Environment International 178 (2023) 108041

Contents lists available at ScienceDirect

Environment International

journal homepage: www.elsevier.com/locate/envint

Full length article

Household air pollution and epigenetic aging in Xuanwei, China

Batel Blechter [a,*,1], Andres Cardenas [b,1], Junming Shi [c,1], Jason Y.Y. Wong [a,2], Wei Hu [a], Mohammad L. Rahman [a], Charles Breeze [a], George S. Downward [d,e], Lützen Portengen [d], Yongliang Zhang [d], Bofu Ning [f], Bu-Tian Ji [a], Richard Cawthon [g], Jihua Li [h], Kaiyun Yang [i], Anne Bozack [b], H. Dean Hosgood [j], Debra T. Silverman [a], Yunchao Huang [i,3], Nathaniel Rothman [a,3], Roel Vermeulen [c,3], Qing Lan [a,3]

[a] Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Rockville, MD, USA
[b] Department of Epidemiology and Population Health, Stanford University, Stanford, CA, USA
[c] Department of Biostatistics, UC Berkeley School of Public Health, Berkeley, CA, USA
[d] Institute for Risk Assessment Sciences, Division of Environmental Epidemiology, Utrecht University, Utrecht, Netherlands
[e] Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, Netherlands
[f] Xuanwei Center of Diseases Control, Xuanwei, Yunnan, China
[g] Department of Human Genetics, University of Utah School of Medicine, Salt Lake City, UT, USA
[h] Quijing Center for Diseases Control and Prevention, Quijing, Yunnan, China
[i] Department of Cardiothoracic Surgery, Third Affiliated Hospital of Kunming Medical University (Yunnan Cancer Hospital), Kunming, Yunnan, China
[j] Division of Epidemiology, Albert Einstein College of Medicine, New York, NY, USA

ARTICLE INFO

Handling Editor: Marti Nadal

Keywords:
Household air pollution
Polycyclic aromatic hydrocarbons
Epigenetic age acceleration
5-methylchrysene
Lung cancer

ABSTRACT

Background: Household air pollution (HAP) from indoor combustion of solid fuel is a global health burden linked to lung cancer. In Xuanwei, China, lung cancer rate for nonsmoking women is among the highest in the world and largely attributed to high levels of polycyclic aromatic hydrocarbons (PAHs) that are produced from combustion of smoky (bituminous) coal used for cooking and heating. Epigenetic age acceleration (EAA), a DNA methylation-based biomarker of aging, has been shown to be highly correlated with biological processes underlying the susceptibility of age-related diseases. We aim to assess the association between HAP exposure and EAA.

Methods: We analyzed data from 106 never-smoking women from Xuanwei, China. Information on fuel type was collected using a questionnaire, and validated exposure models were used to predict levels of 43 HAP constituents. Exposure clusters were identified using hierarchical clustering. EAA was derived for five epigenetic clocks defined as the residuals resulting from regressing each clock on chronological age. We used generalized estimating equations to test associations between exposure clusters derived from predicted levels of HAP exposure, ambient 5-methylchrysene (5-MC), a PAH previously found to be associated with risk of lung cancer, and EAA, while accounting for repeated-measurements and confounders.

Results: We observed an increase in GrimAge EAA for clusters with 31 and 33 PAHs reflecting current ($\beta = 0.77$ y per standard deviation (SD) increase, 95 % CI:0.36,1.19) and childhood ($\beta = 0.92$ y per SD, 95 % CI:0.40,1.45) exposure, respectively. 5-MC (ng/m^3-year) was found to be associated with GrimAge EAA for current ($\beta = 0.15$ y, 95 % CI:0.05,0.25) and childhood ($\beta = 0.30$ y, 95 % CI:0.13,0.47) exposure.

Conclusions: Our findings suggest that exposure to PAHs from indoor smoky coal combustion, particularly 5-MC, is associated with GrimAge EAA, a biomarker of mortality.

* Corresponding author.
E-mail address: batel.blechter@nih.gov (B. Blechter).
[1] These authors contributed equally.
[2] This author is currently at the Epidemiology and Community Health Branch, National Heart Lung and Blood Institute. This work was conducted while formerly at the Division of Cancer Epidemiology and Genetics, National Cancer Institute.
[3] These authors co-supervised this work.

https://doi.org/10.1016/j.envint.2023.108041
Received 6 February 2023; Received in revised form 19 May 2023; Accepted 13 June 2023
Available online 17 June 2023

1. Introduction

Household air pollution (HAP) is a major global health burden affecting about half of the world's population (VOLUME 95 Household Use of Solid Fuels and High-temperature Frying IARC Monographs on the Evaluation of Carcinogenic Risks to Humans WORLD HEALTH ORGANIZATION INTERNATIONAL AGENCY FOR RESEARCH ON CANCER). Individuals in low- and middle-income countries are exposed to HAP through incomplete combustion caused by domestic cooking and/or heating with solid fuels (coal, biomass, wood)(Barone-Adesi, 2012; Lan et al., 2002; Sisti and Boffetta, 2012). It is estimated that over 450 million individuals still use solid fuels in Mainland China(Effects Institute, 2020), and rural counties such as Xuanwei and Fuyuan experience some of the highest lung cancer incidence and mortality rates among never-smoking women(Barone-Adesi, 2012; Mumford, 1987; Chapman, 1988). Individuals in Xuanwei and Fuyuan are exposed to hazardous levels of various toxic constituents, such as polycyclic aromatic hydrocarbons (PAHs) from combustion of smoky (bituminous) coal used for cooking and heating in the home (Lan et al., 2002; Mumford, 1987; Downward, 2014).

PAHs are organic compounds that are emitted from the combustion of fuels and are known for their carcinogenic and genotoxic characteristics(Moorthy et al., 2015). We recently found evidence that cumulative exposure to PAHs, particularly 5-methylchrysene (5-MC), is strongly associated with increased risk of lung cancer in a comprehensive epidemiologic study of HAP constituents in Xuanwei(Vermeulen et al., 2019). However, the mechanisms of HAP in the pathogenesis of lung cancer among never-smokers are unclear.

