

Heterogeneous associations of multiplexed environmental factors and multidimensional aging metrics

Received: 24 October 2023

Accepted: 31 May 2024

Published online: 10 June 2024



Fan Pu¹, Weiran Chen¹, Chenxi Li¹, Jingqiao Fu², Weijing Gao¹, Chao Ma³,
Xingqi Cao¹, Lingzhi Zhang¹, Meng Hao⁴, Jin Zhou⁵, Rong Huang⁵,
Yanan Ma^{5,7}✉, Kejia Hu^{6,7}✉ & Zuyun Liu^{1,7}✉

Complicated associations between multiplexed environmental factors and aging are poorly understood. We manipulated aging using multidimensional metrics such as phenotypic age, brain age, and brain volumes in the UK Bio-bank. Weighted quantile sum regression was used to examine the relative individual contributions of multiplexed environmental factors to aging, and self-organizing maps (SOMs) were used to examine joint effects. Air pollution presented a relatively large contribution in most cases. We also found fair heterogeneities in which the same environmental factor contributed inconsistently to different aging metrics. Particulate matter contributed the most to variance in aging, while noise and green space showed considerable contribution to brain volumes. SOM identified five subpopulations with distinct environmental exposure patterns and the air pollution subpopulation had the worst aging status. This study reveals the heterogeneous associations of multiplexed environmental factors with multidimensional aging metrics and serves as a proof of concept when analyzing multifactors and multiple outcomes.

Accelerated aging is a crucial risk factor for various chronic diseases and mortality¹. Given that aging is a complicated multisystemic process, a single aging biomarker may not comprehensively and accurately portray the whole landscape of the personal aging process due to the individual heterogeneity of cells, tissues, and organs². Moreover, considering that persons may hold varying rates of aging³, using chronological age to measure one's aging process may be arbitrary and opinionated. In fact, aging could be characterized by various physiological phenotypes. Many domain-specific

aging metrics have been developed and widely used, such as the brain (e.g., brain age⁴) and physical functioning (e.g., frailty phenotype score^{5,6}). Additionally, we have recently developed a composite aging metric, phenotypic age (PhenoAge), derived from multi-systemic chemistry biomarkers to reflect changes in multiple dimensions, including body composition, homeostatic mechanisms, and energetics over time. These aging metrics can capture morbidity and mortality risk beyond chronological age^{5,7–11}, thus providing unique opportunities to investigate risk factors for aging processes

¹Center for Clinical Big Data and Analytics of the Second Affiliated Hospital, and Department of Big Data in Health Science School of Public Health, The Key Laboratory of Intelligent Preventive Medicine of Zhejiang Province, Zhejiang University School of Medicine, Hangzhou 310058 Zhejiang, China. ²Ocean College, Zhejiang University, Zhoushan 316021 Zhejiang, China. ³School of Economics and Management, Southeast University, Nanjing 211189 Jiangsu, China. ⁴State Key Laboratory of Genetic Engineering, Collaborative Innovation Center for Genetics and Development, School of Life Sciences and Human Phenome Institute, Fudan University, Shanghai 200433, China. ⁵Key Laboratory of Environmental Stress and Chronic Disease Control & Prevention, Ministry of Education, China Medical University; Department of Biostatistics and Epidemiology, School of Public Health, China Medical University, Shenyang 110122 Liaoning, China. ⁶Department of Big Data in Health Science School of Public Health, Zhejiang University School of Medicine, Hangzhou 310058 Zhejiang, China. ⁷These authors contributed equally: Yanan Ma, Kejia Hu, Zuyun Liu. ✉e-mail: ynma@cmu.edu.cn; kejiahu@zju.edu.cn; zuyunliu@zju.edu.cn

and subsequently inform preventive programs against aging^{8,12–14}. Considering the heterogeneous biological aging process across individuals and organs¹⁵, it is necessary to examine whether certain factors have varying effects on multidimensional aging, rather than evenly and synchronously.

Given that numerous exposures throughout the life course jointly depict the multifaceted picture of aging, understanding the factors contributing to aging is critical to postponing aging and decreasing multiple chronic disease risks. We have previously demonstrated the contributions of various factors to aging, including genetics¹², unhealthy lifestyles^{12,13}, life course adversities^{13,14}, and certain chemicals¹⁶, with limited focus on environmental factors, which are considerably modifiable contributors^{17–19}. Among the few studies on environmental exposures and aging, the majority only evaluated a single environmental factor^{20–23}. However, individuals are exposed to multiplexed environmental factors rather than a single exposure in reality. Focusing on one of the multiplexed correlated factors may result in an erroneous estimation of the potential association of environmental factors with aging. Hence, considering environmental factors holistically and even subdividing potential specific environmental exposure patterns among populations is critical to deepening the understanding of the extent and how multiplexed environmental factors jointly contribute to aging, as well as revealing the potential variance in aging caused by environmental inequality.

In this work, we conducted a proof-of-concept study (Fig. 1) using data from the UK Biobank (UKB), a large population-based cohort study with ~500,000 participants aged 40–69 years²⁴. We show heterogeneous associations of multiplexed environmental factors available in the UKB (i.e., air pollution, green and blue spaces, and noise) with multidimensional aging metrics while air pollution presents a relatively large contribution in most cases. Then, five

subpopulations with distinct environmental exposure patterns are identified, exhibiting the aging inequality, and the air pollution subpopulation has the worst aging status.

Results

Population characteristics

As shown in Supplementary Table 1, the number of participants was 344,088 and 416,998, respectively, in the analysis of PhenoAge and frailty phenotype score. A total of 34,588 participants, with a mean age of 55.46 (standard deviations = 7.37) years, were included in the analysis of brain age. Air pollution (except PM_{2.5–10}) was strongly positively correlated with each other ($r > 0.60$) (Fig. 2A). Green space and blue space in different buffers were strongly correlated with one another, while the correlation between green space (1000 m buffer) and blue space (1000 m buffer) was relatively weak ($r = 0.18$). However, not all aging metrics were highly correlated (Fig. 2B). Particularly, age showed a strong correlation with PhenoAge ($r = 0.85$), as well as the volume of gray matter (GM) ($r = -0.61$) and brain ($r = -0.56$) (Fig. 2B).

