We conducted a retrospective study of 93,200 encounters from 2,434 unique patients (ages 30–75) at a large health system in western Pennsylvania, with a median follow-up of 10.2 years. Levine PhenoAge and KDM biological age were calculated using a nine-biomarker panel, and age acceleration was derived as the residual of biological age regressed on chronological age. Associations between biological age acceleration and clinical outcomes, including mortality, comorbidities, frailty, and 30-day hospital readmission, were evaluated using Cox proportional hazards, logistic, and negative binomial regression models.

Results

The cohort had a mean age of 50.6 years (59.6% female), with a mean PhenoAge of 54.3 years (SD 17.8 years) and a mean KDM age of 50.8 years (SD 21.45 years), both of which were strongly correlated with chronological age ($r = 0.75$ and 0.638, respectively) and with each other ($r = 0.78$). Chronological age alone was not significantly associated with 30-day readmission rate ($p = 0.78$), whereas both PhenoAge and KDM acceleration were associated with increased readmission rate(17.6% and 5.8% per 5-year increase, respectively), indicating that biological age acceleration captures readmission risk that chronological age misses entirely. PhenoAge and KDM acceleration were each independently associated

with mortality beyond chronological age (HR = 1.17 and 1.08 per 5-year increase), significantly improving the model fit over age and sex alone (LRT χ^2 =62.99 and 30.90, respectively, $p < 0.001$). Both PhenoAge and KDM acceleration were independently associated with frailty beyond chronological age, with significant improvement in model fit over an age- and sex-only model (LRT χ^2 =25.80 and 20.00, respectively, $p < 0.005$). PhenoAge and KDM age acceleration were also strongly associated with numerous comorbidities, most prominently renal conditions (chronic kidney disease, hyperkalemia), after correction for multiple comparisons. Longitudinal tertile transition analysis showed that male sex predicted both a greater likelihood of transitioning into the highest acceleration tertile from the lowest acceleration tertile (OR = 1.86) and a lower likelihood of transitioning out of the highest acceleration tertile (OR = 0.56), indicating that males tend toward, and persist within, faster-aging trajectories over time.

Conclusions

This study demonstrates that KDM can be reliably applied to heterogeneous real-world EHR data and directly compared with PhenoAge using an identical biomarker panel within the same clinical population. To our knowledge, KDM has previously been applied primarily within large biobanks or narrow disease-specific cohorts. This study extends KDM to a heterogeneous, multi-disease clinical population and is the first to link biological age acceleration to 30-day hospital readmission across a broad patient population. These findings support the potential utility of EHR-derived biological age as a clinical biomarker of risk stratification.

Introduction

Chronological age is one of the strongest predictors of morbidity, functional decline, and mortality, but individuals of the same chronological age often differ substantially in their physiologic health and risk of adverse outcomes. This heterogeneity has motivated growing interest in biological age, which measures the physiological stage of tissues and organs rather than just the number of years lived. Importantly, unlike chronological age, biological age may be modifiable and partially reversible, making it a potentially actionable marker for health(Fitzgerald et al., 2020; Lu, no date). Environmental and lifestyle factors such as race, socioeconomic deprivation, smoking, and calorie restriction impact biological aging(Harris et al., 2024) with genetic factors responsible for ~ 30% of the change in biological aging(Deelen et al., 2013; Khan, Singer, and Vaughan, 2017).

Biological aging clocks span epigenetic (DNA methylation), transcriptomic, proteomic, and metabolomic, as well as biological clocks derived from organs such as lungs and heart, and clocks based on clinical and functional measures. Clinical clocks are widely used since they are based on routinely collected clinical data, making them readily available and less expensive compared to other types of biological clocks. Two of the most used clinical clocks are Levine PhenoAge(Levine et al., 2018; Liu et al., 2018) and the Klemera-Doubal Method (KDM) (Klemera and Doubal, 2006). Levine PhenoAge was developed as

a mortality-informed phenotypic age measure derived from blood biomarkers and chronological age, with the goal of better capturing mortality and morbidity risk. In contrast, KDM estimates biological age by combining information across multiple biomarkers in relation to chronological age, providing a weighted biological age estimate. While Levine PhenoAge calculation requires a fixed set of biomarkers to be passed, KDM is flexible and can be adapted to any available biomarker sets.

Electronic health records (EHRs) provide a wealth of longitudinal clinical data for biological age modeling. At the individual level, biological age provides prognostic information beyond the calendar age, offering a personalized picture of health (Lu, 2024). Biological age estimates can help clinicians identify patients at higher risk for age-related conditions to guide personalized treatment decisions. At the hospital level, it may support resource allocation and reduce readmissions by identifying vulnerable patients early (Jawad et al., 2026).

While PhenoAge has been previously applied in several clinical settings, including ICU cohorts (Ho, 2023a), disease-specific cohorts such as acute ischemic stroke (Liu et al., 2026), acute myocardial infarction and heart failure (Gao et al., 2025), and COVID-19 (Strazhesko et al., 2022), and a heterogeneous EHR population examining social vulnerabilities (Knobel et al., 2025), applications of KDM to clinical hospital databases have been limited to disease-specific contexts such as COVID-19 (Strazhesko et al., 2022) and psoriasis (Lin et al., 2025). Studies have not estimated KDM biological age from biomarkers extracted from routine EHR workflows across a heterogeneous, multi-disease clinical population, where laboratory values are ordered for clinical rather than research purposes. EHR data present unique challenges, and our study addresses this gap by developing a pipeline for biological age estimation from EHR data and evaluating the resulting measures across multiple outcomes. Such measures have the potential to identify individuals experiencing accelerated biological aging who may be at increased risk for adverse clinical outcomes, using the data that are already routinely collected in clinical care. Specifically, we examined the relationship between biological age and chronological age, quantified biological age acceleration, and evaluated associations with comorbidities, mortality, and hospital readmission.

Methods

Study Design and Population

We conducted a retrospective study using EHR data from participants enrolled in the MedBio project(MedBio Project or University of Pittsburgh, 2017) (Li et al., 2022), an observational study at the University of Pittsburgh that links biological specimens with longitudinal clinical records from one of the largest hospital systems in western Pennsylvania. The MedBio project enrolled participants through a network of investigators across diverse clinical specialties, and, therefore, the resulting cohort reflects a broad spectrum of conditions including chronic pulmonary, renal, inflammatory, autoimmune, and oncologic diseases. The study population included 2,678 participants with inpatient and outpatient encounters, for whom the laboratory values were recorded between 2010 and 2025. Biological age value

was estimated for each encounter using laboratory values drawn at the visit, and all available encounters per patient were included in the longitudinal analyses.

Biological Age Estimation

BioAge toolkit (https://github.com/dayoonkwon/BioAge)(Kwon and Belsky, 2021) was used to calculate Levine PhenoAge (Levine et al., 2018) and KDM (Klemera and Doubal, 2006) biological age. PhenoAge is a mortality-based (Cox-proportional hazard) biological age estimator derived from nine laboratory markers (albumin, C-reactive protein (CRP), glucose, creatinine, lymphocyte %, white blood cell (WBC) count, alkaline phosphatase (ALP), mean corpuscular volume (MCV), and red blood cell distribution width (RDW), adjusted for chronological age to approximate mortality and morbidity risk. The resulting linear predictor is then transformed into an age-equivalent metric by mapping the calculated mortality risk into a Gompertz mortality curve, yielding a biological age in years. KDM is derived from a set of regression models in which each biomarker is regressed on chronological age to estimate the expected biomarker value as a function of age, and these are then used to compute a weighted estimate of biological age. As a flexible framework, KDM is not limited to a fixed set of biomarkers and can be applied using any suitable biomarker panel. To facilitate a direct methodological comparison, the same nine-biomarker panel used for PhenoAge was applied to the KDM framework as well.

Data Processing and Harmonization

Laboratory records were linked to encounter records, and laboratory results with preliminary or incomplete status codes were removed to ensure inclusion of finalized measurements only. Analyses were restricted to specimens collected from blood-derived sources, including peripheral blood, whole blood, serum, and plasma. Within each laboratory measure, results were sometimes reported using different units or measurement conventions, which were harmonized to a common standardized scale before analysis. To exclude the physiologically implausible values, lab measurements falling outside the 1st and 99th percentiles were removed. Laboratory component names varied across the EHR, and to ensure consistency, multiple synonymous laboratory identifiers were mapped to standardized variables as part of data cleansing. The laboratory values were then harmonized into units specified in PhenoAge and the KDM algorithm. After standardization, multiple measurements for laboratory tests within each encounter were collapsed to obtain a single mean value per encounter. For longitudinal analysis, laboratory values were extracted from each encounter independently, allowing biological age to be estimated at multiple time points per patient. For cross-sectional analyses, the clinical data from the first encounter were used. Diagnoses were treated as cumulative, with a condition considered present if it appeared in any encounter record before or in the current encounter. Demographic variables extracted from EHR include age and gender.

