



OPEN Associations of biological aging with the morbidity and all-cause mortality of patients with lung cancer

Qihang Xie^{1,4}, Haoran Qu^{1,4}, Siyu Xie², Wenhao Li¹, Rui Ouyang¹, Chengxiang Zhang³ & Ming Du^{1✉}

Accelerated biological aging has been suggested to be an important risk factor for age-related diseases, but its role in the development of lung cancer is still unclear. The aim of this study was to analyze the relationship between biological aging and lung cancer and its impact on mortality using the NHANES database. We calculated the chronological age-adjusted biological age of 40,765 participants in the U.S. National Health and Nutrition Examination Survey (NHANES) database from 2001 to 2020 using the phenotypic age (PhenoAge) method. Both cross-sectional and longitudinal analyses were performed to assess the associations of biological aging with lung cancer prevalence and mortality in the cohort. Logistic regression and Cox proportional hazards models were used to examine these associations. Sensitivity analysis was conducted using propensity score matching (PSM). Our results indicated that accelerated biological aging was positively associated with an increased risk of developing lung cancer. Individuals with accelerated aging had a greater risk of developing lung cancer than nonaccelerated aging individuals did, with odds ratios of 2.60 (full adjustment 95% CI 1.59–4.22, $p < 0.001$). According to the fully adjusted models, participants in the Q4 subgroup had a 334% greater risk of developing lung cancer. In addition, accelerated biological aging was positively associated with all-cause mortality among individuals with lung cancer. Our study demonstrated that the identification of individuals with accelerated biological aging represents a promising strategy for the identification of populations at elevated risk of developing lung cancer and serves as an independent predictor of mortality in this population.

Keywords Lung cancer, Aging, Biological age, NHANES

Lung cancer is a malignant tumor characterized by uncontrolled cell growth in lung tissues¹. According to 2020 global cancer statistics, lung cancer has the highest incidence (11.4%) and mortality (18%) worldwide². Lung cancer may present with a range of symptoms, including cough, chest pain, dyspnea, anorexia, weight loss, and fatigue³. These studies reported a mean survival time of 13 months, a median survival time of 4.8 months, and a 5-year survival rate of 17%^{4,5}. Age and sex have been identified as predictors of survival (Abedi et al., 2019). Therefore, to reduce the disease burden, identifying risk factors and monitoring the health status of patients with lung cancer are imperative. Accelerated biological aging, characterized by a premature decline in homeostasis, is a major risk factor for age-related diseases and mortality⁶. Aging is a primary contributor to the development and progression of lung cancer⁷, and it is associated with molecular and physiological changes that impair lung function, reduce lung remodeling and regeneration capacity, and increase susceptibility to acute and chronic lung diseases, including lung cancer^{8,9}. However, few population-based studies have explored the relationship between lung cancer and biological aging. Previous studies have shown that increased biological age is associated with increased risks of cancer, mortality, depression, anxiety, and chronic kidney disease^{10–12}. Researchers have developed and validated a multisystem aging measure called phenotypic age (PhenoAge) as a surrogate marker of biological aging¹³. This measure emphasizes the use of clinical biomarkers to approximate an individual's

¹Department of Cardiothoracic Surgery, The First Affiliated Hospital of Chongqing Medical University, 1 Youyi Road, Yuzhong District, Chongqing, China. ²Department of Child Development, Graduate School of Child Development, Osaka University, Suita, Osaka, Japan. ³Department of General Medicine, The First Affiliated Hospital of Chongqing Medical University, Chongqing, China. ⁴Qihang Xie and Haoran Qu contributed equally to this work. ✉email: docxie923@gmail.com

biological aging rate, as opposed to focusing solely on genetic or molecular markers. This approach has been shown to be a reliable indicator of remaining life expectancy, morbidity, and mortality risk, making it a powerful tool for assessing biological aging in population studies¹³. Previous studies have demonstrated a link between aging and an increased risk of various cancers, including lung cancer¹⁴. Recent studies have further explored the relationship between biological aging, as measured by DNA methylation-based biomarkers, and lung cancer risk. Age acceleration measures, such as PhenoAge and GrimAge, have been shown to be associated with increased cancer risk, particularly lung cancer^{13,15,16}. A prospective analysis in the UK Biobank and a Mendelian randomization study revealed that PhenoAgeAccel was positively associated with lung cancer risk, independent of chronological age¹⁷. Additionally, smoking and electronic cigarette use have been linked to accelerated epigenetic aging in lung tissue, with both smokers and vapers showing significantly older GrimAge and shorter telomere lengths than nonsmokers do¹⁸. These findings suggest that epigenetic age acceleration is an independent risk factor for lung cancer and may provide insights into the relationships among biological aging, smoking, and cancer risk^{13,15}. However, compared with DNA methylation-based biological age measurements, PhenoAge is more accessible and less expensive. The present study is the first time that PhenoAge has been used to explore the relationship between biological aging and lung cancer in the U.S. population. The identification of individuals whose biological age exceeds their chronological age may facilitate the implementation of timely interventions, thereby preventing the occurrence of disease and effectively reducing the mortality rate of the disease.

Therefore, the aim of this study was to analyze the relationship between biological aging and lung cancer, as well as its impact on mortality, on the basis of the NHANES database.

Methods

Study population

In this study, we analyzed data to examine the associations among accelerated biological aging, lung cancer risk, and all-cause mortality among individuals with lung cancer. Data were drawn from the U.S. National Health and Nutrition Examination Survey (NHANES), a continuous cross-sectional study conducted over 2-year cycles from 2001 to 2020. We included participants aged > 18 years with available data on the clinical biomarkers required for the calculation of PhenoAge and lung cancer status on the basis of self-reported physician diagnoses. For the mortality analysis, we further excluded individuals with missing death records.