Relative individual contributions of multiplexed environmental factors to aging

As shown in Figs. 3–5, multiplexed environmental factors were significantly associated with all aging metrics, and PM₁₀ was almost always dominant. However, fair heterogeneities in the relative contributions of the same environmental factors to multidimensional aging metrics were observed.

Specifically, PhenoAge had a significant positive association with multiplexed environmental factors ($\beta = 0.043$; 95% CI: 0.014, 0.072, Fig. 3), and PM₁₀ was predominant, with a relative contribution of 40.8%, as a major contributor among the multiplexed environmental factors to the variance in PhenoAge. Meanwhile, PM_{2.5–10} (40.4%)

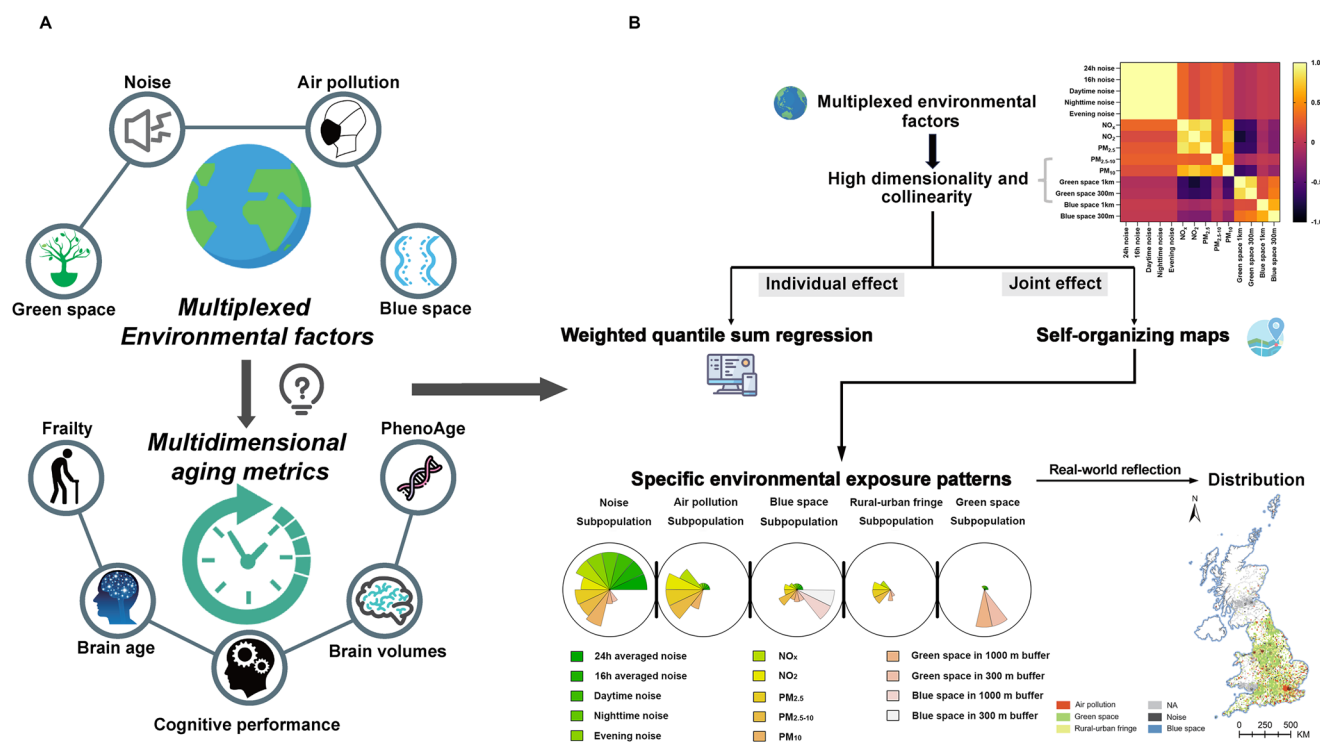


Fig. 1 | Overview of the study design. A The complex associations with multiplexed environmental factors and multidimensional aging metrics. **B** Weighted quantile sum regression (WQS) and self-organizing maps (SOM) were performed to deal with the high dimensionality and collinearity of multiplexed environmental factors and figured out the individual and joint effects of multiplexed

environmental factors on aging, respectively. Five subpopulations with specific environmental exposure patterns were distinguished and were reflected in exact locations on the UK map. Cartoon figures can be freely downloaded at <https://www.iconfont.cn/>.