In addition to laboratory and demographic data, diagnosis information was also extracted from the EHR for phenotype characterization and comorbidity assessment. Individual clinical conditions are often

represented by multiple ICD codes, making direct interpretation of these codes difficult due to the high dimensionality and variability of coding practices. To address this, ICD codes were mapped to Phecodes, which group related diagnoses into standardized and clinically meaningful phenotype categories while reducing dimensionality. We used PheMap (Zheng et al., 2020), a high-throughput phenotyping framework, to automate the identification of clinical phenotypes from EHR using ICD-based mappings. Using these mappings, encounter-level phenotype matrices were generated in which each row represents a patient encounter, and each column corresponds to a clinical prototype.

Handling Missing Data

The PhenoAge calculation required nine biomarkers; however, since many of these lab tests are not routinely collected during hospital visits, missingness was observed for all the biomarkers (Table 1). To address the missingness, we followed a 2-step process: (1) impute the missing data using the mean of the available measurements within a 30-day forward and backward window, and (2) impute remaining missing values using multivariate chained equations (MICE)(Zhang, 2016). MICE is an iterative method for handling missing data by modeling each variable with missing values as a function of other variables, offering a significant advantage over other imputation techniques(Azur et al., 2011). The distribution of all the biomarker values was verified after MICE imputation to ensure that MICE did not alter the distribution of these variables. Distributional similarity between the observed and imputed biomarker values was additionally assessed using density plots provided in the Supplementary Material.

PhenoAge and KDM Age Calculation

Levine PhenoAge and KDM biological age were calculated using the nine biomarkers, and because the BioAge package functions require biomarker inputs to follow specific naming conventions and units, laboratory variables were renamed and harmonized before analysis (Table 1). Both PhenoAge and KDM used the same units, except for creatinine. Creatinine was log-transformed after conversion from umol/L to mg/dL to match the NHANES-trained KDM biomarker specification provided in the BioAge package. PhenoAge was estimated using the NHANES-trained PhenoAge model. Although KDM is conventionally recalibrated separately by sex, a combined-cohort model was used here to maintain methodological consistency with PhenoAge, which applies a single model irrespective of sex. Following the biological age calculation, age acceleration was derived as residuals from linear regression models of biological age on chronological age. To reduce the influence of repeated measurements within the individuals, the regression models were fit using the first encounter per patient, and the resulting coefficients were subsequently applied to all encounters. These residuals represent the deviation of biological age from the age expected based on chronological age. Positive residuals indicate accelerated aging, where biological age exceeds the expected value, while negative residuals indicate decelerated biological aging.

Statistical Analysis

Following the estimation of biological age using Levine PhenoAge and KDM, several analyses were conducted to evaluate the clinical relevance of biological age acceleration. Specifically, we examined the cross-sectional and longitudinal association between biological age acceleration and comorbidities, mortality, frailty, and hospital readmissions.

For the comorbidity analyses, age acceleration was evaluated using the first encounter per patient to capture the baseline biological age acceleration. Conditions were defined using the PheCode mapping described in the previous section, and age acceleration was compared between individuals with and without the condition using linear regression models adjusted for sex. Only those conditions where the number of patients with the condition was greater than 50 were included. For each condition, the reference group consisted of patients without that specific condition, but who may have had other comorbidities present. To account for multiple hypothesis testing across comorbidities, p-values were adjusted using Benjamini–Hochberg (BH) false discovery rate (FDR) procedure, and associations with FDR-adjusted p-value < 0.05 were considered statistically significant. The effect size was defined as the mean difference in age acceleration (with the condition minus without the condition), and p-values were derived accordingly. Results were visualized using volcano plots, with effect size on the X-axis and -log(p-value) on the Y-axis, for both PhenoAge and KDM.

The association of biological age acceleration with mortality was assessed using Cox proportional hazards analysis. All longitudinal encounters were included in this analysis, with survival time measured as the time since the first biological age calculation and the time to death or the time to last known follow-up. A baseline Cox model was fitted with chronological age and sex as predictors of mortality. Two additional models were then constructed by adding PhenoAge acceleration and KDM acceleration to the reference model, allowing assessment of each measure's incremental contribution beyond chronological age. Hazard ratios with 95% confidence intervals were reported. The proportional hazards assumption was evaluated using the Schoenfeld residual-based test. Potential nonlinearity in the association between biological age acceleration and mortality was evaluated using a restricted cubic spline (RCS). RCS are flexible, piecewise cubic polynomial functions used to relax the assumption of linearity between a continuous variable and outcome. Model fit was compared between the linear and spline Cox models using a likelihood ratio test.

We also calculated the frailty index of each patient at the first encounter by summing the number of health deficits identified from their EHR data and dividing by the total number of deficits assessed. Deficits include comorbidities, functional impairments, cognitive disorders, and other clinical conditions identified using ICD-code mappings derived from the eFRAGICAP frailty framework (Orfila et al., 2022). A deficit was considered present if at least one corresponding diagnosis code was identified in the patient's problem list.

$$Fraility\ Score = \frac{Number\ of\ deficits\ present}{Total\ number\ of\ deficits\ assessed}$$

We first converted the frailty score into a binary outcome (fit (frailty score < = 0.12) vs non-fit, where non-fit included mild, moderate, and severe frailty). We then built a set of nested logistic regression models: a baseline model assessing the association between chronological age and frailty status adjusted for sex, followed by two extended models that separately added PhenoAge acceleration and KDM acceleration to the baseline model.

Additionally, the association between biological age acceleration and hospital readmission burden was evaluated using negative binomial regression, restricted to inpatient encounters only. The index encounter was defined as the first inpatient encounter per patient, and the outcome was defined as the number of hospital readmissions occurring within 30 days following the index encounter. Patients with no readmissions within 30 days were retained in the analysis and contributed zero counts to the outcome. Similar to mortality and frailty analysis, a nested modeling approach was used. A baseline model including chronological age and sex was first fitted, followed by two extended models that separately added PhenoAge and KDM acceleration to assess their incremental contribution beyond chronological age and sex. Biological age acceleration measured at the index encounter was used as the predictor, and the results were expressed in incidence rate ratios (IRRs) per 5-year increase in biological age acceleration.

Finally, we also compared the groups of patients, classifying them into tertiles of PhenoAge acceleration (bottom, middle, and top) based on the distribution of acceleration values at the first encounter. The resulting baseline-driven tertile cutoffs were then applied to follow-up visits to classify each patient's tertile membership at each encounter. We then compared the age, sex, and biomarker distribution between patients in the highest and lowest baseline tertiles. In addition, to assess the longitudinal stability of biological aging status over time, we examined transitions in PhenoAge acceleration tertile membership from baseline to follow-up visits.

Sensitivity Analysis

To assess the impact of the missing laboratory data on biological age estimation and downstream analysis, we compared the results obtained using 1) complete-case analysis, 2) a ± 30-day laboratory imputation window approach, and 3) the MICE imputation approach, which have been used in this study. Biological age estimates and distributions were compared across the three approaches. We then re-evaluated the association between biological age acceleration and mortality risk using the Cox proportional hazard model under each missing-data approach, comparing hazard ratio estimates and model discrimination across strategies.

All statistical analyses were done using Python (version 3.13) and R (version 4.5.2).

Results

We first assessed the internal relationships among chronological age, Levine PhenoAge, and KDM Age. The dataset initially consisted of 2678 unique patients. Following biological age estimation, patients with infinite PhenoAge values resulting from extreme biomarker combinations were excluded from the analysis, yielding a final dataset of 2434 unique patients and 93200 encounters. The cohort had a mean age of 50.6 years and was predominantly female (59.6% female, 40.3% male). Missingness varied across various biomarkers, ranging from under 1% for WBC to 44.1% for CRP, reflecting its less frequent use in routine clinical ordering compared to other biomarkers (Table 1).

Table 1
Distribution of lab values and missing percentages

Biomarker	Missing %	Mean (SD)	Median (IQR)
Albumin (g/L)	3.5	39.53 (5.4)	40.0 (37.0–43.0)
Creatinine (umol/L)	1.3	90.12 (78.5)	78.68 (62.76–91.35)
Glucose (mmol/L)	1.8	6.08 (2.41)	5.44 (4.94–6.33)
CRP (mg/dL)	44.1	2.48 (4.77)	0.63 (0.20–2.00)
MCV (fl)	1.5	90.09 (6.26)	90.10 (86.40–93.70)
RDW (%)	2.0	14.37 (2.05)	13.80 (13.20–14.85)
ALP (u/L)	3.7	84.97 (69.03)	71.00 (57.00–91.00)
WBC (1000 cells/uL)	0.9	7.52 (3.04)	6.90 (5.49–8.80)
Lymphocyte (%)	2.1	24.25 (10.8)	24.00 (17.00–31.00)

PhenoAge and KDMAge were both positively correlated with chronological age, with PhenoAge exhibiting a stronger correlation (Spearman $r = 0.755$, $p < 0.001$) and KDM demonstrating a moderate correlation with chronological age ($r = 0.638$, $p < 0.001$) (Fig. 1).

PhenoAge and KDMAge were also positively correlated with each other (Spearman $r = 0.783$, $p < 0.001$), however moderate-to-strong correlation suggests that despite being derived from identical biomarkers, the two measures capture overlapping but not identical dimensions of biological aging. The scatterplot demonstrated approximately linear correlation among all the measures. Against a mean chronological age of 50.6 years, Levine PhenoAge had a mean (SD) of 54.27 years (17.83 years) compared with 50.68 (21.45) years for KDM age.