This study was conducted in accordance with the Declaration of Helsinki and approved by the National Center for Health Statistics (NCHS) Research Ethics Review Board¹⁹. All participants provided written informed consent.

Ascertainment of outcomes

In the NHANES, all participants completed a questionnaire that included a cancer-related question: “Have you ever been told by a doctor or other health professional that you had cancer or a malignancy of any kind?” The participants who responded “yes” were asked a follow-up question (“What kind of cancer was it?”) to identify the participants who were diagnosed with lung cancer.

For the longitudinal analysis, mortality status was ascertained from the NHANES Public Use Linked Mortality File, which provides vital status follow-up data through record linkage, enabling survival analyses of NHANES participants²⁰.

Procedures

PhenoAge was computed for all NHANES participants as a measure of biological aging which captures the influence of clinical biomarkers on health outcomes and aligns with previous work demonstrating its ability to predict disease risk and aging-related outcomes more effectively than chronological age¹³. PhenoAge was derived using the Gompertz proportional hazard model, which is widely recognized in aging research for modeling mortality risk and its exponential increase with age. This model is particularly well-suited for aging studies as it reflects the acceleration of aging-related mortality risk over time. The calculation of PhenoAge follows a well-established protocol, where two Gompertz proportional hazard models are parameterized: The first model incorporates ten biomarkers and chronological age to estimate an individual's biological aging status. The second model uses chronological age alone as the predictor variable, serving as a baseline comparison to evaluate whether biological aging (PhenoAge) provides additional predictive value beyond chronological age alone (Levine, 2013). In accordance with the protocol previously outlined in related studies, the NHANES III dataset was employed as the reference population, with an individual's PhenoAge reflecting the chronological age of an individual with the same mortality risk²¹. The following formula was employed for the purpose of determining the PhenoAge.

$$\begin{aligned} \text{PhenoAge} &= 143.5671 \\ &+ \frac{\ln[-0.0059383581 \times \ln[1 - \text{mortality risk}]]}{0.08548908} \\ \text{mortality risk} &= 1 - e^{-e^{\frac{\exp(120x\gamma) - 1}{\gamma}}} \\ \gamma &= 0.007354285 \end{aligned}$$

where $xb = -19.907 - 0.0336 \times \text{Albumin} + 0.0095 \times \text{Creatinine} + 0.1953 \times \text{Glucose} - 0.0120 \times \text{Lymphocyte Percent} + 0.0268 \times \text{Mean Cell Volume} + 0.3306 \times \text{Red Cell Distribution Width} + 0.00188 \times \text{Alkaline Phosphatase} + 0.0554 \times \text{White Blood Cell Count} + 0.0804 \times \text{Chronological Age}$.

The concept of biological aging is defined as the value of biological age that exceeds an individual's chronological age. Accelerated biological aging was determined through a calculation of the difference between the estimated biological age and the actual chronological age²¹.

In this study, based on their clinical relevance and existing literature^{12,22}, the following variables were included: age, sex (male, female), race (non-Hispanic white, non-Hispanic black, Mexican American, other), education level (grade school or less, high school, more than high school), poverty status (ratio of family income to poverty), smoking status (never, former, current smoker), drinking status (never, former, light or moderate, heavy), and BMI. Additionally, history of emphysema was included in this study. The participants were classified according to their smoking status, which was divided into three categories: current smokers, former smokers, and nonsmokers. The term 'current smoker' was used to describe individuals who had smoked on a regular basis and had smoked at least 100 cigarettes in their lifetime. Former smokers had smoked at least 100 cigarettes and subsequently ceased this practice. Those who had never smoked were classified as nonsmokers, as were those who had smoked fewer than 100 cigarettes. The drinking status of the participants was categorized according to the following definitions: never (had consumed fewer than 12 alcoholic beverages in their lifetime), former (had consumed 12 or more alcoholic beverages in a single year but had not drunk any alcohol in the preceding year or had not drunk any alcohol in the preceding year but had consumed 12 or more alcoholic beverages at some point in their lifetime), light/moderate drinker (≤ 1 alcoholic beverage per day for women or ≤ 2 alcoholic beverages per day for men on average over the past year), or current heavier drinker (> 1 alcoholic beverage per day for women or > 2 alcoholic beverages per day for men on average over the past year). Individuals can be classified according to their body mass index (BMI) into one of three categories: ≤ 25 , > 25 to < 30 , or ≥ 30 kg/m². The medical history of the participants regarding emphysema was established on the basis of the self-reported information provided by the participants.

Study design

In this study, we employed a two-stage design to examine the relationships among accelerated biological aging, lung cancer risk, and all-cause mortality among individuals with lung cancer, utilizing data from the NHANES. In Stage 1, we performed a cross-sectional analysis to investigate the association between PhenoAge acceleration and the prevalence of lung cancer at baseline. Lung cancer status was determined on the basis of self-reported diagnoses collected through NHANES questionnaires. In Stage 2, a longitudinal analysis was conducted to evaluate the association between PhenoAge acceleration and all-cause mortality among participants diagnosed with lung cancer. Mortality data were obtained from the NHANES Public Use Linked Mortality File, enabling survival analysis during the follow-up period.

Statistical analysis

In the present study, we employed NHANES data to ascertain the baseline characteristics of patients with and without lung cancer. The data are expressed as the means \pm SDs ($\bar{x} \pm s$). The statistical analysis of the differences among the groups was conducted using either Student's *t* test or one-way analysis of variance (ANOVA). For the purpose of multiple comparisons, the SNK or LSD method was employed. Qualitative data are presented as percentages and were analyzed via either a chi-square (χ^2) test or Fisher's exact test, as appropriate. Missing data were handled by treating missing values for categorical variables (BMI, smoking status, drinking status, education level, and history of emphysema) as separate categories to preserve the full sample size and avoid potential selection bias. The reported *P* value was two-sided, and a value of less than 0.05 was considered to indicate statistical significance. In alignment with the NHANES analytic guidelines, the present analyses accounted for the complex sampling design, including sampling weights, stratification (strata), and clustering (PSU, primary sampling units). All statistical models incorporated NHANES survey design variables to ensure correct variance estimation and population inference. Furthermore, all analyses were adjusted for potential confounding variables.