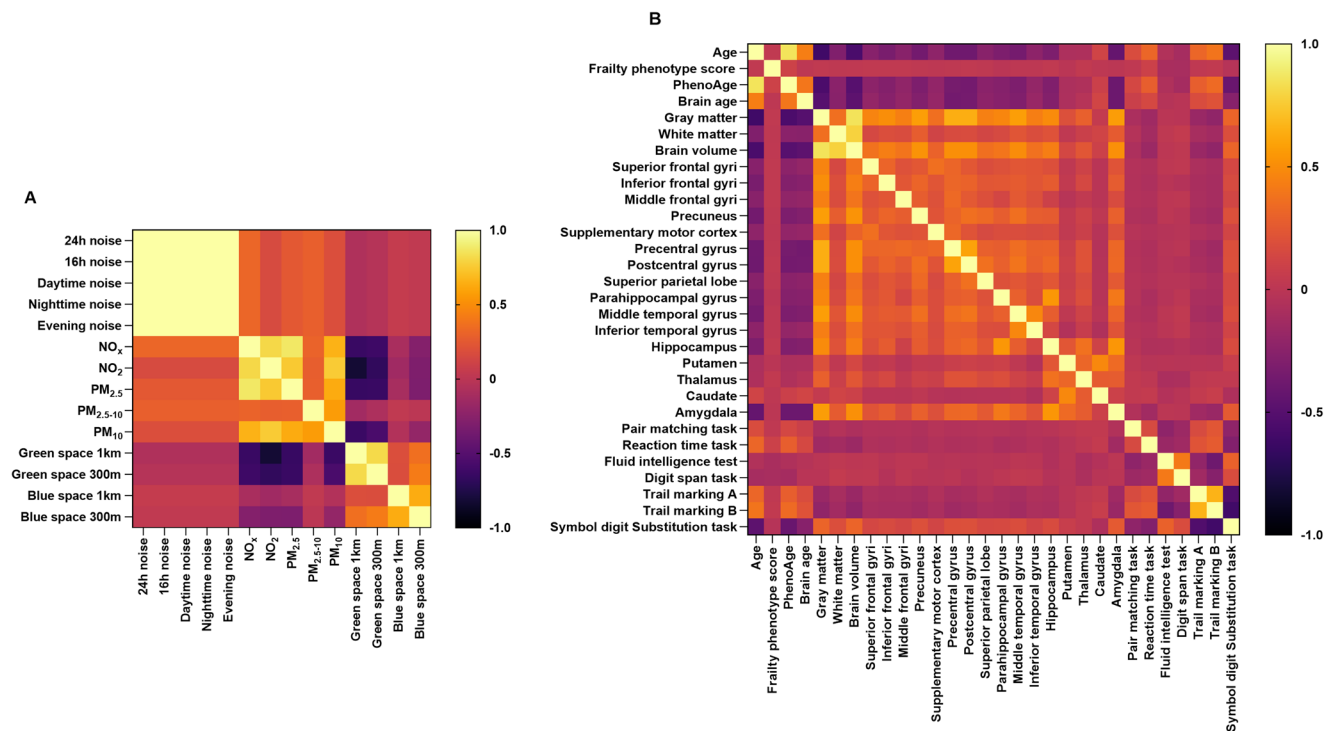


Fig. 2 | Heatmaps of Spearman's correlations among multiplexed environmental factors. (A) and multidimensional aging metrics **(B)**. PM₁₀, particulate matter with aerodynamic diameter $\leq 10 \mu\text{m}$; PM_{2.5}, particulate matter with aerodynamic diameter $\leq 2.5 \mu\text{m}$; PM_{2.5-10}, particulate matter with aerodynamic diameter

between 2.5 and $10 \mu\text{m}$; NO₂, nitrogen dioxide; NO_x, nitrogen oxides. We used Spearman's correlations to assess the correlations among multiplexed environmental factors **(A)** and multidimensional aging metrics **(B)**. Source data are provided as a Source Data file.

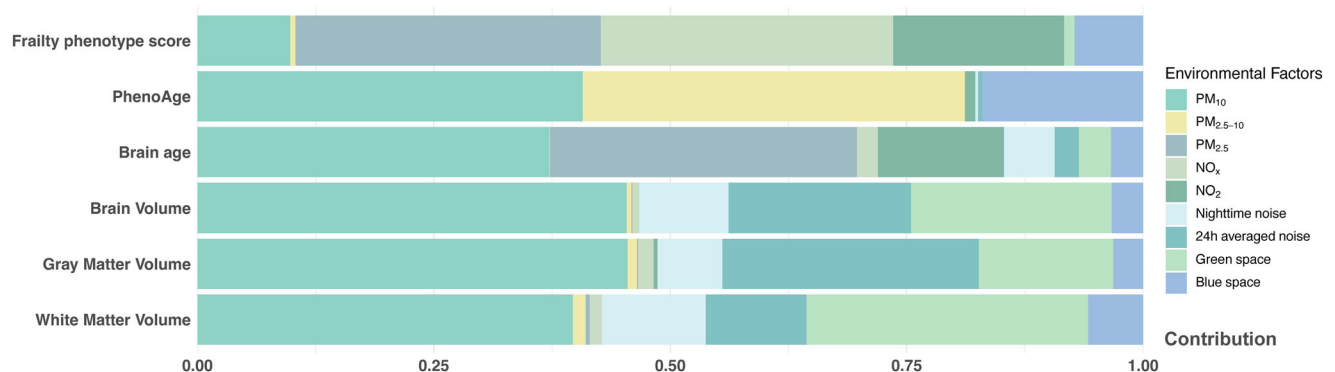


Fig. 3 | Relative individual contributions of multiplexed environmental factors to aging metrics. PM₁₀, particulate matter with aerodynamic diameter $\leq 10 \mu\text{m}$; PM_{2.5}, particulate matter with aerodynamic diameter $\leq 2.5 \mu\text{m}$; PM_{2.5-10}, particulate matter with aerodynamic diameter between 2.5 and $10 \mu\text{m}$; NO₂, nitrogen dioxide;

NO_x, nitrogen oxides. We used WQS to evaluate the relative individual contribution of multiplexed environmental factors to aging metrics. Source data are provided as a Source Data file.

indicated its second-largest contribution to the variance in PhenoAge. The relative contributions were similar to that for brain age (PM₁₀ contributed 37.2% and PM_{2.5} contributed 32.5%).

Regarding frailty phenotype score, PM_{2.5} (32.3%) and NO_x (30.9%) surpassed PM₁₀ (9.8%), taking the dominant place. Although 24-h average noise did not show much contribution to other aging metrics, it presented a considerable contribution to the variance in GM (27.1%) and brain volume (19.3%).

As shown in Fig. 4, air pollution (particularly PM₁₀) still accounted for the major contributions of variance in these brain indicators and cognitive performance. Compared with other regions, the precuneus and postcentral gyrus were more influenced by multiplexed environmental factors. Green space showed considerable contributions to

variance in cognitive performance (e.g., fluid intelligence test, digit span task, and trail making).