To further characterize the clinical significance of biological age beyond overall correlations, we compared the patients at the extremes of the biological age acceleration distribution. This analysis was designed to determine whether individuals with markedly different biological age acceleration also differed in their demographic characteristics, biomarker profiles, and comorbidity burden despite having similar chronological ages. Although the two groups had comparable ages(mean age 57.5 vs 55.5

years), they exhibited substantially different biological age acceleration (mean PhenoAge acceleration: 17.2 years vs 10.7 years). The highest tertile had a higher proportion of males (46.1% vs 30.8%) and a greater comorbidity burden. Compared with the lowest tertile, patients in the highest tertile had lower albumin and lymphocyte percentage and higher glucose, RDW, CRP, ALP, and WBC after adjustment for chronological age and sex (all $p < 0.001$), whereas MCV did not differ significantly ($p = 0.09$).

The availability of repeated measurements over a median follow-up of 10.2 years enabled assessment of longitudinal changes in biological age acceleration at the individual level. During the follow-up, 57.49% of patients changed PhenoAge acceleration tertile membership between baseline and the final available follow-up (Fig. 2), demonstrating that biological age acceleration is dynamic rather than static over time. Although patients in the highest acceleration tertile at baseline were most likely to remain in the same tertile (56.2%), substantial movement between tertiles was observed over time. Only 40.1% patients in the lowest tertile and 31.3% of patients in the middle tertile remained in their original categories. Among patients in the middle tertile, transitions to the highest tertile were the most common (40.8%), whereas patients in the lowest tertile were redistributed almost equally between the middle (30.6%) and highest (29.3%) tertiles at follow-up.

To characterize the predictors of PhenoAge acceleration tertile transitions, we fit binary logistic regression models separately for patients starting in the bottom and top tertiles at baseline, with sex and age as predictors. Among patients starting in the bottom tertile (slowest biological aging) at baseline, males were significantly associated with higher odds of transitioning into a higher tertile at follow-up ($OR = 1.86$, $p < 0.001$), while age was not significantly associated with transition ($OR = 1.01$, $p = 0.081$). Among patients starting in the top tertile (fastest biological aging) at baseline, males were significantly associated with lower odds of transitioning into lower tertile at follow-up ($OR = 0.56$, $p < 0.001$), indicating that male patients were more likely to remain in the fastest-aging tertile. Older age was also significantly associated with lower odds of transitioning out of the top tertile ($OR = 0.98$, $p < 0.001$). Together, these findings suggest that male sex was consistently associated with tendency toward faster biological aging on follow-up and decreasing likelihood of moving out of the fastest-aging tertile.

Having characterized the longitudinal stability of biological aging status, we next examined the overall distribution of biological age acceleration across the cohort, stratified by sex, to determine whether biological age acceleration differed between males and females (Fig. 3). The distributions of both PhenoAge and KDM age acceleration were centered near zero. Females tended to have slightly lower biological age acceleration, whereas males showed a modest rightward shift towards positive acceleration values. The majority of patients clustered within ± 10 years of zero acceleration, with a right tail reflecting subsets of patients with elevated biological age. Overall, sex accounted for only a modest shift in biological age acceleration, while the substantial within-sex spread suggests that rather than sex, individual physiological variation is the primary driver of biological age acceleration in this cohort.

Comorbidities Associated with Age Acceleration

We next sought to identify comorbidities associated with biological age acceleration, to determine which clinical conditions may be driving the individual-level variation observed above, beyond what sex alone explains. These comorbidity analyses, adjusted for sex and visualized using a lollipop plot, revealed several significant associations between biological age acceleration and comorbidities (Fig. 4). Of the 542 conditions identified, 178 met the inclusion threshold of more than 50 patients with the condition being analyzed and were carried forward for analysis. The strongest positive associations with PhenoAge acceleration (Fig. 4a) were predominantly renal conditions (CKD, hyperkalemia, proteinuria), followed by respiratory conditions (respiratory failure, pneumonia) and cardiometabolic conditions (heart failure, type 2 diabetes, hypertension). None of the conditions were associated with negative age acceleration after correcting for multiple comparisons using the Benjamini-Hochberg false discovery rate procedure. The full list of these conditions is provided in Supplementary Materials.

Similar analysis was conducted using KDM Age acceleration (Fig. 4b). The strongest associations were again predominantly renal conditions, followed by a broader range of cardiometabolic conditions. Like PhenoAge, none of the conditions were associated with negative age acceleration after Benjamini-Hochberg correction.

While these cross-sectional analyses identify comorbidities associated with the magnitude of biological age acceleration at a single time point, they do not address whether these conditions also explain why acceleration status changes over time. To investigate this, we returned to the tertile transition analysis introduced in Fig. 2. Among patients who started at the top (fastest aging) tertile, those who remained in the top tertile at follow-up had significantly higher rates of several comorbidities compared to those who moved to the bottom tertile, including CKD (26.1% vs 4.5%, $p < 0.001$), type 2 diabetes (36.0% vs 15.4%, $p < 0.001$), and heart failure (18.7% vs 5.1%, $p < 0.001$).

Longitudinal Trajectories of Biological Age Acceleration

Having identified conditions cross-sectionally and longitudinally associated with biological age acceleration, we next examined whether these associations were also reflected in each patient's individual trajectory over time. We selected CKD and type 2 diabetes for longitudinal trajectory as representative chronic conditions from renal and metabolic categories (from Fig. 4) due to their gradual clinical course, and modeled their longitudinal PhenoAge acceleration trajectories (Fig. 5). By aligning patients at their first visit, we visualize the progressive divergence in PhenoAge acceleration between the diseased and non-diseased cohorts. Many patients classified as having a condition were likely already diagnosed with the disease before the baseline, and the trajectories here represent prevalence rather than incident disease status. Patients with CKD and type 2 diabetes (red trajectories) exhibited a higher age acceleration over time than patients without these conditions (black trajectories).

Association of Biological Ages and Calendar Age with Mortality

Having established that biological age acceleration is associated with increased burden of comorbidity, we next assessed whether this sustained acceleration translates into an increased risk of mortality. The result of the nested Cox proportional hazard model evaluating the associations of biological age accelerations and calendar age with mortality is provided in Table 4.

Table 4
Results of the Cox proportional hazard model

Age	HR per 5-year increase 95% CI	p-value	c-statistics
Chronological Age + Sex	1.522[1.441–1.624]	< 0.05	0.720
Chronological Age + Sex + PhenoAge Acceleration	1.165 [1.128–1.210]	< 0.05	0.745
Chronological Age + Sex + KDM Acceleration	1.084 [1.055–1.113]	< 0.05	0.732

All three measures were significantly associated with mortality ($p < 0.05$). Chronological age was significantly associated with mortality, with each 5-year increase associated with a 52.2% increase in the hazard ratio of death. Adding PhenoAge acceleration to the age- and sex-adjusted model significantly improved the model fit (LRT $\chi^2 = 62.99$, $p < 0.001$), each 5-year increase in PhenoAge acceleration was associated with 16.5% increase in mortality risk. Similarly, the addition of KDM acceleration improved the model fit (LRT $\chi^2 = 30.90$, $p < 0.001$), with each 5-year increase in KDM acceleration associated with 8.4% increase in mortality risk. Discrimination (c-index) improved only modestly with the addition of either measure (C-index 0.74 for PhenoAge, 0.73 for KDM, and 0.72 for chronological age), reflecting that age already orders the patient pairs correctly by risk, whereas acceleration's contribution lies in refining the individual risk estimates. Restricted cubic spline analysis demonstrated no significant departure from linearity for PhenoAge acceleration (LR $\chi^2 = 2.76$, $p = 0.25$) while KDM acceleration showed modest evidence of a nonlinear association with mortality (LR $\chi^2 = 6.50$, $p = 0.039$). As shown in Fig. 6, the mortality risk remained close to the reference line at negative age acceleration and increased steadily with increasing biological age acceleration.

Association of PhenoAge Acceleration with Frailty Index

We next asked whether biological age acceleration, having predicted mortality better than chronological age, was similarly associated with frailty. The frailty index categorized patients as fit (68.5%) and non-fit (mild, moderate, and severe frailty, 31.5%). In an age and sex only logistic regression model, chronological age was significantly associated with frailty status, with each 5-year increase in chronological age associated with a 35% increase in the odds of non-fit status (OR = 1.350, $p < 0.001$). To determine whether biological acceleration provided information beyond chronological age, we compared the nested logistic regression models predicting frailty status. Adding PhenoAge acceleration to a model

containing age and sex significantly improved model fit (likelihood ratio test (LRT) χ^2 = 25.80, p < 0.005), as did adding KDM acceleration (likelihood ratio test (LRT) χ^2 = 20.00, p < 0.005). A 5-year increase in PhenoAge acceleration was associated with 12% higher odds of being fit vs non-fit (OR = 1.120, p < 0.001), and a 5-year increase in KDM acceleration was associated with 7.6% higher odds (OR = 1.076, p < 0.001). The larger LRT for PhenoAge acceleration relative to KDM acceleration suggests that PhenoAge acceleration captures a stronger frailty-related signal beyond chronological age. Although the effect size for chronological age was larger than for biological age acceleration, the significant improvement in model fit (via LRT) demonstrates that age acceleration captures frailty risk not explained by chronological age alone.