In this study, PhenoAge acceleration (PhenoAgeAccel) was initially calculated as a continuous variable using the R package BioAge. To explore the distribution of PhenoAgeAccel across participants, we also categorized it into quartiles on the basis of the PhenoAgeAccel distribution in the study cohort, with cutoffs corresponding to the 25th, 50th (median), and 75th percentiles. Furthermore, PhenoAge acceleration was computed by regressing PhenoAge against chronological age at the time of biomarker measurement, with the residuals defining the PhenoAgeAccel.

In logistic regression models, PhenoAge acceleration was assessed both as a continuous variable and as a binary variable. 'Accelerated biological aging' was defined as a PhenoAge acceleration value exceeding zero, whereas 'nonaccelerated aging' was defined as a PhenoAge acceleration value of zero or less. Given the presence of zero inflation (many participants exhibited no evidence of accelerated aging), we used both continuous and binary classifications of PhenoAge acceleration to assess its impact on lung cancer risk.

The clinical outcome of all-cause mortality over time was described with Kaplan–Meier survival curves and compared by the log-rank test. Furthermore, Cox proportional hazards regression models were used to evaluate the relationships between PhenoAge and mortality in individuals with lung cancer. The confounders were selected on the basis of clinical interest, previous scientific literature and the identification of all significant covariates in the univariate analysis. The objective was to evaluate the association between lung cancer incidence and PhenoAge acceleration, and for this purpose, several models were employed. The crude model was not adjusted for any variables, whereas Model 1 was adjusted for age, sex, and race. Model 2 was further adjusted for education level, the family income-to-poverty ratio, smoking status, drinking status, BMI, and history of emphysema. This stepwise adjustment approach is widely used in epidemiological studies to evaluate whether biological aging provides additional predictive value beyond demographic and lifestyle factors²². It allows for a clearer understanding of the independent association between PhenoAge acceleration and lung cancer risk. The

P value for the trend across increasing exposure groups was calculated using integer values (1, 2, 3, and 4). A restricted cubic spline (RCS) regression model was employed to assess the dose-response relationships between biological age (PhenoAge) and the risk of developing lung cancer, as well as between biological age and all-cause mortality in individuals with lung cancer. The model was fitted using three knots placed at the 25th, 50th, and 75th percentiles of the PhenoAge distribution to capture the nonlinear relationships and better understand how biological age influences both outcomes.

We performed several sensitivity analyses. First, participants who had values that fell within the 1% extremes of the PhenoAge acceleration range were excluded to rule out the effects of extreme values. Second, the associations between PhenoAge acceleration and lung cancer and all-cause mortality were reanalyzed without the consideration of complex sampling designs. Third, to determine the impact of various factors on the relationship between lung cancer and biological aging, we conducted a subgroup analysis and an interactive test. Fourth, we conducted sensitivity analyses to assess whether missingness influenced the study outcomes. Finally, we applied several additional association inference models: propensity score adjustment (PSA), propensity score matching (PSM)²³, inverse probability of treatment weighting (IPTW)²⁴, standardized mortality ratio weighting (SMRW)²⁵, pairwise algorithm (PA)²⁶, and overlap weight (OW)²⁷. The calculated effect sizes and *P* values from all these models were reported and compared. All the analyses were performed with the statistical software packages R3.3.2 (<http://www.R-project.org>, The R Foundation) and the Free Statistics analysis platform (Version 1.9, Beijing, China, <http://www.clinicalscientists.cn/freestatistics>).

Results

Among the 97,657 participants recruited to the NHANES from 2001 to 2020, 46,499 participants had missing data on their lung cancer status and PhenoAge measurements. The percentages of missing data are summarized in Table 1. The rate of missing BMI data was 1.8% (*n* = 726). For smoking status, 2.9% of the data were missing (*n* = 1179). Drinking status had the highest proportion of missing data, with 10.5% missing values (*n* = 4295). Education level had a negligible missing data rate of 0.1% (*n* = 41). History of emphysema had a missing data rate of 3.5% (*n* = 1437). A total of 10,315 participants aged 18 years or younger and 78 people with missing death data were excluded, resulting in 40,765 people being included in the analysis (Fig. 1). The mean age of the participants included in our analyses was 48.3 years (SD 18.7), and 20,046 (51.4%) were male (Table 1).

Overall, participants with lung cancer were older, more likely to be Non-Hispanic White, more likely to be former or current smokers, and more likely to be former drinkers than participants without lung cancer were (*p* < 0.001 for all comparisons). Individuals with a history of emphysema also had a higher incidence of lung cancer than the general population did. There were 10,874 (26.7%) participants who were classified as having PhenoAge acceleration, and those with accelerated biological aging had a greater incidence of lung cancer than those without accelerated biological aging.

Cross-sectional analyses revealed statistically significant associations between biological aging and the prevalence of lung cancer for each model. Each 1-year increase in PhenoAge acceleration increased the risk of developing lung cancer by 4% (*p* < 0.001, Table 2, Model 2).

Individuals with accelerated aging had a greater risk of developing lung cancer than nonaccelerated aging individuals did, with odds ratios of 2.60 (full adjustment 95% CI 1.59–4.22, *p* < 0.001). According to the fully adjusted models, participants in the Q4 subgroup had a 334% greater risk of developing lung cancer.