Specific environmental exposure patterns among populations across the UK

Based on multiplexed environmental factors, self-organizing maps (SOMs) identified five subpopulations (Fig. 5A–C) illustrates the spatially explicit distribution of the five subpopulations in the UK. Participants were not evenly distributed across subpopulations, and we named subpopulations mainly based on each environmental exposure pattern and location characteristics. Most participants experienced moderate air pollution and held moderate green space and blue space, roughly circled the major cities as the transition from cities to suburbs

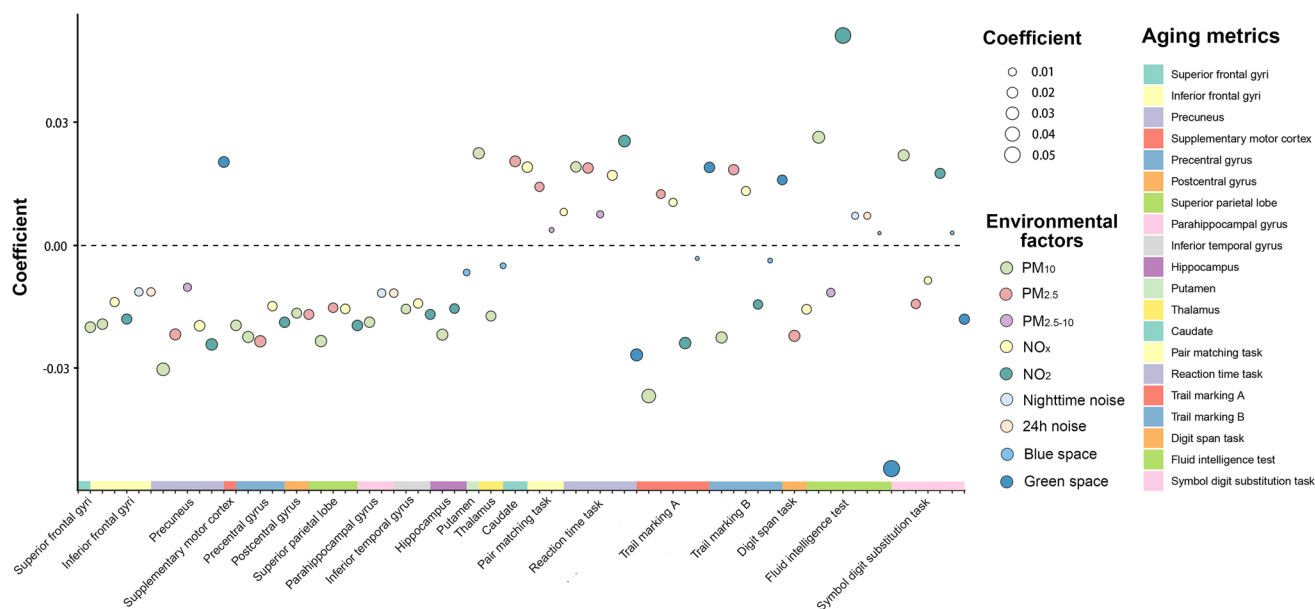


Fig. 4 | Associations of each environmental factor with volumes of aging-related regions and subcortical areas, and cognitive performances. We used linear regression to evaluate the association of each environmental factor with aging-related regions, subcortical areas, and cognitive performances. Benjamini–Hochberg procedure was used to control the family-wise error rate in

the main analyses ($n = 285$). Only associations with an FDR < 0.05 were displayed. The height, color, and size of each data point indicate the coefficient (β) between each environmental factor and one aging metric. The horizontal dashed line denotes the positive and negative correlation boundary. Source data are provided as a Source Data file.

and mountain areas, probably representing the population living in rural–urban fringe areas (the rural–urban fringe subpopulation). The air pollution subpopulation experienced the most serious air pollution and held the least green space, distributed more densely in the southeast UK and mostly located in major cities, such as London, Manchester, Liverpool, and Birmingham (dotted in Fig. 5B, C with circles), probably representing the population living in the center of major cities. The green space subpopulation experienced the lowest air pollution and held the greenest space, likely adjacent to mountain areas (e.g., the Peak District National Park, Mendip Hills, North Wessex Downs AONB, etc.), wrapping around the outside of the rural–urban fringe subpopulation. The noise subpopulation and the blue space subpopulation were distributed in a relatively scattered manner with likely linear tendencies. The noise subpopulation experienced the most severe road traffic noise and seemed to be adjacent to the cross of major roads in relatively open suburbs (Supplementary Fig. 2)²⁵. The blue space subpopulation had the most blue space and was probably distributed along rivers (Otter, Rother, Blyth) (Supplementary Fig. 3)²⁶. In the multiple linear regression model, compared to the green space subpopulation, the blue space subpopulation showed a nonsignificant difference, except for frailty phenotype score ($\beta = 0.025$; 95% CI: 0.005, 0.045) (Fig. 5D). Specifically, the air pollution subpopulation mostly located in major cities consistently had a worse aging status; for example, they had the highest frailty phenotype score ($\beta = 0.045$; 95% CI: 0.037, 0.053), and the worst brain age ($\beta = 0.133$; 95% CI: 0.053, 0.213). Although the rural–urban fringe subpopulation experienced lower air pollution and road traffic noise than the noise subpopulation, less green space among them may result in their worse frailty phenotype score as high as air pollution subpopulation ($\beta = 0.045$; 95% CI: 0.039, 0.051) and accelerated PhenoAge ($\beta = 0.081$; 95% CI: 0.036, 0.126). The noise subpopulation represented the second worst aging status in GM volume ($\beta = -0.045$; 95% CI: -0.069 , -0.021) and the whole-brain volume ($\beta = -0.052$; 95% CI: -0.085 , -0.019), which might imply possible interactive negative effect between air pollution and noise on brain aging. These results provided contribution estimates based on observed contrasts in the multiplexed environmental factors

and suggested synergistic contributions of air pollution, road traffic noise, green space, and blue space.

Additional analyses

The results of the linear regression models were largely consistent with those of WQS (Supplementary Table 2). For example, an interquartile range (IQR) increase in PM₁₀ contributed the most to variance in PhenoAge ($\beta = 0.032$; 95% CI: 0.085, 0.056), while an IQR increase in PM_{2.5} contributed the most to variance in frailty phenotype score ($\beta = 0.023$; 95% CI: 0.019, 0.027). The findings mainly remained the same while adjusting for iSES, instead of nSES. Particular matters and nitrogen oxides continued to show the dominant contribution to multidimensional aging metrics in WQS models. The air pollution subpopulation had the worst aging status, e.g., they had the highest PhenoAge ($\beta = 0.189$; 95% CI: 0.134, 0.244). Detailed results were reported in Supplementary Results and Supplementary Tables 3–6. After eliminating participants living in the current location for < 5 years, the results still largely remained the same (Supplementary Tables 7–8).