Epigenetic changes can alter gene expression levels without changing the underlying DNA sequence. These changes include DNA methylation (DNAm) and have been shown to be an important pathway through which environmental factors, such as air pollutants, exert their effects(Alfano, 2018). Diseases such as lung cancer are often associated with epigenetic modifications(Weinhold, 2006; Baglietto, 2017; Fasanelli, 2015), and exposure to environmental toxins may result in epigenetic changes, impacting gene expression and disease risk(Breton and Marutani, 2014).

Aside from environmental exposures, aging is also considered a major risk factor for lung cancer risk, which may act through age-related decline in immune function(Derhovanessian et al., 2008); cellular senescence(Rodier and Campisi, 2011) and accumulation of DNA damage from carcinogens(López-Otín et al., 2013). As the rate by which these changes occur varies across individuals, it is believed that chronological age may not be the ideal marker to capture this variability. Epigenetic age acceleration (EAA) calculated from epigenetic clocks is a DNAm-based biomarker for aging that has been shown to be highly correlated with biological processes underlying the susceptibility to age-related diseases, such as lung cancer(Li et al., 2022; Dugué, 2021; Levine, 2015).

We recently found that EAA calculated from the GrimAge clock, a strong biomarker of mortality, was associated with increased risk of lung cancer in a prospective cohort study of never-smoking women in Shanghai, China(Rahman et al.). Further, EAA clocks have also been correlated with environmental toxins, including PAHs(Li, 2018) and particulate matter with aerodynamic diameter $\leq$ 2.5 μm ($PM_{2.5}$)(Nwanaji-Enwerem, 2017). To further explore whether DNAm is influenced by indoor ambient air pollutants, we investigated the association of exposure to combustion emissions from solid fuel and its constituents with EAA among never-smoking women from a study in Xuanwei, China.

2. Material and methods

2.1. Study population and design

The study population has been previously described in detail(Hu, 2014; Wong, 2017). Briefly, in the Xuanwei Exposure Assessment Study, we collected data on household air pollutants and exposures that may be related to the combustion of solid fuels through cooking or heating in the home. The study enrolled a total of 163 healthy never smoking women between August 2008 and June 2009 from 30 villages across Xuanwei and Fuyuan counties in the Yunnan province. The following criteria were used to select five households in each village: 1) having a stove that utilized solid fuel; 2) the residence was>10-years-old; 3) household used the same equipment for cooking or heating for the last 5 years; and 4) household included a healthy never-smoking woman aged 20–80 years who was responsible for cooking. Two air measurements were taken 24 h apart and whole blood samples were collected on the second day. Written informed consent was provided by all participants. This study was approved by the National Cancer Institute Special Studies Institutional Review Board (#06CN092).

2.2. Solid fuel use

In-person interviews were conducted by two trained study staff and information on demographic, anthropometric and household characteristics was collected. An activity questionnaire was used to record women's household activity during the measurements. The questionnaire collected information on household stove and ventilation type, cooking activities, heating practices, type of coal mine that supplied household fuel and fuel usage. Fuel used at measurement was categorized as smokeless coal, smoky coal, and wood and/or plant, and fuel used during the participant's childhood was categorized as smokeless coal, smoky coal, wood, or mixed fuel (combination of coal, wood and plant material).

2.3. Estimation of individual household air pollutant exposures

Exposure assessment has been previously described in detail(Vermeulen et al., 2019; Hu, 2014; Downward, 2016; Seow, 2016; Rousseeuw, 1987). Briefly, measurements of PAHs, $PM_{2.5}$, black carbon (BC), sulfur dioxide (SO_2) and nitrogen dioxide (NO_2) were collected over two sequential 24-hour periods, with about half of subjects visited in a second season to allow for seasonal adjustments. PAHs were measured using personal and indoor measurements, and particulate matter was collected using a 37 mm Teflon filter. Self-reported information on stove use, fuel, and mine from which coal was sourced was collected for each participant. Supervised stepwise predictive linear mixed-effect models were used to predict the annual average exposure of each pollutant and applied to self-reported histories of stove and fuel use, treating village and individual subject as random effects. Variables used as fixed effects in the final predictive model included fuel usage, stove design, room volume and season in which measurements were conducted. For each model predicting the individual air pollutants, overall goodness of fit was determined through the Akaike information criterion (AIC), and the ratio of the variance of the predicted values over the variance of the observed values (R^2) was calculated. The 43 imputed individual air pollutants and their corresponding R^2 values are listed in Supplementary Table 1.

2.4. Derivation of cluster prototypes

To account for the strong correlation between individual pollutants, we derived exposure prototypes by clustering as previously described (Vermeulen et al., 2019). Briefly, a hierarchical cluster analysis identified clusters of the 43 individual air pollutants, for which we used standard Euclidean distances and the complete linkage method to determine the cluster sequence. The number of clusters exacted was based on the silhouette score(Aryee, 2014), where a larger number of clusters was favored to better resemble the individual pollutants belonging to each cluster. A cluster prototype score for each of the clusters was derived using the first component score coefficient of a

principal component analysis, and each score was mean centered and scaled. Clusters were derived for different exposure timepoints, including current (i.e., at measurement), childhood (i.e., age 0–18) and cumulative exposures (i.e., lifetime). Six cluster prototypes were generated for current pollutant exposure, including a cluster of 31 PAHs (PAH31), a cluster of BC and 6 PAHs (BC & PAH6), a cluster of $PM_{2.5}$ and retene ($PM_{2.5}$ & RET), as well as individual clusters for NO_2, SO_2 and Naptho(*2,3,k*)fluoranthene (NkF). Five cluster prototypes were generated for childhood exposure, including a cluster of 33 PAHs (PAH33), a cluster of $PM_{2.5}$, RET, BC and 4 PAHs ($PM_{2.5}$, RET, BC & PAH4), as well as individual clusters for NO_2, SO_2 and NkF. Lastly, six cluster prototypes were generated for cumulative exposure, including a cluster of 36 PAHs (PAH36), a cluster of BC, NO_2 and $PM_{2.5}$, (BC, NO_2, & $PM_{2.5}$), as well as individual clusters for dibenzo(*a,l*)pyrene (DlP), RET, NkF. A list of constituents included in each cluster is outlined in Supplementary Table 1.