Association of PhenoAge Acceleration with Hospital Readmission

Having shown that biological age acceleration is associated with both mortality and frailty, we next assessed whether it also predicted short-term healthcare utilization, specifically the burden of hospital readmission. During a 30-day follow-up period after the first encounter, the majority of the patients (95%) had two or fewer readmissions, with a highly right-skewed distribution reflecting a large proportion of patients with no readmissions. To evaluate the contribution of biological age acceleration to readmission risk beyond chronological age, we compared the nested negative binomial regression models predicting hospital readmissions within 30 days. In an age- and sex-only negative binomial model, age was not significantly associated with hospital readmission (Incidence Rate Ratio (IRR)per 5-year increase = 1.005, p = 0.78). Addition of PhenoAge acceleration improved the model-fit (LRT χ^2 = 98.6, p < 0.001), with each 5-year increase in PhenoAge acceleration associated with 17.6% increase in readmission rate. Similarly, the addition of KDM acceleration to the base model improved the model fit relative to the base model (LRT χ^2 = 21.8, p < 0.001), with each 5-year increase associated with 5.8% increase in readmission rate. As in frailty analysis, the magnitude of improvement was substantially greater for PhenoAge acceleration than for KDM acceleration.

Sensitivity Analysis

We compared the biological age estimation and downstream analysis of the results obtained using 1) complete-case analysis, 2) a ± 30-day laboratory imputation window approach, and 3) the MICE imputation approach. The ± 30-day window approach yielded biological age estimates and distributions highly similar to those of complete-case analysis, suggesting that the inclusion of temporally adjacent laboratory values had minimal impact on the calculated biological age measures. Though the application of MICE substantially increased the sample size compared to complete-case analysis, the distributions of both PhenoAge and KDM remained broadly comparable to those observed in complete-case analysis, supporting the stability of biological age estimates across missing-data strategies. Hazard ratio estimates from the Cox proportional hazard models evaluating the association between the biological age acceleration and mortality risk were highly consistent across the missing-data approaches. The complete case analysis (N = 902) yielded the highest discrimination, with c-statistic increasing from 0.75 (chronological age and sex alone) to 0.81 with the addition of PhenoAge acceleration and to 0.82 for

KDM acceleration. The ± 30-day imputation approach (N = 1065) showed a similar pattern, with c-statistic increasing from 0.74 to 0.79 (PhenoAge) and 0.80 (KDM), whereas with MICE imputation, the c-statistic increased from 0.72 to 0.74 (PhenoAge) and 0.73 (KDM). Although the model discrimination was highest in the complete-case analysis and decreased moderately with the use of imputation, the relative improvement in discrimination by adding biological age acceleration to the baseline model was consistent across the three approaches.

Discussion

Biological age has long been recognized as a more meaningful indicator of health than the number of years lived, yet its clinical utility has remained largely confined to research settings. This study demonstrates that biological age can be estimated directly from clinical data and that it carries a meaningful prognostic value across a broad range of patient outcomes using two established biomarker-based approaches, Levine PhenoAge and the Klemera-Doubal Method (KDM), applied to a heterogeneous real-world EHR population with fifteen years of longitudinal follow-up. PhenoAge and KDM were positively correlated with chronological age and were also strongly correlated with each other, indicating that they capture overlapping but not identical dimensions of biological aging. This is supported by our downstream analyses, where both measures identified overlapping patterns of age-related comorbidities, particularly renal and cardiometabolic diseases, and both were independently associated with mortality, frailty, and hospital readmission. However, the differences in effect sizes, disease rankings, and predictive performance indicate that the two clocks are not interchangeable. Prior studies have similarly reported positive correlation and with each other, though both showed broader distributions than typically reported in research cohorts such as NHANES and UK Biobank(He et al., 2024). Males had a modestly higher biological age acceleration than females for both PhenoAge and KDM. Patients in the highest and lowest biological acceleration tertiles differed only modestly by two years in chronological age, yet exhibited markedly different biological age acceleration. The highest tertile had a higher proportion of males than the lowest tertile. This divergence was also reflected at the biomarker level, patients in the fast-aging tertile had lower albumin, lymphocyte percentage and higher levels of glucose, RDW, CRP, ALP, and WBC compared to the lowest tertile. Biological aging also appeared dynamic as more than half of the patients changed PhenoAge acceleration tertiles over time. Male sex was consistently associated with a tendency toward faster biological aging on follow-up and a decreasing likelihood of moving out of the fastest-aging tertile.

Beyond capturing the dynamic nature of biological aging, both PhenoAge and KDM age acceleration were also clinically meaningful. Each was significantly associated with mortality, and PhenoAge showed the highest concordance, outperforming both KDM and chronological age. Restricted cubic spline analysis further revealed a non-linear association between biological age acceleration and mortality risk, with hazard ratios increasing more sharply at higher biological age values. Biological age acceleration was also associated with a broad range of comorbidities, with substantial overlap in the conditions implicated by both PhenoAge and KDM, although the magnitude and relative ranking of these associations differed slightly between the two measures. Additionally, biological age acceleration

provided incremental predictive value for both frailty and increased hospital readmission burden beyond chronological age and sex, with nested model comparisons confirming that the addition of PhenoAge and KDM acceleration improved model fit for each outcome. These findings collectively support biological age as a feasible and clinically relevant metric derived from routine EHR data.

Achieving these findings in a real-world EHR setting required addressing several challenges that distinguish clinical data from research data. The correlations observed between biological age and chronological age were lower than those typically reported in research cohorts such as NHANES and UK Biobank (Kuo et al., 2021; Mak et al., 2023), reflecting greater clinical heterogeneity of the hospital population rather than a failure of measures themselves. Both KDM and PhenoAge showed right-skewed distributions, with KDM having a broader spread and occasionally negative outcomes, reflecting its sensitivity to extreme biomarker combinations in a clinically heterogeneous population. PhenoAge, by contrast, yielded a higher mean biological age than KDM. Collectively, these findings suggest that while both measures are informative, they quantify biological age from different perspectives. While recent evidence suggests that biological age metrics perform similarly across cohorts (Provost et al., 2025), in this disease-enriched EHR population, PhenoAge demonstrated modestly stronger associations with mortality, frailty, and readmission count. While PhenoAge has been previously applied in several clinical settings, including ICU cohorts (Ho, 2023a), disease-specific cohorts such as acute ischemic stroke (Liu et al., 2026), acute myocardial infarction and heart failure (Gao et al., 2025), and COVID-19 (Strazhesko et al., 2022), and a heterogeneous EHR population examining social vulnerabilities (Knobel et al., 2025), applications of KDM to clinical hospital databases have been limited to disease-specific contexts such as COVID-19 (Strazhesko et al., 2022) and psoriasis (Lin et al., 2025). Studies have not estimated KDM biological age from biomarkers extracted from EHR workflows across a heterogeneous clinical population, where laboratory values are ordered for clinical rather than research purposes. Successfully translating these algorithms into EHR data required extensive laboratory result harmonization, unit standardization, management of heterogeneous naming conventions, and structured handling of missing data, challenges typically not encountered in controlled research settings.

This study observed modest sex-based differences in biological age acceleration, with males showing a slight tendency towards positive acceleration for both PhenoAge and KDM, though the distribution overlapped substantially between sexes. Beyond the modest sex difference in acceleration, our longitudinal analysis of tertile transitions revealed a more pronounced sex-related pattern. Males in the lowest baseline tertile were nearly twice as likely as females to transition into a higher acceleration tertile at follow-up, while males in the highest baseline tertile were significantly less likely to move out of that tertile. Older age was also associated with a lower likelihood of transitioning out of the highest tertile, suggesting that both male sex and older age independently contribute to the persistence of accelerated biological aging over time.

Both biological age measures were clinically relevant in outcome analyses. Using a nested Cox proportional hazard model, we found that chronological age was a strong predictor of mortality on its own, and that the addition of PhenoAge acceleration and KDM acceleration improved prediction,

consistent with prior literature (Zuo et al., 2026). Importantly, Levine PhenoAge and KDM acceleration remained significantly associated with mortality even after adjusting for chronological age, indicating that these measures capture physiological information relevant to mortality risk beyond what the chronological age alone provides. Levine PhenoAge showed the highest discrimination for mortality risk, outperforming both KDM and calendar age. The stronger performance of PhenoAge is likely explained by its origins as a mortality model, whereas KDM was designed primarily to estimate physiological age from biomarker values. While the difference in concordance was modest, biological age acceleration offers an important advantage over chronological age. Two patients of the same chronological age but with different physiological health would receive identical risk estimates based on chronological age alone, whereas biological age acceleration would differentiate between them. These associations were robust to the method used for handling missing biomarker data, with the sensitivity analyses using complete-case analysis and a 30-day imputation window approach yielding results consistent with our primary multiple imputation approach. Beyond linear Cox model associations, we further explored the shape of the relationship between the biological age and mortality risk using restricted cubic spline analysis, consistent with approaches used in prior studies(Talifu et al., 2025; Yang et al., 2026; Zhang et al., 2026). Restricted spline analysis suggested that the association between biological age and mortality was not strictly linear on the log-hazard scale, indicating that the effect of increasing biological age may vary across age ranges.