According to the stratified analyses (Supplementary Tables 9–16), overall, air pollution remained the most important contributor to multidimensional aging metrics. Males seemed to be a vulnerable subpopulation in which each aging metric was significantly associated with multiplexed environmental factors. In males, PM₁₀ made the dominant contribution to PhenoAge (45.5%), white matter (WM) volume (54.5%), and brain volume (52.7%). Meanwhile, heavy drinkers (drinking daily or almost daily) were more likely to be influenced by multiplexed environmental factors, especially for their brain volumes. For example, multiplexed environmental factors were significantly associated with the whole-brain volume among heavy drinkers ($\beta = -3505.70$; 95% CI: -5730.89 , -1280.51) and moderate drinkers (drinking one to four times per week) ($\beta = -2010.94$; 95% CI: -3342.82 , -678.98).

Compared to the green space subpopulation, the air pollution subpopulation remained to have the worst aging status in most aging metrics, especially for frailty phenotype score and the rural–urban subpopulation also showed a worse aging status in stratified analyses.

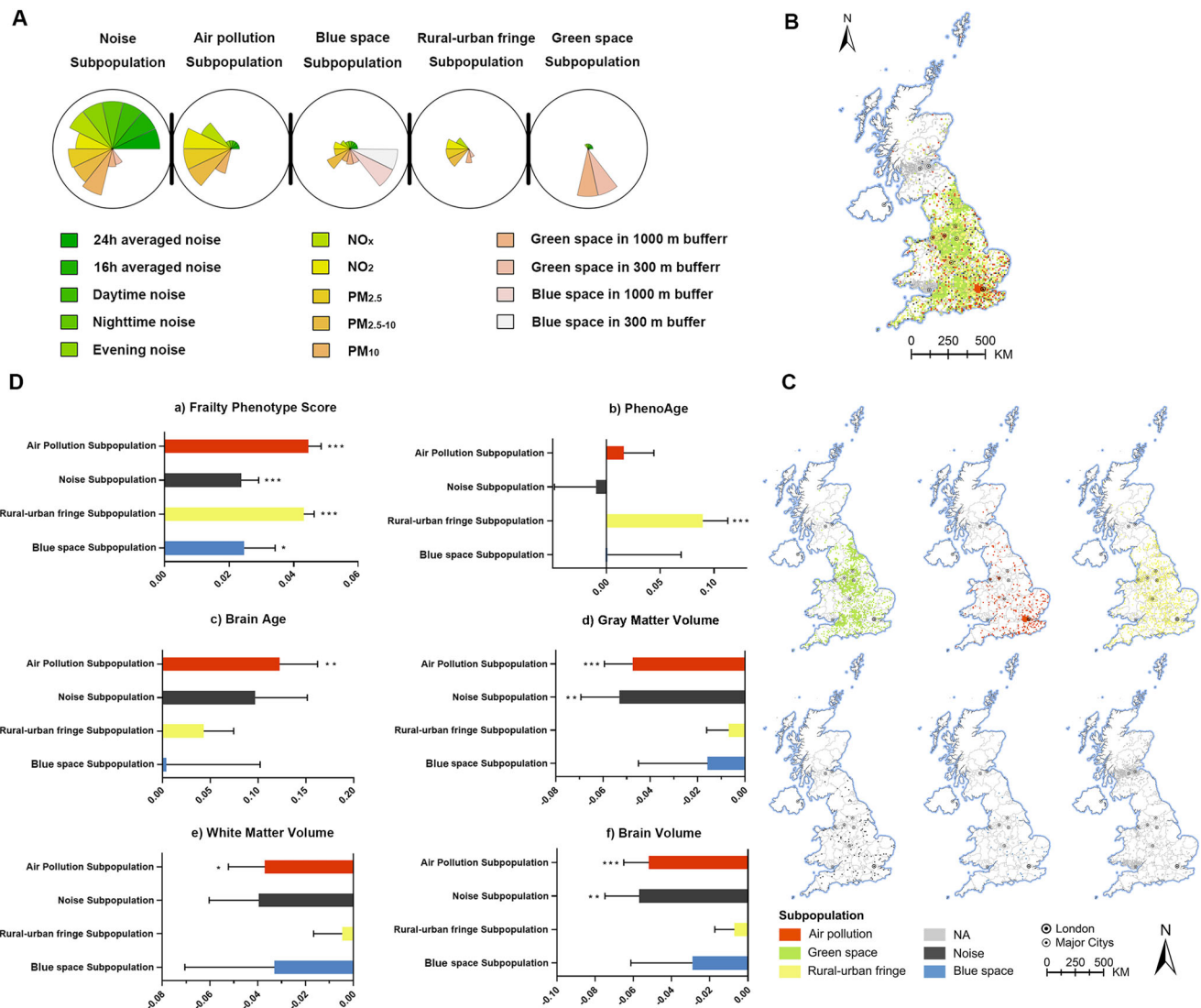


Fig. 5 | The characteristics (A) and distributions of identified subpopulations (B, C), and the associations of subpopulations with various aging metrics (D). SOM, self-organizing map. We used SOM analyses to recognize group structure. Populations were differentiated into air subpopulation (red, specific name was used only to refer to the main environmental exposure feature, not all features), green space subpopulation (green), rural-urban fringe subpopulation (yellow), noise subpopulation (black), blue space subpopulation (blue), and others (gray). **A** Characteristics of subpopulations. A larger sector size represents the larger amount of a specific environmental factor. **B** Distributions of subpopulations and **C** Distributions separately. We used multiple linear regression models to evaluate the associations of subpopulations with various aging metrics (**D**). Linear

regression models were performed to examine the associations of subpopulations with various aging metrics. All models were adjusted for age, sex, ethnicity, neighborhood socioeconomic status (nSES), smoking status, BMI (category variable), alcohol intake frequency, regular exercise, healthy diet, history of cardiovascular disease (CVD), and cancer at baseline. Benjamini–Hochberg procedure was used to control the family-wise error rate in the main analyses ($n = 285$). Two-sided P value of <0.05 was considered statistically significant (values are represented as a coefficient $\pm$ standard error of the mean. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.005$; different subpopulations vs. green space subpopulation). The blank map of the UK can be freely downloaded from GADM version 4.1 (<https://www.gadm.org/>). Source data are provided as a Source Data file.