2.5. DNA extraction and methylation measurements

Leukocyte genomic DNA was extracted from a sample of blood collected on the second visit. DNA was extracted from whole blood using standard procedures. Bisulfite-converted DNA samples were randomized across Infinium HumanMethylation450 BeadChip by Illumina according to the manufacturer's protocol (San Diego, CA, USA) and DNAm levels at > 485,000 cytosine-phosphoguanine (CpG) sites were quantified. IDAT files generated from Illumina were processed in R using the minfi package(Tian, 2017). We used the ChAMP pipeline(Fortin, 2014) for quality control using default parameters. We removed samples that performed poorly based on detection p-values < 0.01 or with intensities < 10.5. We checked for concordance between predicted and annotated sex. Lastly, we used functional normalization to remove technical variability(Horvath, 2013).

2.6. Calculation of epigenetic aging biomarkers

We calculated clocks and EAA using the Horvath online calculator (https://dnamage.genetics.ucla.edu/) for the following DNA methylation clocks: the Horvath Pan Tissue(Horvath, 2018), Horvath Skin-Blood(Hannum, 2013), Hannum Blood(Levine, 2018), PhenoAge(Lu, 2019), and GrimAge(Horvath, 2016) clocks. We derived the outcome of interest, EAA, as the residuals resulting from regressing each clock on the chronological age of each participant, where a positive EAA indicated that the estimated epigenetic age is higher than the chronological age, suggesting increased biological aging. We also tested Intrinsic EAA (IEAA) and Extrinsic EAA (EEAA), representing intrinsic cellular aging and age-related changes in blood cell counts derived from the Horvath Pan Tissue and Hannum Blood clocks, respectively(R: a language and environment for statistical computing).

2.7. Statistical analysis

We assessed the association between categorical fuel types at each exposure time point and EAA biomarkers using generalized estimating equations (GEE) accounting for repeated measurements. We set smokeless coal use as the reference category for fuel type used at measurement and fuel used at childhood. We further used GEE to assess the association between each exposure cluster, 5-methylchrysene (5-MC) (an individual PAH constituent that has been found to have the strongest association with lung cancer risk in previous studies in Xuanwei independent of other PAHs (Vermeulen et al., 2019)) and each EAA biomarker. All models were adjusted for chronological age, county (Xuanwei or Fuyuan), body mass index (BMI; kg/m^2), education (no education, attended elementary school, graduated elementary school, attended middle school or higher), and socioeconomic status (SES; no luxury items and at least one luxury item such as a bicycle, sewing machine, radio, watch, phone, motorcycle, TV set or tractor). We ran additional GEE models assessing the association between exposure clusters and EAA clocks, while also mutually adjusting for all clusters within the exposure period. To separate the effect of early life exposure from smoky coal use on EAA, we derived an adulthood exposure (age ≥ 18) to 5-MC by subtracting the childhood exposure from the cumulative exposure to 5-MC and including both exposures in a model to obtain their independent effect on EAA clocks.

Sensitivity analyses were conducted using linear regression by restricting analyses to measurements at initial visit only. We conducted additional sensitivity analyses to further adjust all models for exposure to environmental tobacco smoke (ETS), as well as for season. Information on exposure to ETS was collected by asking participants if anyone in their smoked tobacco in their household before and after marriage. We categorized early life exposure and ever exposure to ETS if a family member smoked before marriage, and if a family member smoked before and/or after marriage, respectively. Season was determined based on the date the measurement was conducted and categorized into spring, summer, fall and winter. P-values < 0.05 were considered statistically significant. All analyses were performed using the R statistical software (Gao, 2022) (version 4.2.2).

3. Results

3.1. Study population

The characteristics of the study population are shown in Table 1. A total of 106 subjects had available self-reported fuel type, imputed

Table 1
Characteristics of the Xuanwei Exposure Assessment study participants with available methylation data (N = 106).

	N[a] (%)
Age (years), Mean (SD)	56.2 (15.0)
BMI (kg/m^2), Mean (SD)	22.0 (3.46)
County	
Xuanwei	53 (50.0 %)
Fuyuan	53 (50.0 %)
Education	
No school	72 (67.9 %)
Attended elementary school	17 (16.0 %)
Graduated elementary school	13 (12.3 %)
Attended middle school or higher	4 (3.8 %)
Socioeconomic status	
No luxury items	53 (50.0 %)
At least 1 luxury item	53 (50.0 %)
Exposure to environmental tobacco smoke	
No	5 (4.7 %)
Yes	99 (93.4 %)
Missing	2 (1.9 %)
Fuel type at measurement[b]	
Smokeless coal	13 (12.3 %)
Smoky coal	82 (77.4 %)
Wood and/or plant	11 (10.4 %)
Childhood fuel type	
Smokeless coal	4 (3.8 %)
Smoky coal	43 (40.6 %)
Wood	9 (8.5 %)
Mixed fuel[c]	50 (47.2 %)
Current exposure to 5-MC, Mean (SD)	8.13 (4.14)
Childhood exposure to 5-MC, Mean (SD)	5.14 (2.81)
Cumulative exposure to 5-MC, Mean (SD)	266 (2.81)

Abbreviations: SD, standard deviation; BMI, body mass index; 5-MC, 5-methylchrysene.

[a] A total of 106 subjects with available methylation data, out of which 23 have repeated measurements.

[b] Values represent fuel type at first visit.

[c] Mixed fuel types include unspecified coal, plant, beehive, and a mix of fuel types (e.g., wood and smokeless coal).