While mortality represents a major aging-related endpoint, we additionally observed significant associations between biological age acceleration measures and morbidity-related outcomes. The comorbidity analyses were particularly informative. Consistent with prior UK Biobank and NHANES studies, demonstrating that accelerated biological aging is associated with increased risk of chronic disease conditions across multiple organ systems(Talifu et al., 2025) (He et al., 2024; Tian et al., 2025), our analysis also showed similar associations between accelerated biological age and several major comorbidities spanning renal, cardiometabolic, respiratory, and hematologic categories. Positive associations were largely consistent across both PhenoAge and KDM, suggesting they are robust and model-independent findings.However, these findings are associative and do not imply causation and should be interpreted with caution, as a patient might have multiple coexisting comorbidities. Consequently, the observed associations may reflect the cumulative burden of multimorbidity rather than a single condition. Further studies involving Bayesian modeling can be helpful, which can simultaneously account for multiple existing conditions.

The cross-sectional findings raise the question of whether the biological age differences associated with a disease are persistent over time. Longitudinal analysis demonstrated that biological age trajectories also differed according to the disease status over time. Patients with chronic diseases such as CKD and type 2 diabetes showed consistently higher PhenoAge acceleration across the follow-up period compared to patients without these conditions. One possible explanation for the persistent elevation in biological age with these conditions is that these continuously perturb the biomarkers underlying biological age estimates and are not fully reversible by current treatments. This is consistent with the concept of inflammaging, where chronic diseases drive accelerated biological aging independent of

disease management (Campisi, 2013; Franceschi, 2014). While current treatments for chronic conditions may partially improve the underlying biomarkers, the persistent elevation in biological age acceleration observed in our cohort suggests that disease treatment does not fully restore biomarker profiles to levels seen in disease-free individuals.

In addition to association with morbidity and mortality, biological age acceleration also demonstrated important relationships with frailty status, with greater acceleration associated with higher odds of being non-fit. Biological age acceleration provided meaningful incremental predictive value for frailty status even after accounting for chronological age and sex, with nested model comparisons confirming that both PhenoAge acceleration and KDM acceleration improved the model fit beyond age alone. Prior studies have similarly reported associations between accelerated biological aging and increased frailty risk in aging populations (Phyo et al., 2024; Tay et al., 2025). Consistent with prior studies showing PhenoAge has greater sensitivity in identifying high-risk individuals, PhenoAge showed a stronger and more robust association with frailty than KDM in our dataset.

Beyond frailty, both biological age measures were also associated with healthcare utilization. Higher PhenoAge and KDM acceleration were both associated with higher hospital readmission burden, indicating that patients with higher biological age may be at increased risk of subsequent readmissions. Notably, chronological age alone was not significantly associated with readmission risk, whereas both PhenoAge acceleration and KDM acceleration were significant predictors, suggesting that biological age acceleration captures readmission risk information that chronological age alone does not provide. As with the frailty outcome, the improvement in model fit was pronounced for PhenoAge acceleration than for KDM acceleration. Prior studies examining biological age acceleration and hospital readmission have been focused on specific hospital settings, such as ICU readmission or postoperative outcomes(Ho, 2023; Jawad et al., 2026). To our knowledge, our study is the first to evaluate the readmission burden with the EHR population using longitudinal clinical data, demonstrating consistent associations for both PhenoAge and KDM acceleration measures.

Applying biological age to EHR presents several challenges, including missing and irregular measurements, varying laboratory units, and limited generalizability across health systems. Direct application of KDM BioAge package on our EHR dataset presented a set of challenges. First, the presence of multiple encounters per patient interfered with the KDM algorithm, reducing its correlation with chronological age. Second, including patients below 30 and above 75 years of age also reduced the correlation. A similar issue also arose when calculating age acceleration. Deriving residuals using all encounters introduced measurements within individuals, which we resolved by calculating residuals using only the first encounter per patient and applying the resulting parameters to all subsequent encounters. Unlike KDM, which required several methodological adjustments to apply to longitudinal EHR data, PhenoAge's primary challenge was data availability rather than algorithmic sensitivity. Once the laboratory values were correctly harmonized to the required units, the model could be applied directly using its fixed biomarker specification. Although KDM is designed to flexibly incorporate any set of

available biomarkers, we constructed it to the same nine-biomarker panel used for PhenoAge to enable direct methodological comparison between the two measures.

Despite these challenges, biological age could still be integrated into EHR systems using routinely collected clinical data. The associations observed in our study with mortality, frailty, and hospital readmission suggest that biological age acceleration carries prognostic information beyond chronological age. However, showing that biological age is useful in clinical practice requires more than association. The next step would be to determine whether biological age can meaningfully stratify patients into risk groups who would benefit from closer monitoring or targeted interventions, beyond what chronological age and standard clinical assessments already provide. Further studies should also evaluate whether interventions aimed at slowing biological aging, such as lifestyle modification, exercise, and dietary interventions, can reduce biological age acceleration and improve clinical outcomes. This perspective is consistent with emerging views in geroscience that aging may be considered as a disease rather than merely a risk factor for disease (Bulterijs et al., 2015), providing a rationale for interventions aimed at slowing biological aging. If such a utility is demonstrated, biological age could be integrated into EHR dashboards and updated automatically at each patient encounter, providing clinicians with a longitudinal view of the patient's aging trajectory and helping identify those whose health is declining faster than expected for their age. Prospective studies would ultimately be needed to determine whether biological age estimates lead to measurable improvements in patient outcomes.

A few limitations of this study should be considered. This study cohort is not a general clinical population, but rather participants enrolled in an observational study with available EHR data, and was enriched for patients with chronic disease. The elevated biological age and stronger comorbidity associations may therefore partly reflect this disease enrichment. External validation with a large, independent cohort (eg:- AllOfUS) is essential to generalize the findings from this study. Additionally, our analysis did not account for socioeconomic status or environmental factors such as smoking and diet, all of which are established contributors to biological aging. The absence of these variables limits our ability to determine whether the observed outcomes in our study are confounded by them. An additional limitation relates to the biological age algorithms themselves. PhenoAge and KDM may produce extreme or undefined estimates when biomarker values are extremely abnormal, which can occur in critically ill patients. A practical limitation of PhenoAge is that it relies on biomarkers such as CRP, which are not routinely available for all patients. As a result, applying PhenoAge in real-world clinical settings may require imputation or may be limited to subsets of patients with sufficient laboratory data. Although we used a structured two-step imputation strategy and verified post-imputation distributions, residual bias from missing data may remain.

Conclusion

We showed that biological age can be estimated from routine clinical data and is meaningfully associated with patient outcomes, including mortality, frailty, and hospital readmission. The successful

application of both PhenoAge and KDM to a heterogeneous, real-world EHR population, with longitudinal follow-up over fifteen years, demonstrates that biological age is a feasible and clinically relevant metric. As EHR data has become more accessible and standardized, biological age modeling has the potential to evolve from a research tool into a scalable clinical metric that, when integrated into routine clinical workflows, could support personalized care across diverse patient populations.

Corresponding Author

Smitha Edakalavan, sme57@pitt.edu

Human Ethics and Consent

This study was conducted in accordance with the Declaration of Helsinki. The study protocol was reviewed and approved by the University of Pittsburgh Institutional Review Board (

STUDY19030104). Given the retrospective nature of the study and the use of existing de-identified electronic health record data, the requirement for informed consent was waived by the IRB.

Funding Declaration

This work was supported by the National Institutes of Health (T32 HL007563)

Declarations

Author Contribution

S.E performed the data analysis and wrote the manuscript. S.M.N provided technical methodological guidance and reviewed the manuscript. A.M, J.R., and M.K. reviewed the manuscript. All authors have reviewed and approved the manuscript.

Acknowledgement:

We thank Mrs. Melissa Saul for her contributions to the acquisition and curation of the dataset used in this study.

Code availability:

The underlying code for this study is not publicly available but may be made available to qualified researchers on reasonable request from the corresponding author.

Data Availability

The datasets analyzed during the current study are not publicly available due to IRB restrictions on the sharing of protected health information under HIPAA compliance. Aggregated data supporting the findings of this study are available from the corresponding author on reasonable request.