Kaplan–Meier analyses of individuals diagnosed with lung cancer revealed a greater risk of all-cause mortality for those with accelerated aging than for those with nonaccelerated aging (log rank test *P* = 0.028, Supplementary_Figure 1). Among lung cancer patients, the multivariate Cox proportional hazards model results, adjusted for potential confounders, revealed that each 1-unit increase in PhenoAge acceleration was associated with a 12% increase in all-cause mortality (95% CI 1.04–1.20, *P* = 0.003) (Table 3, Model 2).

Moreover, individuals in the Q4 quartile presented a greater incidence of all-cause mortality (HR = 11.82, 95% CI 1.48–94.3, *P* = 0.02) than those in the Q1 quartile did, following adjustment for potential confounding variables. The mortality rate was 219% higher in the accelerated aging cohort than in the nonaccelerated aging cohort (95% CI 1.28–7.94, *P* = 0.013) in Model 2.

The estimated dose-response curves revealed a statistically significant linear correlation between PhenoAge acceleration and both lung cancer risk (Supplementary_Figure 2, *P* value for nonlinearity = 0.22) and all-cause mortality (Supplementary_Figure 3, *P* value for nonlinearity < 0.001). These analyses were adjusted for multiple variables using RCS regression and excluded the highest and lowest 0.5% of the PhenoAge acceleration measurements to ensure robustness. As PhenoAge acceleration increases, both lung cancer risk and mortality show a noticeable upward trend, suggesting that greater biological aging is associated with an increased risk of developing lung cancer and a greater likelihood of death in individuals with lung cancer.

In cross-sectional analyses, the correlation between PhenoAge acceleration and lung cancer was stronger in male participants than in their female counterparts. Additionally, this correlation was more pronounced in individuals without a history of emphysema than in those with a history of emphysema (Fig. 2). The associations between PhenoAge acceleration and the risk of developing lung cancer were more pronounced in Non-Hispanic White and Mexican American individuals and in those with BMIs between 25 and 30 kg/m² than in those with other BMIs. No statistically significant interaction was observed in the cross-sectional analysis between the other stratified variables. Sensitivity analyses revealed that the results remained consistent and robust after excluding participants with PhenoAge acceleration values in the upper and lower 1% extremes (as detailed in Supplementary_Table 1). Furthermore, repeating the main analyses without considering complex sampling designs (as detailed in Supplementary_Table 2) yielded similar results. After PSM, a statistically significant positive association was identified between PhenoAge acceleration and lung cancer (OR = 2.32, 95% CI 1.41–3.83, *p* = 0.001). The results remained consistent when weighting analysis was employed with SMRW, PA, and OW (Supplementary_Table 3). For the analysis of all-cause mortality, PSM revealed that PhenoAge acceleration was

	Overall (<i>n</i> = 40765)	No lung cancer (<i>n</i> = 40674)	Prevalent lung cancer (<i>n</i> = 91)	<i>p</i>
Age, years (mean (SD))	48.3 (18.7)	48.3 (18.7)	68 (13.9)	< 0.001
Sex (%)				
Female	19,819 (48.6)	19,766 (48.6)	53 (58.2)	0.066
Male	20,946 (51.4)	20,908 (51.4)	38 (41.8)	
Race or ethnicity (%)				
Non-Hispanic White	18,407 (45.2)	18,349 (45.1)	58 (63.7)	< 0.001
Non-Hispanic Black	8286 (20.3)	8264 (20.3)	22 (24.2)	
Mexican American	7206 (17.7)	7205 (17.7)	1 (1.1)	
Other	6866 (16.8)	6856 (16.9)	10 (11.0)	
BMI, kg/m ² (%)				
≤ 25	12,386 (30.4)	12,355 (30.4)	31 (34.1)	0.822
> 25 to < 30	13,352 (32.8)	13,322 (32.8)	30 (33.0)	
≥ 30	14,301 (35.1)	14,272 (35.1)	29 (31.9)	
Missing	726 (1.8)	725 (1.8)	1 (1.1)	
Smoking status (%)				
Never	21,454 (52.6)	21,443 (52.7)	11 (12.1)	< 0.001
Former	9754 (23.9)	9688 (23.8)	66 (72.5)	
Current	8378 (20.6)	8364 (20.6)	14 (15.4)	
Missing	1179 (2.9)	1179 (2.9)	0 (0)	
Drinking status (%)				
Never	5452 (13.4)	5446 (13.4)	6 (6.6)	< 0.001
Current light/moderate drinker	16,846 (41.3)	16,815 (41.3)	31 (34.1)	
Current heavier drinker	7236 (17.8)	7226 (17.8)	10 (11.0)	
Former	6936 (17.0)	6899 (17.0)	37 (40.7)	
Missing	4295 (10.5)	4288 (10.5)	7 (7.7)	
Education (%)				
Less than high school	10,932 (26.8)	10,904 (26.8)	28 (30.8)	0.014
High school or equivalent	9546 (23.4)	9523 (23.4)	23 (25.3)	
College and above	20,246 (49.7)	20,207 (49.7)	39 (42.9)	
Missing	41 (0.1)	40 (0.1)	1 (1.1)	
Poverty (mean (SD))	2.5 (1.6)	2.5 (1.6)	2.3 (1.3)	0.172
History of emphysema (%)				
No	38,521 (94.5)	38,456 (94.5)	65 (71.4)	< 0.001
Yes	807 (2.0)	782 (1.9)	25 (27.5)	
Missing	1437 (3.5)	1436 (3.5)	1 (1.1)	
PhenoAge, years (mean (SD))	−2.1 (6.2)	−2.1 (6.2)	2 (7.0)	< 0.001
PhenoAge, nonaccelerated aging	29,891 (73.3)	29,853 (73.4)	38 (41.8)	< 0.001
PhenoAge, accelerated aging	10,874 (26.7)	10,821 (26.6)	53 (58.2)	

Table 1. Baseline characteristics of participants by lung cancer status. Data are *n* (%), mean (SD). PhenoAge = phenotypic age.

significantly associated with increased all-cause mortality (HR = 2.27, 95% CI 2.18–2.36; $P < 0.001$). Weighted Cox regression models yielded consistent results, with HRs ranging from 1.91 to 1.93 across the IPTW, SMRW, PA, and OW methods (Supplementary_Table 4). Additionally, when excluding participants with missing data, the results remained unchanged, confirming the robustness of our findings (Supplementary_Table 5).