For example, among participants with age ≥ 60 years, both the air pollution subpopulation ($\beta = 0.047$; 95% CI: 0.035, 0.059) and the rural-urban fringe subpopulation ($\beta = 0.042$; 95% CI: 0.032, 0.052) had a higher frailty phenotype score.

Discussion

Based on unique data from the UK, we disentangled the complex individual and joint contributions of the multiplexed environmental factors to aging, indicative of multidimensional aging metrics. The results suggested synergistic contributions of multiplexed environmental factors to aging, and the largest contributor was air pollution. Moreover, we also found heterogeneities in the relative contributions of environmental factors to multidimensional aging. Particulate matter (i.e., PM_{2.5}, PM_{2.5-10}, and PM₁₀) showed the predominant contribution

to variance in multidimensional aging metrics, while noise and green space showed considerable contribution to brain volumes. SOM reduced the dimensionality of the data and identified subpopulations exhibiting multiple environmental exposure patterns. Interestingly, the associations between identified subpopulations and multidimensional aging metrics were largely consistent. Compared to the green space subpopulation, the air pollution subpopulation had the worst aging status, indicative of almost all aging metrics (e.g., highest frailty phenotype score and brain age). The findings provide strong evidence of the joint contribution of multiplexed environmental factors to aging and heterogeneity in the contributions of the same environmental factors to multidimensional aging and serve as a proof-of-concept study for disentangling multifactor and multioutcome issues.

Despite many studies examining the associations of some individual environmental factors (e.g., $PM_{2.5}$ ^{21,27,28}, NO_2 ²¹, and green space^{29–32}) with aging, few have examined the relative contributions of multiple environmental factors to aging. Evaluating single individual factors' contributions without considering other environmental factors may ignore potential confounding and interaction contributions, partially due to the strong correlations among the multiplexed environmental factors, as we observed in Fig. 2. For instance, higher levels of green space tend to be related to lower levels of air pollution and traffic noise³³. The geographical relations between air pollution and road traffic noise may be universal because road traffic noise is connected to traffic intensity, which impacts traffic-associated air pollutants^{33,34}. Consequently, conclusions solely evaluating individual factors may be limited and biased, which makes them difficult to be interpreted and validated. This study holistically considered multiplexed environmental factors and disentangled their complex associations with aging to some extent, which may serve as a window for predicting various health outcomes earlier and more accurately, as well as another angle when interpreting results from previous studies that only considered a single environmental factor.

Given the global disparity in environmental pollution^{35,36}, individuals may have heterogeneous environmental exposure patterns³⁷ and thus face differentiated health risks. In this study, SOM reduced the dimensionality of the data and identified five subpopulations across the UK. These subpopulations with distinct environmental exposure patterns showed significant aging disparities. Compared to the green space subpopulation, the air pollutant subpopulation mostly located in major cities in the UK had the worst aging status throughout various aging metrics, indicating the important contribution of air pollution to aging, which was consistent with the results from WQS. Interestingly, although the rural–urban fringe subpopulation had lower air pollution and road traffic noise than the noise subpopulation, it still suffered from a poor aging status (e.g., the second highest frailty phenotype score and the highest PhenoAge), which may be due to less green space. Because the rural–urban fringe subpopulation was mainly located between the air pollutant subpopulation and the green space subpopulation, it may imply that the poor environmental qualities in major cities could have a radiative effect on persons in surrounding areas, whose accelerated aging has probably been ignored previously. This study may serve as a conceptual framework for more accurately identifying subpopulations with the same environmental exposure pattern and locations, helping develop targeted policies to improve persons' surrounding environmental qualities and further relieve the aging burden.

In addition to the oversimplified choice of multiplexed environmental factors, the metrics of aging were also not comprehensive in previous studies. As a multifaceted process with variation among individuals and organs¹⁵, aging is unlikely to be interpreted by a single-dimensional aging metric. Using various aging metrics, which focus on multisystemic and domain-specific dimensions, our study revealed heterogeneity in the relative contributions of the same environmental factors to multidimensional aging. Such heterogeneity further stressed the complexity of the aging process and the different biological mechanisms of how environmental factors may affect the aging process. First, recent studies have reported differentiated aging clocks and biomarkers involving various organs (e.g., liver and kidney) and systems (e.g., immune systems and metabolic system)^{15,38,39}, which imply that there might be systemic aging drivers/clocks overlaid with organ/tissue-specific counterparts³⁸, exhibiting complicated interactions. Second, taking road traffic noise as an example, we found that it showed more considerable contributions to the brain aging process (e.g., brain age acceleration and gray volumes decrease). When noise travels through the auditory pathway to the brain, it triggers the paraventricular nucleus of the hypothalamus to release corticotrophin-releasing hormone. This process leads to the activation

of proinflammatory cytokines and oxidative stress, which in turn stimulate the synthesis, secretion, and neurotoxicity of neurotransmitters. Additionally, many studies have demonstrated that the brain's susceptibility to noise can be attributed to the impact of stress on the higher cortical and limbic structures^{40,41}. Besides, several genes (e.g., $SOX2$ ^{42,43}, $iNOS$ ⁴⁴, and NXN ^{45,46}) were significantly associated with exposure to specific air pollutants (e.g., $PM_{2.5}$ and PM_{10}), which may further influence various biological pathways (e.g., insulin resistance and TNF signaling pathway)⁴⁷. Further, asynchronous inter- and intra-organ gene expression during aging process^{39,48} could be differentially affected by specific pathways. Future studies are poised to more systematically unravel the heterogeneous contributions of specific exposures to different aging pathways and molecules.

Overall, as a multisystemic process, aging is typically a multidimensional outcome that provides a window into disease prediction and tracing. Quantifying the contribution of environmental factors to multidimensional aging holds substantial promise for precision healthy aging and related environmental management. It should be noted that although this study used several classic aging metrics, as the development of new aging metrics of various organs and systems (recent advanced aging measurements), the present results may be biased to some extent.