References

1. Azur, M.J. *et al.* (2011) “Multiple imputation by chained equations: what is it and how does it work?,” *International Journal of Methods in Psychiatric Research*, 20(1), pp. 40–49. Available at: https://doi.org/10.1002/mpr.329.
2. Bulterijs, S. *et al.* (2015) “It is time to classify biological aging as a disease,” *Frontiers in Genetics*, 6. Available at: https://doi.org/10.3389/fgene.2015.00205.
3. Campisi, J. (2013) “Aging, Cellular Senescence, and Cancer,” *Annual review of physiology*, 75, pp. 685–705. Available at: https://doi.org/10.1146/annurev-physiol-030212-183653.
4. Deelen, J. *et al.* (2013) “Identifying the genomic determinants of aging and longevity in human population studies: Progress and challenges,” *BioEssays*, 35(4), pp. 386–396. Available at: https://doi.org/10.1002/bies.201200148.
5. Fitzgerald, K.N. *et al.* (2020) “Reversal of Epigenetic Age with Diet and Lifestyle in a Pilot Randomized Clinical Trial.” medRxiv, p. 2020.07.07.20148098. Available at: https://doi.org/10.1101/2020.07.07.20148098.
6. Franceschi, C. (2014) “Chronic Inflammation (Inflammaging) and Its Potential Contribution to Age-Associated Diseases.”
7. Harris, K.M. *et al.* (2024) “Sociodemographic and Lifestyle Factors and Epigenetic Aging in US Young Adults: NIMHD Social Epigenomics Program,” *JAMA Network Open*, 7(7), p. e2427889. Available at: https://doi.org/10.1001/jamanetworkopen.2024.27889.
8. He, Y. *et al.* (2024) “Accelerated biological aging: unveiling the path to cardiometabolic multimorbidity, dementia, and mortality,” *Frontiers in Public Health*, 12. Available at: https://doi.org/10.3389/fpubh.2024.1423016.
9. Ho, K.M. (2023) “Biological age as a predictor of unplanned intensive care readmission during the same hospitalization,” *Heart & Lung*, 62, pp. 249–255. Available at: https://doi.org/10.1016/j.hrtlng.2023.08.010.
10. Jawad, B.N. *et al.* (2026) “Acute biological age as a determinant of adverse outcomes requiring hospitalization in Danish emergency department patients,” *Communications Medicine*, 6(1), p. 156. Available at: https://doi.org/10.1038/s43856-026-01428-6.
11. Khan, S.S., Singer, B.D. and Vaughan, D.E. (2017) “Molecular and physiological manifestations and measurement of aging in humans,” *Aging Cell*, 16(4), pp. 624–633. Available at: https://doi.org/10.1111/acel.12601.
12. Klemera, P. and Doubal, S. (2006) “A new approach to the concept and computation of biological age,” *Mechanisms of Ageing and Development*, 127(3), pp. 240–248. Available at:

https://doi.org/10.1016/j.mad.2005.10.004.

13. Kuo, C.-L. *et al.* (2021) “Biological Aging Predicts Vulnerability to COVID-19 Severity in UK Biobank Participants,” *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 76(8), pp. e133–e141. Available at: https://doi.org/10.1093/gerona/glab060.
14. Kwon, D. and Belsky, D.W. (2021) “A toolkit for quantification of biological age from blood chemistry and organ function test data: BioAge,” *GeroScience*, 43(6), pp. 2795–2808. Available at: https://doi.org/10.1007/s11357-021-00480-5.
15. Levine, M.E. *et al.* (2018) “An epigenetic biomarker of aging for lifespan and healthspan,” *Aging (Albany NY)*, 10(4), pp. 573–591. Available at: https://doi.org/10.18632/aging.101414.
16. Li, K. *et al.* (2022) “Gut and oral microbiota associations with viral mitigation behaviors during the COVID-19 pandemic,” *Frontiers in Cellular and Infection Microbiology*, 12, p. 966361. Available at: https://doi.org/10.3389/fcimb.2022.966361.
17. Liu, Z. *et al.* (2018) “A new aging measure captures morbidity and mortality risk across diverse subpopulations from NHANES IV: A cohort study,” *PLOS Medicine*, 15(12), p. e1002718. Available at: https://doi.org/10.1371/journal.pmed.1002718.
18. Lu, W.-H. (2024) “Effect of Modifiable Lifestyle Factors on Biological Aging,” *JAR life*, 13, pp. 88–92. Available at: https://doi.org/10.14283/jarlife.2024.13.
19. Lu, W.-H. (no date) “Effect of Modifiable Lifestyle Factors on Biological Aging,” *JAR life*, 13, pp. 88–92. Available at: https://doi.org/10.14283/jarlife.2024.13.
20. Mak, J.K.L. *et al.* (2023) “Clinical biomarker-based biological aging and risk of cancer in the UK Biobank,” *British Journal of Cancer*, 129(1), pp. 94–103. Available at: https://doi.org/10.1038/s41416-023-02288-w.
21. MedBio Project or University of Pittsburgh (2017) “MedBio.” Available at: https://medbio.microbiome.pitt.edu/.
22. Orfila, F. *et al.* (2022) “Validation of an electronic frailty index with electronic health records: eFRAGICAP index,” *BMC Geriatrics*, 22, p. 404. Available at: https://doi.org/10.1186/s12877-022-03090-8.
23. Phyo, A.Z.Z. *et al.* (2024) “Epigenetic age acceleration and the risk of frailty, and persistent activities of daily living (ADL) disability,” *Age and Ageing*, 53(6), p. afae127. Available at: https://doi.org/10.1093/ageing/afae127.
24. Provost, G. *et al.* (2025) “No winners or losers: clinical chemistry-based biological aging metrics perform similarly across cohorts and health outcomes,” *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 80(11), p. glaf179. Available at: https://doi.org/10.1093/gerona/glaf179.
25. Talifu, Z. *et al.* (2025) “The Association Between Accelerated Biological Aging and the Physical, Psychological, and Cognitive Multimorbidity and Life Expectancy: Cohort Study,” *Aging Cell*, 24(9), p. e70142. Available at: https://doi.org/10.1111/acel.70142.

26. Tay, J.H. *et al.* (2025) “Biological age measured by DNA methylation clocks and frailty: a systematic review and meta-analysis,” *The Lancet. Healthy Longevity*, 6(10), p. 100773. Available at: https://doi.org/10.1016/j.lanhl.2025.100773.
27. Tian, Y. *et al.* (2025) “Biological Age Acceleration Associated with the Progression Trajectory of Cardio-Renal–Metabolic Multimorbidity: A Prospective Cohort Study,” *Nutrients*, 17(11), p. 1783. Available at: https://doi.org/10.3390/nu17111783.
28. Yang, Y. *et al.* (2026) “Association of biological age acceleration with cardiovascular disease and premature mortality: a population-based prospective cohort study,” *Frontiers in Public Health*, 14. Available at: https://doi.org/10.3389/fpubh.2026.1816971.
29. Zhang, D. *et al.* (2026) “The Mediating Role of Biological Age Advance in the Association Between Periodontitis and Mortality: Biological Aging Links Periodontitis to Mortality,” *Clinical and Experimental Dental Research*, 12(1), p. e70305. Available at: https://doi.org/10.1002/cre2.70305.
30. Zhang, Z. (2016) “Multiple imputation with multivariate imputation by chained equation (MICE) package,” *Annals of Translational Medicine*, 4(2), p. 30. Available at: https://doi.org/10.3978/j.issn.2305-5839.2015.12.63.
31. Zheng, N.S. *et al.* (2020) “PheMap: a multi-resource knowledge base for high-throughput phenotyping within electronic health records,” *Journal of the American Medical Informatics Association*, 27(11), pp. 1675–1687. Available at: https://doi.org/10.1093/jamia/ocaa104.
32. Zuo, S. *et al.* (2026) “Biological age acceleration, longitudinal change and mortality risk in the Dutch Lifelines cohort,” *GeroScience* [Preprint]. Available at: https://doi.org/10.1007/s11357-026-02241-8.