Discussion

In this study, we explored the associations between accelerated biological aging and both lung cancer risk and all-cause mortality among individuals with lung cancer via data from the NHANES cohort. Our findings indicate that PhenoAge acceleration, a clinical measure of biological aging, is significantly associated with an increased risk of developing lung cancer and increased all-cause mortality in lung cancer patients.

The relationship between aging and lung cancer risk has been postulated to be attributable to three primary causal factors: first, the progressive accumulation of unresolved damage resulting from exposure to carcinogenic substances (such as nicotine)²⁸; second, the age-related deterioration in immune function⁹; and third, augmented cellular senescence²⁹. Furthermore, oxidative stress has been demonstrated to alter the function of DNA repair mechanisms through epigenetic modifications. This process is also associated with premature aging, which can be attributed to telomere shortening caused by impaired DNA repair. Notably, telomere shortening

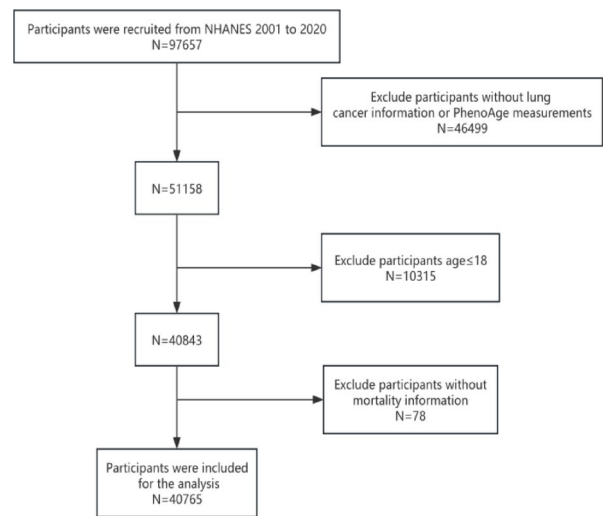


Fig. 1. Flowchart of participant selection from the US NHANES.

	Crude model			Model 1			Model 2		
	OR	CI	P value	OR	CI	P value	OR	CI	P value
PhenoAge acceleration									
Continuous	1.06	1.05–1.07	<0.001	1.05	1.04–1.07	<0.001	1.04	1.02–1.06	<0.001
Q1	Reference			Reference			Reference		
Q2	1.8	0.58–5.54	0.303	1.86	0.60–5.73	0.276	1.74	0.58–5.26	0.323
Q3	2.28	0.75–6.93	0.147	2.12	0.71–6.31	0.176	1.88	0.59–6.04	0.283
Q4	8.11	3.08–21.34	<0.001	5.82	2.21–15.32	<0.001	4.34	1.48–12.75	0.008
P for trend			<0.001			<0.001			0.002
Nonaging	Reference			Reference			Reference		
Aging	4.55	2.95–7.01	<0.001	3.29	2.13–5.08	<0.001	2.6	1.59–4.22	<0.001

Table 2. Association between biological aging and the risk of developing lung cancer. Crude model: nonadjusted model; Model 1: adjusted for age, sex, race; Model 2: adjusted for covariates of Model 1, and smoking status, drinking status, education level, body mass index, history of emphysema, and family income-to-poverty ratio. Nonaging: PhenoAge acceleration ≤ 0. Aging: PhenoAge acceleration > 0. Abbreviations: Q, quartile.

	Crude model			Model 1			Model 2		
	HR	CI	P value	HR	CI	P value	HR	CI	P value
PhenoAge acceleration									
Continuous	1.1	1.04–1.17	<0.001	1.09	1.03–1.15	0.002	1.12	1.04–1.20	0.003
Q1	Reference			Reference			Reference		
Q2	2.36	0.92–6.05	0.073	1.95	0.73–5.22	0.185	8.42	1.26–56.40	0.028
Q3	2.93	1.12–7.70	0.029	2.37	0.80–7.03	0.121	7.07	1.14–43.82	0.036
Q4	4.9	1.76–13.67	0.002	3.44	1.24–9.54	0.018	11.82	1.48–94.35	0.02
P for trend			0.002			0.019			0.019
Nonaging	Reference			Reference			Reference		
Aging	2.18	1.13–4.22	0.02	1.81	0.88–3.73	0.105	3.19	1.28–7.94	0.013

Table 3. Association between biological aging and the risk of all-cause mortality among lung cancer patients. Crude model: nonadjusted model; Model 1: adjusted for age, sex, race; Model 2: adjusted for covariates of Model 1, and smoking status, drinking status, education level, body mass index, history of emphysema, and family income-to-poverty ratio. Nonaging: PhenoAge acceleration ≤ 0. Aging: PhenoAge acceleration > 0. Abbreviations: Q, quartile.