[d]104 participants had data imputed for 5-MC, out of which 22 had repeated measurements. Units are per ng/m^3-year.

individual air pollutants, and methylation data, 23 of which had repeated measurements. The mean chronological age was 56.3 (standard deviation (SD) = 15.0) years and the mean BMI was 22.0 (SD = 3.46) kg/m^2. Half of the women resided in Xuanwei county (50.0 %), half had at least 1 luxury item (50.0 %), and most women did not graduate from elementary school (83.9 %) and were exposed to ETS (93.4 %). The majority of women used smoky coal at measurement (77.4 %) and almost half used smoky coal during childhood (40.6 %).

We observed strong pairwise correlations (r > 0.92) among all estimates of epigenetic age assessed in this study, as well as between all aging markers and chronological age ($r > 0.93$), suggesting these biomarkers perform well in our study population (Supplementary Fig. 1A). Both GrimAge and the Skin-Blood clock had the highest correlation ($r = 0.96$) with chronological age and all epigenetic aging markers were strongly correlated with each other ($r > 0.93$). Further, we observed low to moderate pairwise correlations ($r < 0.59$) among all the residual estimates of epigenetic age calculated in relation to chronological age (Supplementary Fig. 1B).

3.2. Associations between categorical fuel type and EAA clocks

We found greater GrimAge EAA comparing smoky coal use to smokeless coal use for current (β = 1.84 years, (y), 95 % confidence interval (CI): 0.59,3.09, P-value = 0.004) and childhood (β = 4.14 y, 95 % CI: 1.63, 6.64, P-value = 0.001) exposure (Table 2). No other measures of EAA were associated with categorical fuel type exposure (P-value > 0.05).

3.3. Associations between exposure clusters, 5-methylchrysene and EAA clocks

To disentangle which PAHs are driving the association between smoky coal and EAA, we tested exposure clusters. In models adjusted for demographics characteristics only, we observed an increasing relationship between current exposure clusters PAH31 (β = 0.77 y per standard deviation (SD) change, 95 % CI: 0.36, 1.19, P-value = 3E−04), BC & PAH6 (β = 0.72 y per SD, 95 % CI: 0.21, 1.23, P-value = 0.006) and GrimAge EAA (Fig. 1A, Supplementary Table 2). Similarly, we observed an increasing relationship between childhood exposure clusters PAH33 (β = 0.92 y per SD, 95 % CI: 0.40, 1.45, P-value = 0.001) and PM2.5, RET & PAH5 (β = 0.72 y per SD, 95 % CI: 0.27, 1.18, P-value = 0.002), cumulative exposure clusters PAH36 (β = 1.12 y per SD, 95 % CI: 0.63, 1.60, P-value = 6E−06), BC, NO_2 & $PM_{2.5}$, (β = 0.92 y per SD, 95 % CI: 0.24, 1.60, P-value = 0.008), NkF (β = 0.62 y per SD, 95 % CI: 0.15, 1.10, P-value = 0.01) with GrimAge EAA (Fig. 1B & C, Supplementary Table 2). Childhood exposure cluster PAH33 was also associated with increase in PhenoAge EAA (β = 1.01 y per SD, 95 % CI: 0.19,1.82, P-value = 0.015). We found a similar pattern in analyses assessing the association between exposure clusters and EAA clocks mutually adjusting for clusters within the same exposure time (Supplementary Table 3).

We observed an increasing monotonic relationship between 5-MC and GrimAge EAA for current (β = 0.15 y, 95 % CI: 0.05, 0.25, P-value = 0.003), childhood (β = 0.30 y, 95 % CI: 0.13, 0.47, P-value = 4.7×10^{-4}) and cumulative exposure (β = 0.006 y, 95 % CI: 0.003,0.009, P-value = 2.9×10^{-4}) (Fig. 2, Supplementary Table 2). In analyses separating the effect of childhood and adulthood exposure to 5-MC, we observed an association between childhood 5-MC and GrimAge EAA (β = 0.02 y, 95 % CI: 0.004,0.03, P-value = 0.01), while the association between adulthood exposure to 5-MC and GrimAge EAA was null (β = −0.001 y, 95 % CI: −0.009,0.006, P-value = 0.73). Linear regression analyses restricting to initial measurements, as well as analyses further adjusting for season, early life exposure and ever exposure to ETS yielded similar results (Supplementary Table 4).

Table 2
Association between fuel type and epigenetic age acceleration.

	Grim Age		Horvath		Hannum		Pheno Age		Skin & blood		IEAA		EEAA	
	β (95 % CI)[a]	P-value[b]	β (95 % CI)[a]	P-value[b]	β (95 % CI)[a]	P-value[b]	β (95 % CI)[a]	P-value[b]	β (95 % CI)[a]	P-value[b]	β (95 % CI)[a]	P-value[b]	β (95 % CI)[a]	P-value[b]
Current fuel type														
Smokeless coal	REF		REF		REF		REF		REF		REF		REF	
Smoky coal	1.84 (0.59, 3.09)	**0.004****	−1.64 (−4.87, 1.59)	0.32	−0.95 (−0.95, 1.53)	0.45	1.20 (−0.94, 3.33)	0.27	0.67 (−1.51, 2.84)	0.55	−1.28 (−4.16, 1.60)	0.38	−1.03 (−3.85, 1.79)	0.47
Wood and/or plant	2.36 (0.20, 4.51)	**0.032***	0.28 (−3.40, 3.96)	0.88	−1.52 (−1.52, 1.34)	0.30	1.41 (−1.60, 4.41)	0.36	0.49 (−2.77, 3.75)	0.77	0.92 (−2.17, 4.02)	0.56	−1.66 (−5.27, 1.94)	0.37
Fuel type at childhood														
Smokeless coal	REF		REF		REF		REF		REF		REF		REF	
Smoky coal	4.14 (1.63, 6.64)	**0.001****	−0.10 (−2.52, 2.33)	0.94	0.38 (0.38, 3.98)	0.84	3.73 (−1.46, 8.91)	0.16	−0.48 (−2.65, 1.69)	0.66	0.59 (−3.46, 4.63)	0.78	1.21 (−3.30, 5.73)	0.60
Wood	3.85 (1.24, 6.46)	**0.004****	0.33 (−2.91, 3.56)	0.84	−0.08 (−0.08, 3.38)	0.97	2.95 (−2.84, 8.75)	0.32	−1.60 (−4.93, 1.74)	0.35	2.39 (−2.17, 6.94)	0.30	0.43 (−4.13, 4.99)	0.85
Mixed fuel[c]	2.63 (0.20, 5.05)	**0.034***	0.38 (−2.04, 2.81)	0.76	−0.22 (−0.22, 3.13)	0.90	2.88 (−2.07, 7.84)	0.25	−0.73 (−2.74, 1.28)	0.48	1.40 (−2.62, 5.43)	0.50	0.44 (−3.77, 4.65)	0.84