To our knowledge, this was the first study to link multiplexed environmental factors to multidimensional aging, providing a proof-of-concept study for dealing with multifactor and multioutcome issues. The heterogeneity that we observed is actually a common dilemma when including multidimensional factors and outcomes simultaneously. The intriguing finding is that the results turn to consistency in the associations of subpopulations with multidimensional aging. The air pollution group remained at the worst aging status, indicative of multidimensional aging metrics, while the green space group remained the best. This implies that the heterogeneity diminished or, to some extent, was concealed when reducing the dimensionality of data on multiplexed environmental factors. In this way, we were able to focus on the latent inequality in the macro dimension, which made more sense in public health. We present efforts in dealing with multifactor and multioutcome issues, but more challenges remain. First, the concept of exposomes and phenomics is attractive. UKB provides a unique opportunity to analyze complex associations of multiple factors and outcomes, but we face the problem of having varying sample sizes for various outcomes. Moreover, few databases have such comprehensive data. Although several large cohorts, such as CHIMGEN⁴⁹, ABCD⁵⁰, cVEDA⁵¹, and Generation R⁵², provide relatively complete exposures or outcomes, it is still difficult to standardize and harmonize the data. This is one of the key reasons that the findings such as ours could not be verified externally. Some organizations, such as Gateway to Global Aging Data⁵³, are making efforts, but it remains to be a long way. Second, although we used cutting-edge statistical methods and found some interesting results in this study, the increasing data dimensions and sample size will bring great statistical challenges. Whether other new methods, including network analysis⁵⁴ and artificial intelligence, could address these challenges is unclear. Finally, many studies assume that exposure affects outcomes linearly, but this is not necessarily true. Whether multiple factors affect outcomes in a nonlinear or even systematic way requires further investigation. For example, Cohen et al. proposed understanding aging from the perspective of complex systems, and many aging processes are characterized by the interaction of multiple systems^{15,54–56}.

This study has several strengths. First, the large sample size of the middle-aged and older adults from UKB and its spatial distribution across the UK allowed for higher precision and power than smaller studies, especially in multifactor and multioutcome analyses. The diversity of the UK also allowed the SOM to identify a complete collection of environmental exposure patterns. Second, cutting-edged statistical methods provide a more comprehensive interpretation of

the individual and joint contributions of the multiplexed environmental factors to aging. Third, the actual distributions of subpopulations on the UK map further validated the reasonability and accuracy of the SOM clustering results. Fourth, using multidimensional aging metrics, we comprehensively depicted the whole landscape of aging and indicated the heterogeneity of the contribution of the same environmental factors to multidimensional aging, which highlights the necessity and challenges when dealing with multifactor and multi-outcome issues.

However, there are still some limitations. First, as a cross-sectional study, the causal inference of this study was not as strong. To further verify our findings, effective methods to analyze the longitudinal contribution of baseline multiplexed environmental factors to subsequent health outcomes are urgently needed. Second, the data on more environmental exposures such as night light, indoor air pollution, and temperature, which may be also critical to aging, were not available in the UKB. Future studies are needed to depict more comprehensive individual environmental factors. Third, due to the assumed national traffic flow baseline value and UKB's noise algorithm only considering major road noise and not considering other noise types (e.g., the impact of secondary roads), noise exposures for those at low exposure levels may be overestimated, and exposure for those living in areas with heavily trafficked minor roads may be underestimated. This may also partially explain the scattered distribution of the noise subpopulation in this study. Fourth, our study's participants were mostly White, healthier, and had higher socioeconomic status (SES) than the general population in the UK. Finally, although the data possessed were consistent with recent research, the time duration between different data collection, especially between baseline data collection and imaging assessment, may cause potential confounding effects.

In this large sample of a UK population, we captured the relative contributions of multiplexed environmental factors to aging and revealed the heterogeneity in the same environmental factors to multidimensional aging. Moreover, we identified five subpopulations with different environmental exposure patterns across the UK and observed their differentiated distribution and associations with multidimensional aging, with the air pollution group having the worst aging status. This proof-of-concept study reveals how imperative it is to holistically consider multiplexed environmental factors in analyses of their associations with aging and further points out the necessity and challenges when dealing with multifactor and multioutcome issues.

Methods

Study population

UKB is a national cohort study conducted from 2006 to 2010 and recruited ~500,000 participants aged 40–69 years in the UK. The data on the baseline questionnaire and anthropometric measures were collected at 22 assessment centers across England, Wales, and Scotland. A detailed description of the sampling design and data quality of UKB was published elsewhere^{24,57}. In the analysis of each aging metric, participants with missing data needed to calculate the aging metric were excluded. Details were provided in Supplementary Fig. 1. The North West Multi-Centre Research Ethics Committee as a Research Tissue Bank approved the UKB. Each participant provided written informed consent before the study, and researchers were allowed to use data from UKB without additional ethical clearance.

Exposure assessment

Air pollution measurements. The land use regression model integrating multisource predictors such as road, land use, and topography, developed as part of the European Study of Cohorts for Air Pollution Effects project^{58,59}, was adopted to estimate the annual average

concentration of particulate matter with a diameter of 10 μm or less (PM_{10}), particulate matter with a diameter of 2.5 and 10 μm ($\text{PM}_{2.5-10}$), and particulate matter with a diameter of 2.5 μm or less ($\text{PM}_{2.5}$), nitrogen dioxide (NO_2), and nitrogen oxides (NO_x) at a spatial resolution of 100 m. We used annual average air pollutant concentrations at participants' residential locations collected at the baseline visit to assess individual exposures to air pollution^{60,61}. Annual air pollution concentrations for NO_2 and PM_{10} were available during 2005–2010 and 2007–2010, respectively, and the annual averages of NO_2 and PM_{10} were used as exposures in the analysis. While $\text{PM}_{2.5}$, $\text{PM}_{2.5-10}$, and NO_x were only available for 2010; and thus, we used the 2010 data of these pollutants to assess individual exposures⁶².