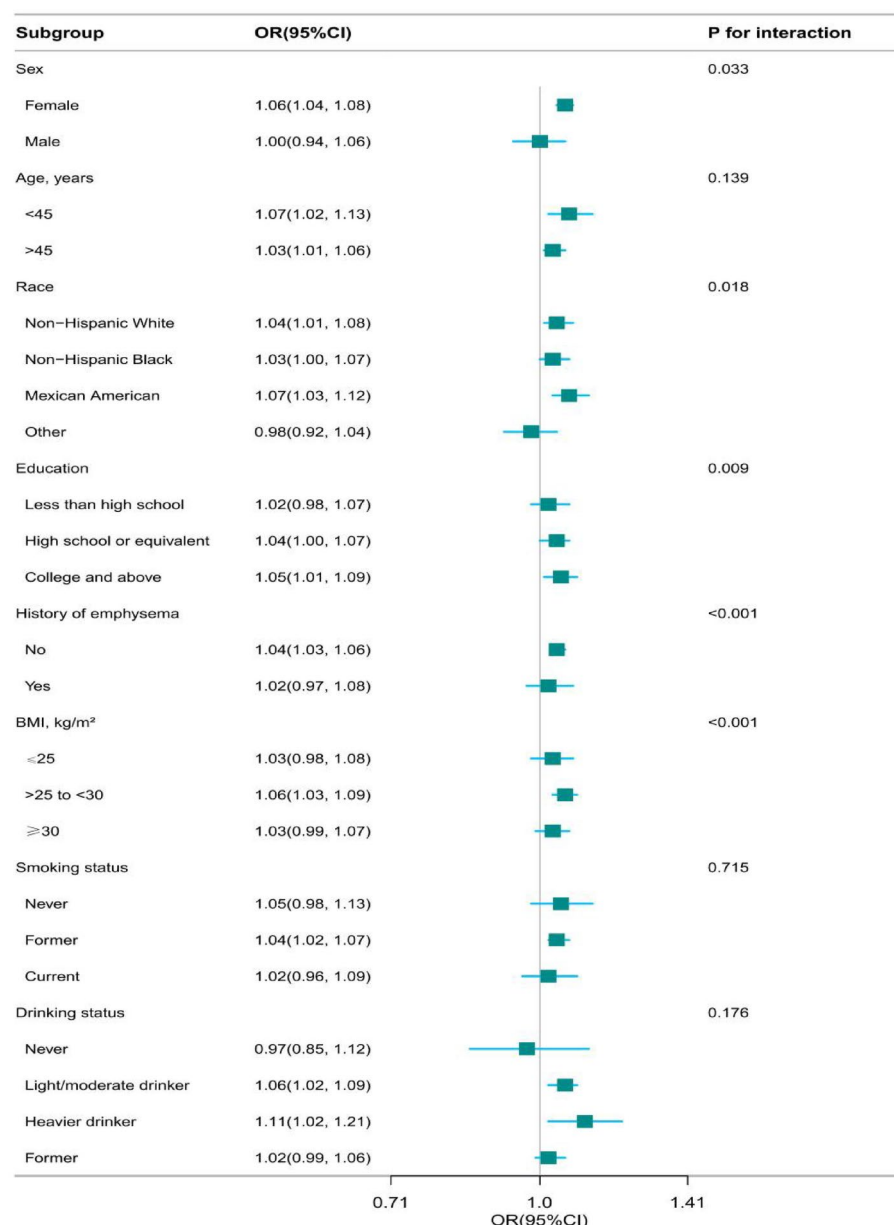


Fig. 2. Stratified analyses of the associations between biological aging and the risk of developing lung cancer.

has been linked to the development of lung cancer³⁰. The concept of ‘lung age’, derived from the first second of forced expiration, body height and age, has been demonstrated to function as an indicator of pulmonary obstructive impairment and to be significantly associated with postoperative respiratory complications and patient survival in cases of lung cancer³¹. In addition, previous studies at UK Biobank have shown that biological aging increases the risk of developing lung cancer, which is consistent with our findings^{12,17}. Biological age, as determined by the PhenoAge method, is calculated on the basis of numerous factors, including metabolism, immunity, inflammation and organ homeostasis. This approach can be used to accurately identify individuals who appear to be physiologically older than their chronological age would suggest^{32,33}. These findings reinforce the understanding that accelerated biological aging increases the risk of developing lung cancer and suggest that higher levels of accelerated biological aging could serve as valuable predictors of lung cancer development.

To date, the relationship between lung cancer and accelerated biological aging has been explored in only a limited number of studies^{12,17}. Our findings provide additional evidence linking lung cancer to biological aging and are the first to analyze how biological aging impacts mortality in lung cancer patients. The evidence suggests that aging plays a key role in the risk of death across populations^{34,35}. The process of aging is typified by a series of gradual physiological declines, a reduction in cellular repair mechanisms, and the eventual occurrence of death³⁶. The age component offers insight into the impact of physiological alterations resulting from the aging process over time³⁷. The term “biological aging” is used to describe the decline in tissue and organismal function that occurs over time. In contrast, the term “chronological aging” is used to indicate the amount of

time that has passed since an individual's birth³⁸. Therefore, it is more accurate to use biological age to predict mortality. Previous studies have shown that accelerated biological aging increases mortality in individuals with Alzheimer's disease, type 2 diabetes and multivessel coronary artery disease^{39–41}. Furthermore, individuals who were diagnosed with lung cancer and exhibited accelerated biological aging presented an increased mortality risk compared with those without lung cancer and without accelerated biological aging. Therefore, monitoring the biological aging status of individuals with and without lung cancer may be beneficial for preventing mortality.