Abbreviations: IEAA, Intrinsic epigenetic age acceleration; EEAA, Extrinsic epigenetic age acceleration; CI, confidence interval.
*P-value < 0.05; **P-value < 0.01; ***P-value < 0.001.
[a] Coefficients estimated separately for each exposure using generalized estimating equations accounting for repeated measures, adjusted for age, county, body mass index, education, and socioeconomic status.
[b] bolded P-values were below the false discovery rate < 0.2 threshold.
[c] Mixed fuel types include unspecified coal, plant, beehive, and a mix of fuel types (e.g., wood and smokeless coal).

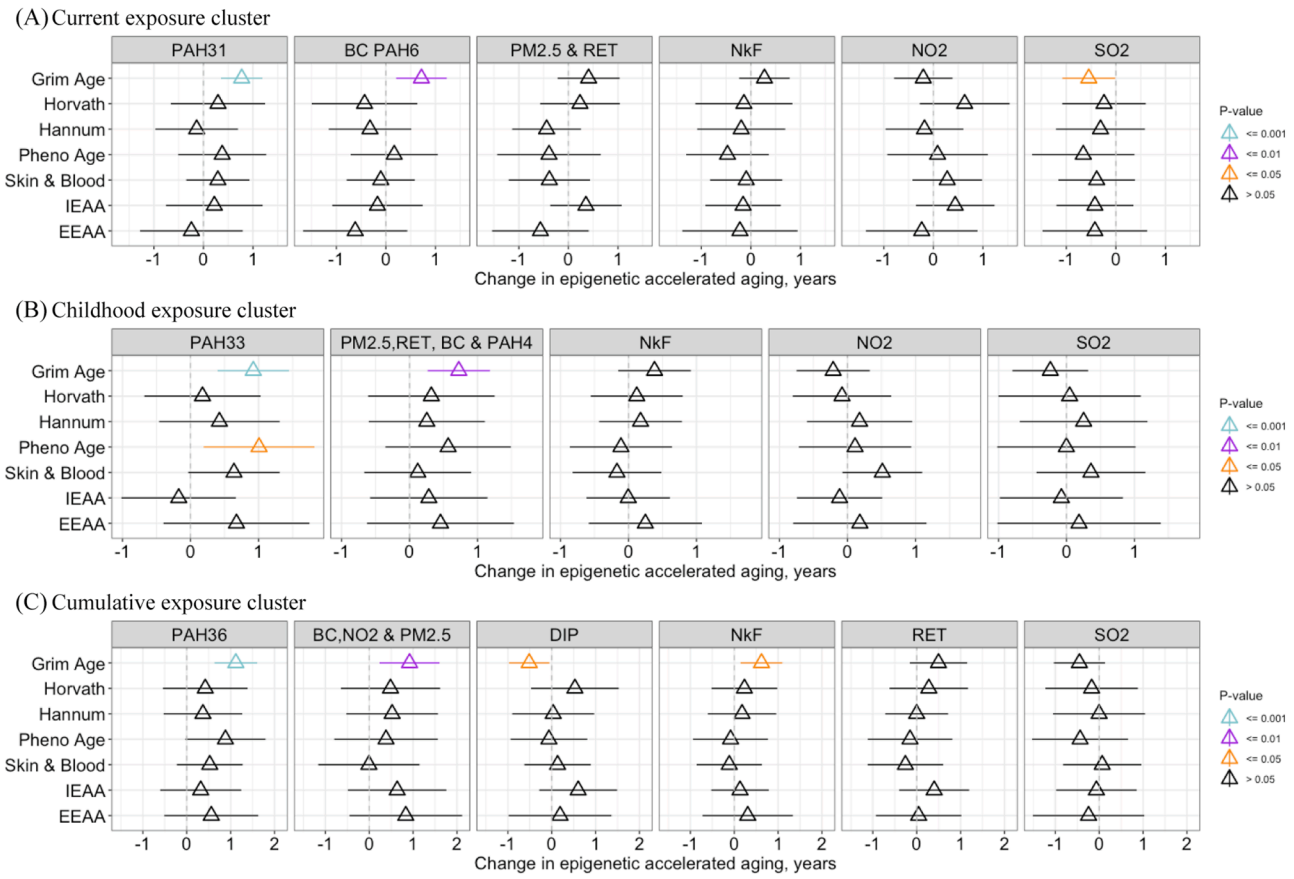

Fig. 1. Associations between clusters for current (A), childhood (B) and cumulative (C) exposures derived from imputed individual air pollutants and epigenetic age acceleration. Estimates represent change in years of epigenetic age acceleration due to increase in 1 standard deviation for each cluster and were derived using generalized estimating equations adjusting for age, county, body mass index, education, and socioeconomic status. Estimate values are available in Supplementary Table 2.