Figures

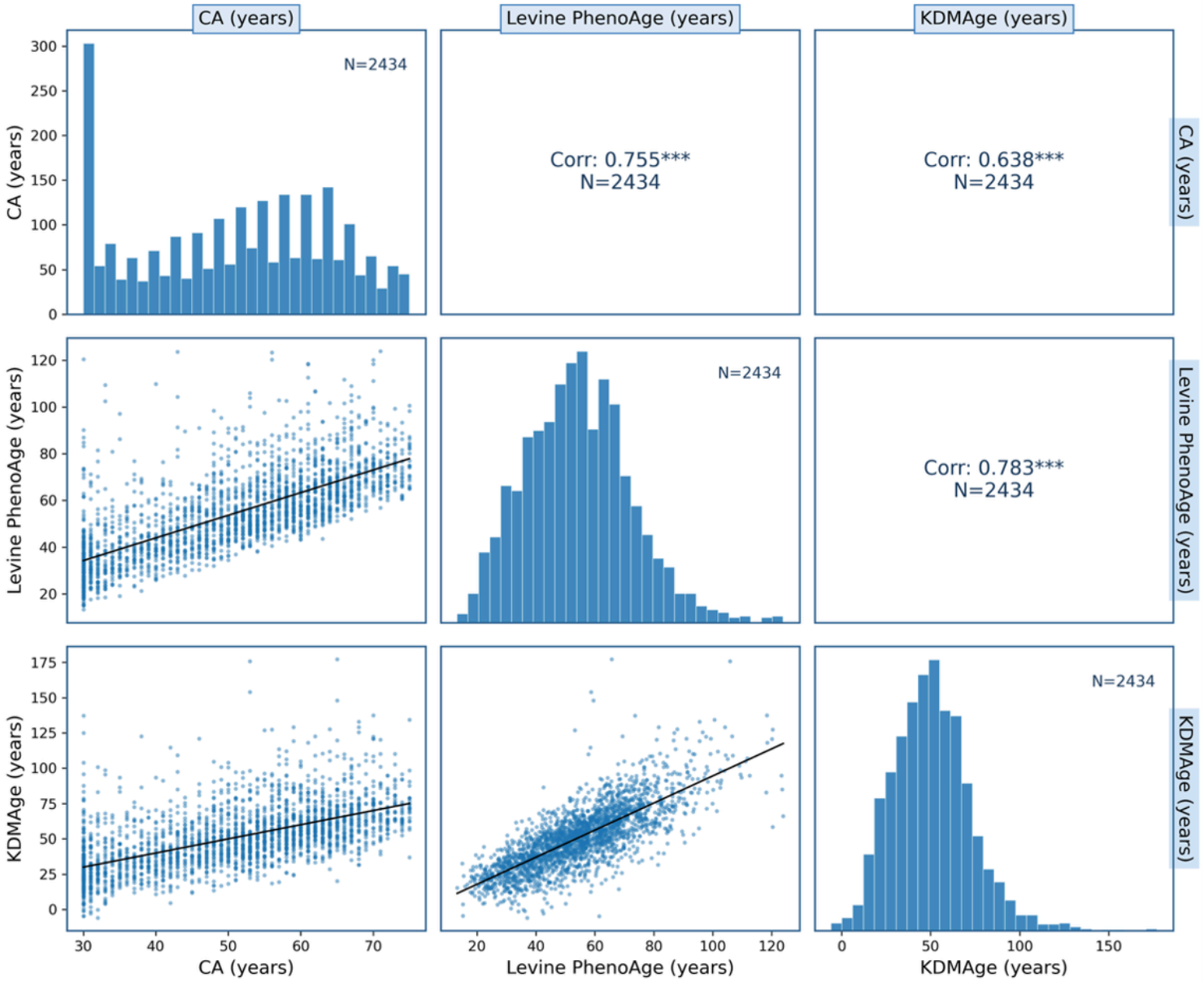


Figure 1

Pairwise relationships between chronological age, Levine PhenoAge and KDMAge. The distribution is represented along the diagonal, scatter plots in the lower panels with fitted regression lines, and Spearman correlations in the upper panels.

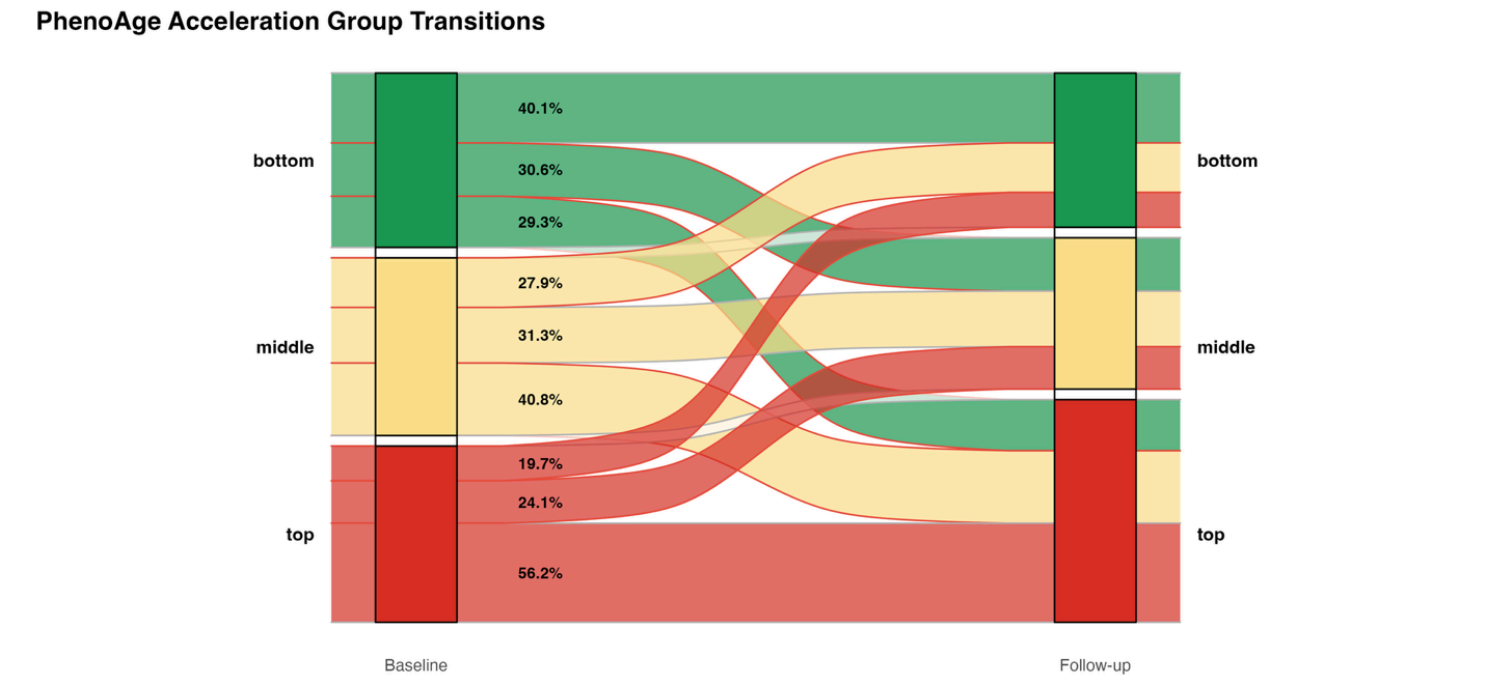


Figure 2

Transitions in PhenoAge acceleration tertile membership from baseline to follow-up. Patients were classified into tertiles of PhenoAge acceleration (bottom, middle, and top thirds) based on the distribution of acceleration values.