The strengths of our study lie in the large sample size and the extended survey periods of the NHANES. These factors provide a robust foundation for our findings. Our study also has several limitations. First, owing to the cross-sectional nature of the NHANES dataset, causal relationships between accelerated biological aging and lung cancer cannot be established. It remains unclear whether accelerated aging precedes lung cancer development or if lung cancer itself accelerates biological aging. This temporal ambiguity limits the ability to draw definitive conclusions. Additionally, the study does not account for the differing rates of biological aging across individuals with the same chronological age, which could lead to variation in the relationship between biological aging and lung cancer risk. Future longitudinal studies are needed to clarify these temporal relationships. Second, our study focused on prevalent lung cancer cases, namely, individuals who were already diagnosed with lung cancer at the time of data collection. While this approach provides insights into the relationship between biological aging and lung cancer outcomes, it does not address how biological aging contributes to the onset of lung cancer. Incorporating incident cases in future research would help elucidate the role of biological aging in the early stages of lung cancer development. Third, lung cancer diagnoses in the NHANES are based on self-reported data, which may introduce reporting bias and affect the accuracy of the results. Additionally, NHANES lacks detailed clinical information on cancer subtype, histology, and aggressiveness, which are crucial for understanding the heterogeneity of lung cancer and its interaction with biological aging. Fourth, while PhenoAge has been shown to predict disease and mortality outcomes effectively, it may not capture the full complexity of biological aging mechanisms. Biomarkers such as DNA methylation-based measures could provide deeper insights into the molecular processes underlying aging and offer additional predictive power. Future studies integrating clinical and molecular data are necessary to improve the accuracy of biological aging assessments. Fifth, regarding missing data, we treated missing values as separate categories for categorical variables. This approach assumes that individuals with missing data are systematically different from those with complete data, which could introduce potential bias. However, we conducted sensitivity analyses to assess whether missingness influenced the study outcomes. The consistency of our findings across multiple analytic approaches suggests that missing data did not substantially impact the overall conclusions. Sixth, this study spans multiple NHANES cycles from 2001 to 2020, during which changes in healthcare policies, screening programs, and treatment advancements may have influenced both biological aging and lung cancer risk. For example, Medicaid expansion and the implementation of low-dose CT lung cancer screening could have affected early detection and disease incidence. However, our study did not adjust for these long-term population-level shifts, which could introduce unmeasured confounding. Future research should consider temporal adjustments to evaluate whether the observed associations remain stable over time. Finally, the number of lung cancer cases in our cohort was relatively small, and the results should be considered preliminary. However, the results remained consistent across multiple sensitivity analyses, including PSM, unweighted analyses, and RCS regression. These methods confirmed the robustness of the observed associations. Future studies with larger numbers of incident lung cancer cases and diverse populations are needed to confirm these findings and explore their applicability to broader demographic groups. The incorporation of molecular biomarkers and longitudinal designs would further enhance the understanding of how biological aging influences lung cancer risk and outcomes.

Conclusion

Despite these limitations, this study underscores the potential of accelerated biological aging, as measured by PhenoAge, to serve as an independent predictor of lung cancer risk and mortality. The consistent associations observed across multiple sensitivity analyses suggest that interventions aimed at slowing biological aging may offer promising strategies for reducing lung cancer risk and improving outcomes in lung cancer patients. Further research is needed to establish causal links and identify the mechanisms driving these associations, paving the way for targeted prevention and treatment approaches.

Data availability

Publicly available datasets were analyzed in this study. This data can be found here: <https://wwwn.cdc.gov/nchs/nhanes/>.

Received: 8 November 2024; Accepted: 25 April 2025

Published online: 23 May 2025

References

- Gavhane, S. B. Review: Advances in lung cancer research. *Int. J. Res. Appl. Sci. Eng. Technol.* **11**, 921–929 (2023).
- Liu, S., Wei, W., Wang, J. & Chen, T. Theranostic applications of selenium nanomedicines against lung cancer. *J. Nanobiotechnol.* **21**, 96 (2023).
- Acevedo, A. M. I. symptom management in early stage lung cancer. *J. Thorac. Oncol.* **14**, S1156–S1157 (2019).
- Abazari, M. et al. Estimation of survival rates in patients with lung cancer in West Azerbaijan, the Northwest of Iran. *Asian Pac. J. Cancer Prev. APJCP.* **16**, 3923–3926 (2015).
- Abedi, S. et al. Estimating the survival of patients with lung cancer: What is the best statistical model? *J. Prev. Med. Public. Health Yebang Uihakhoe Chi.* **52**, 140–144 (2019).
- North, B. J. & Sinclair, D. A. The intersection between aging and cardiovascular disease. *Circ. Res.* **110**, 1097–1108 (2012).