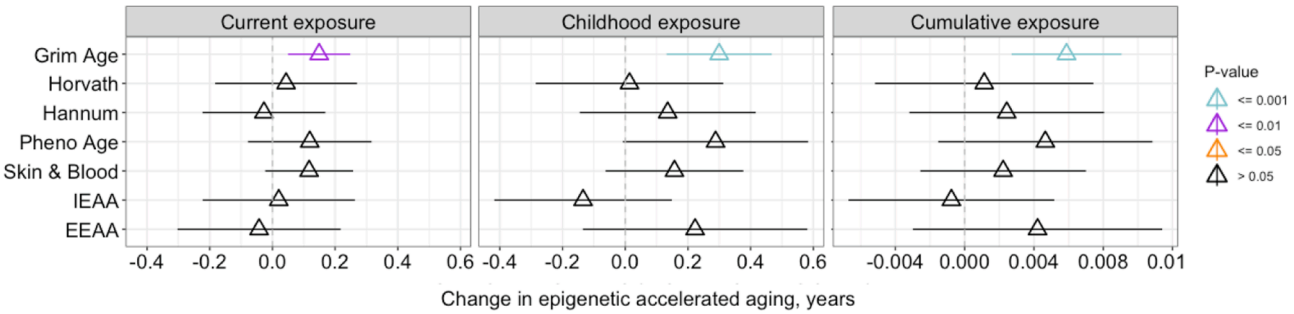

Fig. 2. Associations between current (A), childhood (B) and cumulative (C) exposure to 5-methylchrysene and epigenetic age acceleration. Estimates represent change in years of epigenetic age acceleration and were derived using generalized estimating equations adjusting for age, county, body mass index, education, and socioeconomic status. Estimate values are available in Supplementary Table 2.

4. Discussion

To investigate the relationship between coal combustion emissions, their constituents, and epigenetic aging, we conducted analyses in an epidemiologic study with detailed exposure assessment among never-smoking women in rural China. We observed an association between current, childhood and cumulative PAH clusters, and greater GrimAge EAA, a strong biomarker for mortality. Further, we found that exposure to 5-MC, an individual PAH constituent selected *a priori*, was positively associated with GrimAge EAA. While we had limited sample size, we also observed an association between current use of smoky coal as well as childhood exposure, with greater GrimAge. To our knowledge, this is the first study to find a link between indoor coal combustion emission constituents and epigenetic aging.

The GrimAge clock includes data from 1,030 CpGs that are associated with health-related plasma proteins, sex and chronological age, as well as smoking pack-years, which may be particularly relevant to understanding the GrimAge associations in our analyses (Cardenas, 2022). GrimAge has been shown to be a predictor of lifespan, including for never-smokers(Jones, 2012), as well as to be associated with time to any cancer(Oberdoerffer and Sinclair, 2007). Notably, we recently found that the GrimAge clock was associated with increased risk of lung cancer in a prospective cohort study among never-smoking women in China (Rahman et al.). Further, greater GrimAge EAA is strongly associated

with air pollutants such as PM2.5(Horvath, 2013) and smoking(Fraga and Esteller, 2007),which is not surprising given its incorporation of pack years in the training data. Interestingly, this biomarker might indeed reflect PAH exposure and risk among never smoking participants as highlighted in our study.

There has been a growing body of evidence showing that DNAm, an epigenetic modification that plays a major role in gene regulation and genomic stability (Christensen, 2009) is closely related to aging(Esteller, 2008; Schadt, 2009; Ruiz-Hernandez, 2015; Wong et al., 2019; Wong et al., 2021) and is associated with various age-related diseases, such as cancer(Blechter, 2023). Given that DNAm may be influenced by environmental factors, epigenetic age may reflect environmentally induced effects on aging- a major risk factor for cancer (Schadt, 2009). Exposure to air pollutants, an environmental risk factor for lung cancer, generates oxidative stress resulting from the release of reactive oxygen species and induces chemical damage to DNA, which may alter DNAm and expression patterns, leading to adverse health effects (Ruiz-Hernandez, 2015). Our results suggest that this highly accurate clock, GrimAge, is sensitive to ambient exposures, including during childhood, a sensitive developmental window.

We previously found that lifetime smoky coal users had a nearly 100-fold increased risk of lung cancer mortality compared with lifetime smokeless coal users(Barone-Adesi, 2012), and that the increased risk was especially pronounced in early life exposure (Wong et al., 2019). Further, to identify the specific components of smoky coal that potentially drive this excess risk, we previously identified a cluster of PAHs related to lung cancer risk, with 5-MC having the strongest association (Vermeulen et al., 2019). Subsequently, we found that coal combustion emissions, and particularly 5-MC, are associated with urinary mutagenicity (Wong et al., 2021) and markers of genomic instability, such as mitochondrial DNA(Wong et al., 2017) and Alu retroelement copy number (Blechter, 2023)). Similarly, in our current study we observed that women using smoky coal and exposed to PAHs, in particularly 5-MC, have accelerated epigenetic aging. Notably, we found that early life use of smoky coal had a greater effect on EAA compared to adult use of smoky coal, and that the effect of childhood exposure to 5-MC on current measure of EAA is independent of more recent exposure. This finding is consistent with our case-control (Vermeulen et al., 2019; Wong et al., 2019) and cohort studies (Portengen et al., 2023), which showed that exposure to cumulative 5-MC before the age of 18 was associated with a higher lung cancer risk compared to adulthood exposure. Taken together with a previous study linking exposure to air pollutants with epigenetic age,(Li, 2018) our findings suggest that exposure to PAHs, in particular during childhood, may be a contributor to accelerated aging.

Our study has a number of strengths. First, the subjects in the study had comprehensive personal air monitoring, which we used to predict exposure using robust statistical methods, limiting the potential for exposure misclassification. Additionally, the study population was only composed of never smoking Chinese women, which removes potential confounding due to sex, race/ethnicity, and cigarette smoking. This study also has limitations. The sample size for certain comparisons was limited (e.g., smoky vs. smokeless coal), which reduced our statistical power to detect modest to small associations. However, given the potent carcinogenicity of smoky coal, out *a priori* expectation is that we would potentially see strong biologic effects for biomarkers relevant to cancer etiology and mortality, such as GrimAge. Further, there may be residual confounding from other sources of air pollution. However, we were able to conduct sensitivity analyses incorporating ETS, one of the major potential confounders, and found that the associations between HAP and GrimAge EAA remain.