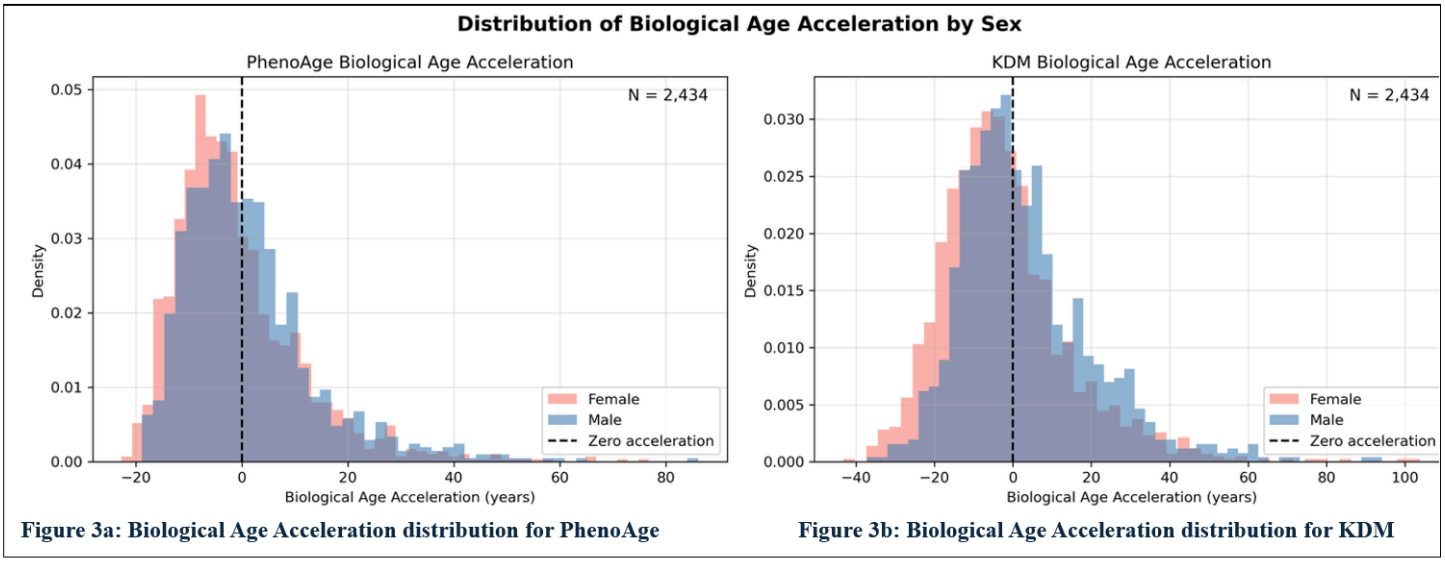


Figure 3a: Biological Age Acceleration distribution for PhenoAge

Figure 3b: Biological Age Acceleration distribution for KDM

Figure 3

See image above for figure legend

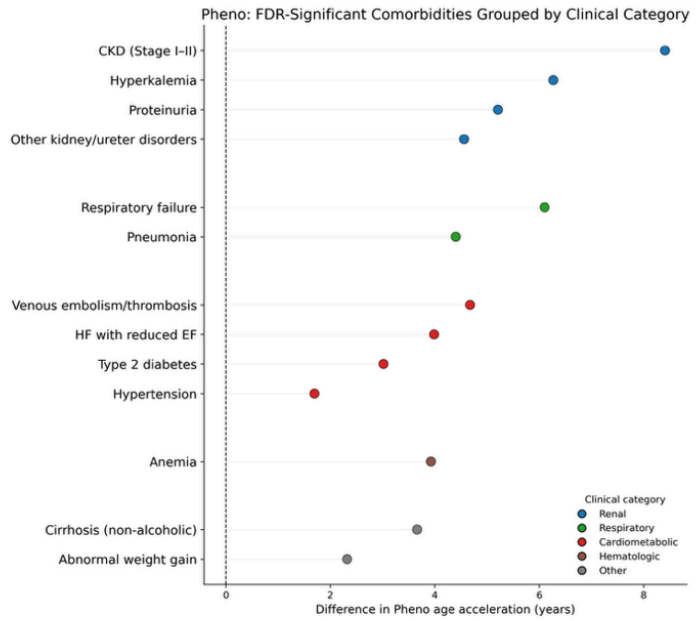


Fig 4a: Levine PhenoAge Acceleration

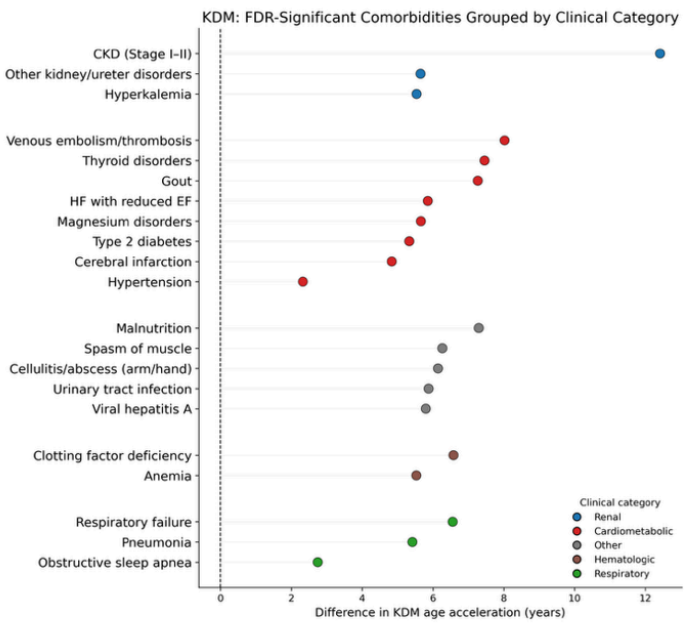


Fig 4b: KDM Acceleration

Figure 4

***P*lot displaying the association between Levine PhenoAge *and KDM* acceleration and comorbidities**. The x-axis represents the mean difference in age acceleration (years) between patients with and without each comorbidity.

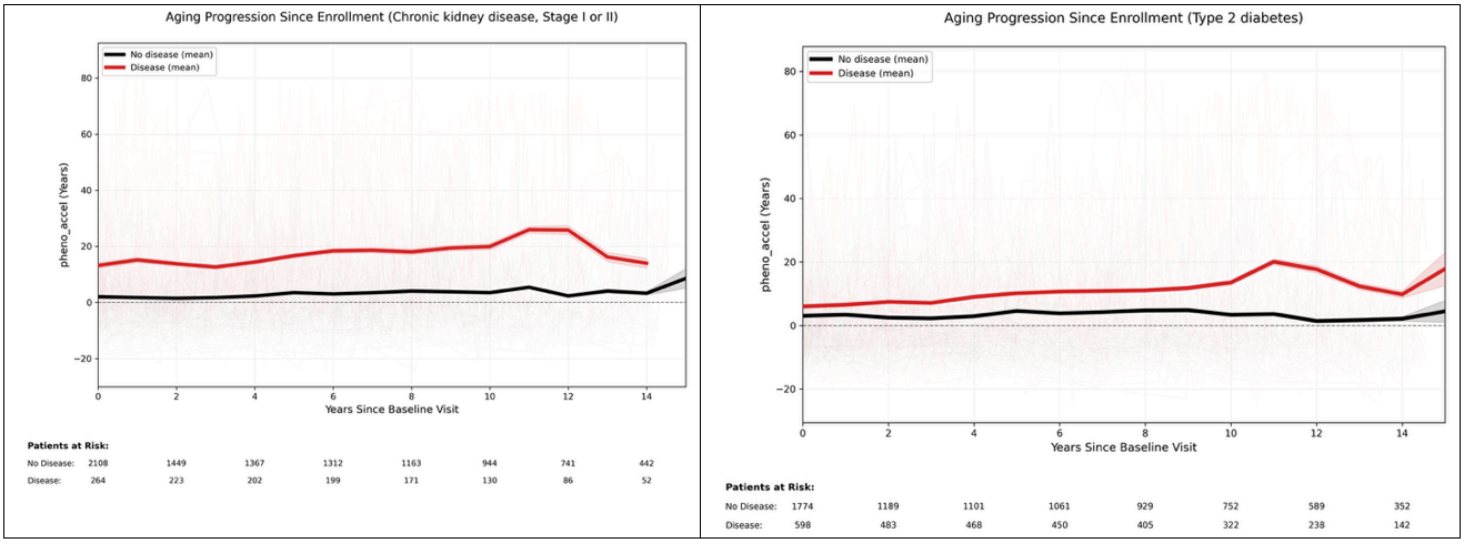


Figure 5

Longitudinal Trajectories of Biological Age Acceleration

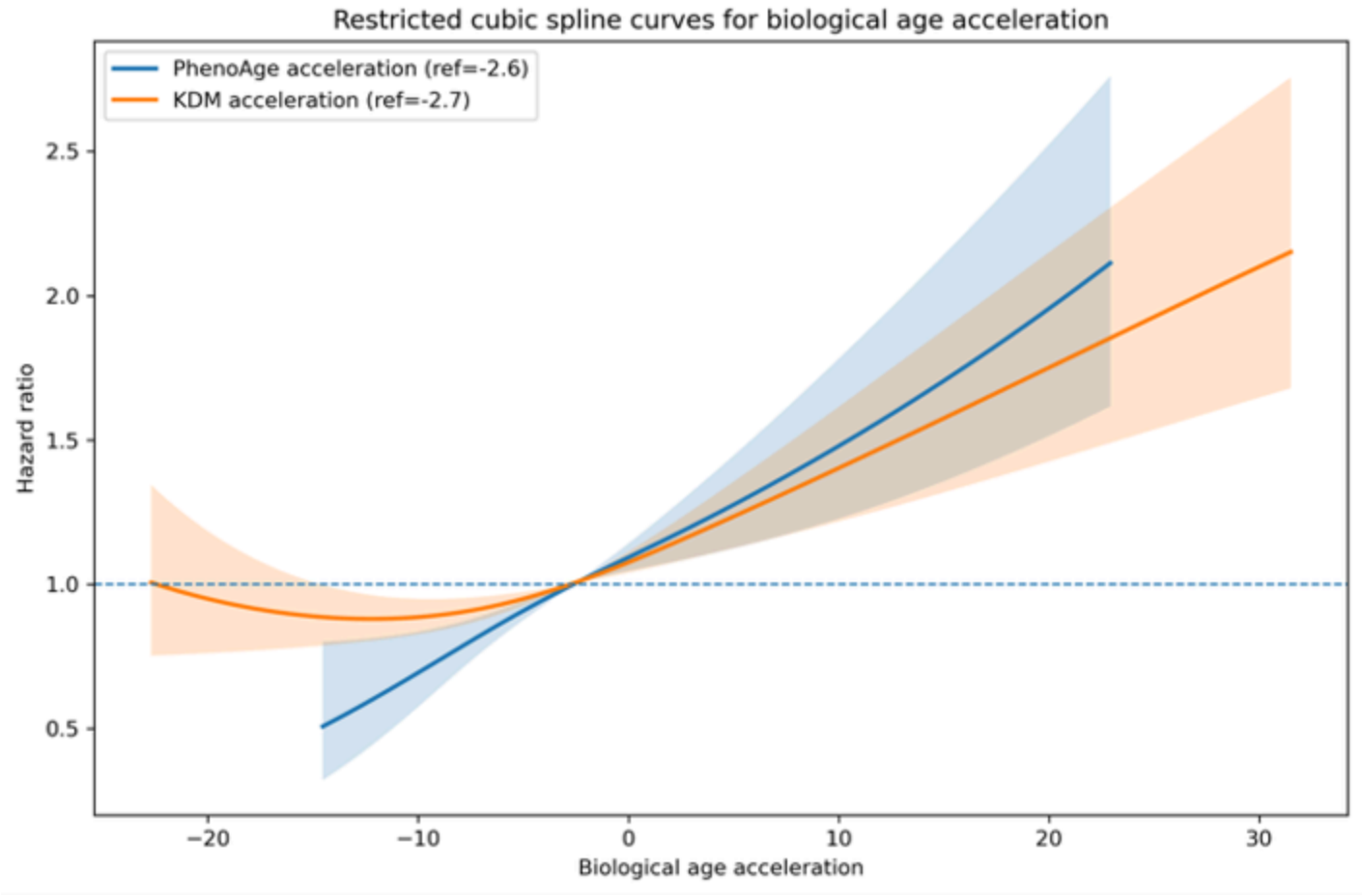


Figure 6

Restricted Cubic Splines for PhenoAge and KDM Age and Acceleration

Supplementary Files

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

- Supplementary.docx
- volcanopheno2.csv
- volcanokdm1.csv