7. Angelidis, I. et al. An atlas of the aging lung mapped by single cell transcriptomics and deep tissue proteomics. *Nat. Commun.* **10**, 963 (2019).
8. Abu Bakar, S. A., Mohamed Shahrudin, S., Ismail, S. N. S. & Wan Md Adnan, W. A. H. Biological age for chronic kidney disease patients using index model. *PeerJ* **10**, e13694 (2022).
9. Cho, S. J. & Stout-Delgado, H. W. Aging and lung disease. *Annu. Rev. Physiol.* **82**, 433–459 (2020).
10. Gao, X. et al. Accelerated biological aging and risk of depression and anxiety: Evidence from 424,299 UK biobank participants. *Nat. Commun.* **14**, 2277 (2023).
11. Levine, M. E. et al. An epigenetic biomarker of aging for lifespan and healthspan. *Aging* **10**, 573–591 (2018).
12. Mak, J. K. L. et al. Clinical biomarker-based biological aging and risk of cancer in the UK biobank. *Br. J. Cancer.* **129**, 94–103 (2023).
13. Dugué, P. A. et al. DNA methylation-based biological aging and cancer risk and survival: Pooled analysis of seven prospective studies. *Int. J. Cancer* **142**, 1611–1619 (2018).
14. Hofman, P. et al. The oncoage consortium: Linking aging and oncology from bench to bedside and back again. *Cancers* **11**, 250 (2019).
15. Dugué, P. A. et al. Biological aging measures based on blood DNA methylation and risk of cancer: A prospective study. *JNCI Cancer Spectr.* **5**, pkaa109 (2021).
16. Lau, C. H. E. & Robinson, O. DNA methylation age as a biomarker for cancer. *Int. J. Cancer.* **148**, 2652–2663 (2021).
17. Ma, Z. et al. Association between biological aging and lung cancer risk: Cohort study and Mendelian randomization analysis. *iScience* **26**, 106018 (2023).
18. Song, M. A. et al. Accelerated epigenetic age, inflammation, and gene expression in lung: Comparisons of smokers and vapers with non-smokers. *Clin. Epigenet.* **15**, 160 (2023).
19. Na, L. et al. Feasibility of reidentifying individuals in large National physical activity data sets from which protected health information has been removed with use of machine learning. *JAMA Netw. Open* **1**, e186040 (2018).
20. Sun, Y. et al. Association of seafood consumption and mercury exposure with cardiovascular and All-Cause mortality among US adults. *JAMA Netw. Open* **4**, e2136367 (2021).
21. Kwon, D. & Belsky, D. W. A toolkit for quantification of biological age from blood chemistry and organ function test data: BioAge. *GeroScience* **43**, 2795–2808 (2021).
22. Zhao, M., Veeranki, S. P., Magnussen, C. G. & Xi, B. Recommended physical activity and all cause and cause specific mortality in US adults: Prospective cohort study. *BMJ* **370**, 2031 (2020).
23. Austin, P. C. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivar. Behav. Res.* **46**, 399–424 (2011).
24. Austin, P. C. & Stuart, E. A. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Stat. Med.* **34**, 3661–3679 (2015).
25. Sato, T. & Matsuyama, Y. Marginal structural models as a tool for standardization. *Epidemiol. Camb. Mass.* **14**, 680–686 (2003).
26. Li, L. & Greene, T. A weighting analogue to pair matching in propensity score analysis. *Int. J. Biostat.* **9**, 215–234 (2013).
27. Li, F., Morgan, K. L. & Zaslavsky, A. M. Balancing covariates via propensity score weighting. *J. Am. Stat. Assoc.* **113**, 390–400 (2018).
28. López-Otin, C., Blasco, M. A., Partridge, L., Serrano, M. & Kroemer, G. The hallmarks of aging. *Cell* **153**, 1194–1217 (2013).
29. Calcinotto, A. et al. Cellular senescence: Aging, cancer, and injury. *Physiol. Rev.* **99**, 1047–1078 (2019).
30. Sears, C. R. et al. Xeroderma pigmentosum group C deficiency alters cigarette smoke DNA damage cell fate and accelerates emphysema development. *Am. J. Respir. Cell. Mol. Biol.* **58**, 402–411 (2018).
31. Haruki, T. et al. Lung age predicts post-operative complications and survival in lung cancer patients. *Respirol. Carlton Vic.* **15**, 495–500 (2010).
32. Belsky, D. W. et al. Quantification of biological aging in young adults. *Proc. Natl. Acad. Sci. U. S. A.* **112**, E4104–E4110 (2015).
33. Brown, P. J. et al. Biological age, not chronological age, is associated with late-life depression. *J. Gerontol. Biol. Sci. Med. Sci.* **73**, 1370–1376 (2018).
34. Berlanga-Acosta, J. A. et al. Cellular senescence as the pathogenic hub of diabetes-related wound chronicity. *Front. Endocrinol.* **11**, 573032 (2020).
35. Tsai, H. Z., Lin, R. K. & Hsieh, T. S. Drosophila mitochondrial topoisomerase III alpha affects the aging process via maintenance of mitochondrial function and genome integrity. *J. Biomed. Sci.* **23**, 38 (2016).
36. Cruz-Topete, D., Oakley, R. H. & Cidlowski, J. A. Glucocorticoid signaling and the aging heart. *Front. Endocrinol.* **11**, 347 (2020).
37. Swerdel, J. N. et al. Ischemic stroke rate increases in young adults: Evidence for a generational effect? *J. Am. Heart Assoc.* **5**, e004245 (2016).
38. Hamczyk, M. R., Nevado, R. M., Barettino, A., Fuster, V. & Andrés, V. Biological versus chronological aging: JACC focus seminar. *J. Am. Coll. Cardiol.* **75**, 919–930 (2020).
39. Gonzales, M. M. et al. Biological aging processes underlying cognitive decline and neurodegenerative disease. *J. Clin. Invest.* **132**, e158453 (2022).
40. Ma, Q. et al. Association between phenotypic age and mortality in patients with multivessel coronary artery disease. *Dis. Markers* **2022**, 4524032 (2022).
41. Sabbatinelli, J. et al. DNA Methylation-derived biological age and long-term mortality risk in subjects with type 2 diabetes. *Cardiovasc. Diabetol.* **23**, 250 (2024).

Acknowledgements

We would like to express our gratitude to the participants and researchers who participated in the NHANES. The use of the Free Statistics analysis platform makes it easier for us to explore the NHANES database.

Author contributions

Ming Du designed the study. Ming Du and Qihang Xie supervised the study. Qihang Xie, Haoran Qu, Siyu Xie, Wenhao Li, Rui Ouyang, and Chengxiang Zhang extracted the data, interpreted the study results, and wrote the first draft of the manuscript. Ming Du revised the manuscript.

Funding

No funding was received for this study.

Declarations

Competing interests

The authors declare no competing interests.

Ethical statement

Data for this study were sourced from datasets available to the public with the approval of the National Center for Health Statistics' research ethics review committees.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-00114-2>.

Correspondence and requests for materials should be addressed to M.D.

Reprints and permissions information is available at www.nature.com/reprints.

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

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

© The Author(s) 2025