In summary, we observed greater GrimAge EAA for never-smoking women exposed to PAHs from using smoky coal for cooking and heating in the home. To our knowledge, this is among the first epidemiologic studies to suggest that exposure to indoor air pollutants may be associated with accelerated epigenetic aging. The mechanisms underlying the associations between HAP exposure with EAA, as well as the relationship between HAP-related EAA and downstream health outcomes, warrant further investigation.

CRediT authorship contribution statement

Batel Blechter: Methodology, Formal analysis, Visualization, Writing – original draft, Writing – review & editing. **Andres Cardenas:** Methodology, Formal analysis, Visualization, Writing – original draft, Writing – review & editing. **Junming Shi:** Methodology, Formal analysis, Visualization, Writing – original draft, Writing – review & editing. **Jason Y.Y. Wong:** Methodology, Data curation, Formal analysis, Writing – review & editing. **Wei Hu:** Project administration, Investigation, Data curation, Writing – review & editing. **Mohammad L. Rahman:** Writing – review & editing. **Charles Breeze:** Writing – review & editing. **George S. Downward:** Investigation, Data curation, Writing – review & editing. **Lützen Portengen:** Investigation, Data curation, Writing – review & editing. **Yongliang Zhang:** Writing – review & editing. **Bofu Ning:** Project administration, Investigation, Data curation, Writing – review & editing. **Bu-Tian Ji:** Investigation, Writing – review & editing. **Richard Cawthon:** Writing – review & editing. **Jihua Li:** Project administration, Investigation, Data curation, Writing – review & editing. **Kaiyun Yang:** Data curation, Writing – review & editing. **Anne Bozack:** Writing – review & editing. **H. Dean Hosgood:** Investigation, Writing – review & editing. **Debra T. Silverman:** Conceptualization, Writing – review & editing, Supervision. **Yunchao Huang:** Methodology, Writing – review & editing, Supervision. **Nathaniel Rothman:** Conceptualization, Writing – review & editing, Supervision. **Roel Vermeulen:** Conceptualization, Methodology, Formal analysis, Writing – review & editing, Supervision. **Qing Lan:** Conceptualization, Methodology, Investigation, Formal analysis, Writing – review & editing, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

The authors acknowledge the research contributions of the Cancer Genomics Research Laboratory for their expertise, execution, and support of this research in the areas of project planning, wet laboratory processing of specimens, and bioinformatics analysis of generated data. This project has been funded in whole or in part with Federal funds from the National Cancer Institute, National Institutes of Health, under NCI Contract No. 75N910D00024. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envint.2023.108041.

References

R: a language and environment for statistical computing. https://www.gbif.org/tool/81287/r-a-language-and-environment-for-statistical-computing.

Alfano, R., et al., 2018. The Impact of Air Pollution on Our Epigenome: How Far Is the Evidence? (A Systematic Review). Curr. Environ. Health Rep. 5, 544–578.

Aryee, M.J., et al., 2014. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics 30, 1363–1369.
Baglietto, L., et al., 2017. DNA methylation changes measured in pre-diagnostic peripheral blood samples are associated with smoking and lung cancer risk. Int. J. Cancer 140, 50.
Barone-Adesi, F., et al., 2012. Risk of lung cancer associated with domestic use of coal in Xuanwei, China: Retrospective cohort study. BMJ (Online) 345.
Blechter, B., et al., 2023. Exposure to smoky coal combustion emissions and leukocyte Alu retroelement copy number. Carcinogenesis. https://doi.org/10.1093/CARCIN/BGAD027.
Cardenas, A., et al., 2022. Epigenome-wide association study and epigenetic age acceleration associated with cigarette smoking among Costa Rican adults. Sci. Rep. 12.
Chapman, R.S., et al., 1988. The epidemiology of lung cancer in Xuan Wei, China: current progress, issues, and research strategies. Arch. Environ. Health 43, 180–185.
Christensen, B.C., et al., 2009. Aging and Environmental Exposures Alter Tissue-Specific DNA Methylation Dependent upon CpG Island Context. PLoS Genet 5, 1000602.
Derhovanessian, E., Solana, R., Larbi, A., Pawelec, G., 2008. Immunity, ageing and cancer. Immun. Ageing 5.
Downward, G.S., et al., 2014. Polycyclic Aromatic Hydrocarbon Exposure in Household Air Pollution from Solid Fuel Combustion among the Female Population of Xuanwei and Fuyuan Counties, China. *Environ. Sci. Technol.* 48, 14632.
Downward, G.S., et al., 2016. Outdoor, indoor, and personal black carbon exposure from cookstoves burning solid fuels. Indoor Air 26, 784–795.
Dugué, P.-A., et al., 2021. Biological Aging Measures Based on Blood DNA Methylation and Risk of Cancer: A Prospective Study. JNCI Cancer Spectr 5.
Effects Institute, H., 2020. Health Effects Institute. State of Global Air 2020. Special. Report 2020.
Esteller, M. Epigenetics in Cancer. *https://doi.org/10.1056/NEJMra072067* **358**, 1148–1159 (2008).
Fasanelli, F., et al., 2015. Hypomethylation of smoking-related genes is associated with future lung cancer in four prospective cohorts. Nat. Commun. 6.
Breton, C. v. & Marutani, A. N. Air Pollution and Epigenetics: Recent Findings. *Curr Environ Health Rep* **1**, 35–45 (2014).
Fortin, J.P., et al., 2014. Functional normalization of 450k methylation array data improves replication in large cancer studies. GenomeBiol 15, 503.
Fraga, M.F., Esteller, M., 2007. Epigenetics and aging: the targets and the marks. Trends Genet. 23, 413–418.
Gao, X., et al., 2022. Short-Term Exposure of PM2.5and Epigenetic Aging: A Quasi-Experimental Study. Environ. Sci. Technol. 56, 14690–14700.
Hannum, G., et al., 2013. Genome-wide methylation profiles reveal quantitative views of human aging rates. Mol. Cell 49, 359–367.
Horvath, S., 2013. DNA methylation age of human tissues and cell types. GenomeBiol 14, R115.
Horvath, S., 2013. DNA methylation age of human tissues and cell types. GenomeBiol 14, 1–20.
Horvath, S., et al., 2016. An epigenetic clock analysis of race/ethnicity, sex, and coronary heart disease. GenomeBiol 17, 1–23.
Horvath, S., et al., 2018. Epigenetic clock for skin and blood cells applied to Hutchinson Gilford Progeria Syndrome and ex vivo studies. Aging 10, 1758–1775.
Hu, W., et al., 2014. Personal and Indoor PM2.5 Exposure from Burning Solid Fuels in Vented and Unvented Stoves in a Rural Region of China with a High Incidence of Lung Cancer. Environ. Sci. Technol. 48, 8456.
International Agency for Research on Cancer. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans: A Review of Human Carcinogens: Personal Habits and Indoor Combustions. **100**, (2012).
Jones, P.A., 2012. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat. Rev. Genet. 13, 484–492.
Lan, Q., Chapman, R.S., Schreinemachers, D.M., Tian, L., He, X., 2002. Household stove improvement and risk of lung cancer in Xuanwei, China. *J. Natl. Cancer Inst* 94, 826–835.
Levine, M.E., et al., 2015. DNA methylation age of blood predicts future onset of lung cancer in the women's health initiative. Aging (Albany NY) 7, 690.
Levine, M.E., et al., 2018. An epigenetic biomarker of aging for lifespan and healthspan. Aging 10, 573–591.
Li, J., et al., 2018. Exposure to Polycyclic Aromatic Hydrocarbons and Accelerated DNA Methylation Aging. Environ. HealthPerspect 126.
Li, X., Schöttker, B., Holleczek, B., Brenner, H., 2022. Associations of DNA methylation algorithms of aging and cancer risk: Results from a prospective cohort study. EBioMedicine 81, 104083.
López-Otín, C., Blasco, M.A., Partridge, L., Serrano, M., Kroemer, G., 2013. The Hallmarks of Aging. Cell 153, 1194.
Lu, A.T., et al., 2019. DNA methylation GrimAge strongly predicts lifespan and healthspan. Aging 11, 303–327.
Moorthy, B., Chu, C., Carlin, D.J., 2015. Polycyclic Aromatic Hydrocarbons: From Metabolism to Lung Cancer. Toxicol. Sci. 145, 5.
Mumford, J.L., et al., 1987. Lung cancer and indoor air pollution in Xuan Wei, China. *Science* 235, 217–220.
Nwanaji-Enwerem, J.C., et al., 2017. Associations between long-term exposure to PM2.5 component species and blood DNA methylation age in the elderly: The VA normative aging study. Environ. Int. 102, 57.
Oberdoerffer, P., Sinclair, D.A., 2007. The role of nuclear architecture in genomic instability and ageing. Nat. Rev.Mol Cell Biol 8, 692–702.
Portengen, L., et al., 2023. Methylated polycyclic aromatic hydrocarbons from household coal use across the life course and risk of lung cancer in a large cohort of 42,420 subjects in Xuanwei, China. Environ Int 173.
Rahman, M. L. *et al.* Epigenome-Wide Association Study of Lung Cancer Among Never Smokers in Two Prospective Cohorts in Shanghai. doi:10.2139/SSRN.4395139.
Rodier, F., Campisi, J., 2011. Four faces of cellular senescence. J. Cell Biol. 192, 547–556.
Rousseeuw, P.J., 1987. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. J. Comput. Appl. Math. 20, 53–65.
Ruiz-Hernandez, A., et al., 2015. Environmental chemicals and DNA methylation in adults: a systematic review of the epidemiologic evidence. Clin. Epigenetics 7.
Schadt, E.E., 2009. Molecular networks as sensors and drivers of common human diseases. Nature 461, 218–223.
Seow, W.J., et al., 2016. Indoor concentrations of nitrogen dioxide and sulfur dioxide from burning solid fuels for cooking and heating in Yunnan Province, China. *Indoor Air* 26, 776.
Sisti, J., Boffetta, P., 2012. What proportion of lung cancer in never-smokers can be attributed to known risk factors? Int. J. Cancer 131, 265–275.
Tian, Y., et al., 2017. ChAMP: updated methylation analysis pipeline for Illumina BeadChips. Bioinformatics 33, 3982–3984.
Vermeulen, R., et al., 2019. Constituents of Household Air Pollution and Risk of Lung Cancer among Never-Smoking Women in Xuanwei and Fuyuan, China. *Environ. Health Perspect.* 127.
VOLUME 95 Household Use of Solid Fuels and High-temperature Frying IARC Monographs on the Evaluation of Carcinogenic Risks to Humans WORLD HEALTH ORGANIZATION INTERNATIONAL AGENCY FOR RESEARCH ON CANCER.
Weinhold, B., 2006. Epigenetics: The Science of Change. Environ. HealthPerspect 114, A160.
Wong, J.Y.Y., et al., 2017. Personal exposure to fine particulate matter and benzo[a] pyrene from indoor air pollution and leukocyte mitochondrial DNA copy number in rural China. Carcinogenesis 38, 893.
Wong, J.Y.Y., et al., 2019. Lung cancer risk by geologic coal deposits: A case–control study of female never-smokers from Xuanwei and Fuyuan, China. *Int. J. Cancer* 144, 2918–2927.
Wong, J.Y.Y., et al., 2021. Elevated urinary mutagenicity among those exposed to bituminous coal combustion emissions or diesel engine exhaust. Environ. Mol. Mutagen 62, 458–470.