Residential green and blue spaces. The percentages of green spaces and blue spaces in the participants' home neighborhood (300 m and 1 km radius around residential location) were estimated based on a land use map from the 2005 Generalized Land Use Database (GLUD), provided by the UK Department for Communities and Local Government (<https://www.gov.uk/government/statistics>). The GLUD records the dominant land use types based on a ten-class typology (e.g., domestic buildings, nondomestic buildings, roads, green space, blue, etc.) at a spatial resolution of 1 km⁶³. Each participant's exposure to green spaces or blue spaces was computed by overlaying the mapped green and blue spaces with the circle buffers surrounding residential locations in the geographic information system software to calculate the percentage of each buffer that contained these land cover types. Considering that the residential green (and blue) spaces in the buffers with radii from 300 to 1000 m were highly correlated (Spearman's correlation coefficient (ρ) = 0.813 for green spaces and 0.629 for blue spaces), we only reported the associations with the green (and blue) space exposures using a buffer with a radius of 1000 m.

Road traffic noise. The 2009 annual mean road traffic noise of all roads in the participant's home neighborhood (500 m radius around residential location) was modeled using the Common Noise Assessment Methods, developed from the European Union noise modeling framework^{64,65}. The model considered detailed information on absorption from buildings, noise propagation and the distance between receptor and source, land use and angle of view, building heights, meteorology, road network geography, and land cover when calculating hourly vehicle flows using a daily average traffic profile¹⁹ and was widely used in previous studies^{19,66}. We used the 24 h averaged noise (weighted average 24 h noise sound level, with a penalty of 5 and 10 dB added to the evening hours and night hours, respectively) and night-time noise (average sound pressure level during night-time hours 23:00–07:00) to be comparable to previous studies¹⁹.

Multidimensional aging metrics

A multisystemic aging metric—PhenoAge. Derived from multi-systemic chemistry biomarkers, PhenoAge serves as a relatively comprehensive aging metric. The biomarkers obtained from blood samples at the time of participant enrollment were used to calculate PhenoAge⁶⁷. Within 24 h of the blood draw, the samples were normally analyzed at the UKB central laboratory using Beckman Coulter LH750 instruments. The laboratory results were then recorded in the participant's data files. The UKB website provides more details about biomarker data processing^{68,69}. PhenoAge was developed by regressing mortality hazard on 42 clinical biomarkers and chronological age^{7,8}, and has been widely used and demonstrated to capture morbidity and mortality risks across diverse subpopulations from various countries^{48,70}.

Based on the Gompertz distribution, chronological age and nine clinical biomarkers were selected into a parametric proportional

hazards model, and 10-year mortality risk was converted into units of years. The equation to calculate PhenoAge is presented as follows

$$\text{Phenotypic Age} = 141.50225 + \frac{\ln(-0.00553 \times \ln(1 - \text{Mortality Risk}))}{0.090165}$$

where

$$\text{Mortality Risk} = 1 - e^{\frac{-e^{xb}(\exp(120 \times y) - 1)}{y}}$$

$$y = 0.0076927$$

$$\begin{aligned} xb = & -19.907 - 0.0336 \times \text{albumin} + 0.0095 \times \text{creatinine} \\ & + 0.1953 \times \text{glucose} + 0.0954 \times \ln(C - \text{reactive protein}) \\ & - 0.012 \times \text{lymphocyte percent} + 0.0268 \times \text{mean corpuscular volume} \\ & + 0.3306 \times \text{red cell distribution width} + 0.00188 \times \text{alkaline phosphatase} \\ & + 0.0554 \times \text{white blood cell count} + 0.0804 \times \text{chronological age} \end{aligned}$$

Domain-specific aging metrics—physical functioning. Being widely used as a valid metric of the aging process in geriatrics and gerontology^{71,72}, frailty is characterized by an increased vulnerability to stressor events caused by cumulative diminished reserve and dysregulation in multiple physiological systems⁶. We used frailty phenotype score, a widely used physical frailty measurement proposed by Fried et al.⁶ Frailty phenotype score was evaluated using five criteria (unintentional weight loss, exhaustion, weakness, slow gait speed, and low physical activity) and was used previously in the UKB³. Of the five criteria, weakness was assessed using objectively measured handgrip strength; the other four criteria were assessed using a self-report questionnaire (see details in the Supplementary Method). Frailty phenotype score ranged from 0 to 5, with a higher score indicating more severe frailty^{5,6}.

Domain-specific aging metrics—brain

Brain volumes. Full details on the UKB neuroimaging data are provided here: https://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/brain_mri.pdf. In brief, T1-weighted MRI used an MPRAGE sequence with 1-mm isotropic resolution. Brain volumes in mm³ were extracted from T1-structural brain MRI images, which were provided by ongoing research that started in 2014, to acquire high-quality imaging data from 100,000 predominantly healthy participants in the UKB Study⁴. In this study, we used imaging-derived phenotypes (IDPs) according to a previous study, in which the brain MRI processing pipelines were described in detail⁷³. In total, 19 aging-related IDPs were involved, including GM volume, WM volume, brain volume (GM + WM), regional GM volumes (that is, volumes of GM in the superior frontal gyrus, inferior frontal gyrus, middle frontal gyrus, supplementary motor cortex, precentral gyrus, postcentral gyrus, precuneus, superior parietal lobe, parahippocampal gyrus, middle temporal gyrus, and inferior temporal gyrus), and volumes of several subcortical areas (including the hippocampus, putamen, thalamus, caudate, and amygdala)⁷⁴. In particular, GM, WM, and total brain volumes were normalized by head size⁷⁵. For the other IDPs, the sum of the volumes in the left and right hemispheres was calculated.

Brain age. Brain age is a widely used index for quantifying individuals' brain health as deviation from a normative brain aging trajectory. Higher-than-expected brain age is thought to partially reflect the above-average rate of brain aging^{10,76}. In general, based on a set of regional and global features extracted from T1w sequences (including 365 structural magnetic resonance imaging features partitioned into 68 features of cortical thickness, area, and gray-white matter contrast, 66 features of cortical volume, 41 features of subcortical intensity, and 54 features of subcortical volume), the individuals' brain age was estimated using machine learning method, gradient tree boosting as

implemented in XGBoost (<https://xgboost.readthedocs.io>) and optimized using tenfold cross-validation and a randomized hyperparameter search⁷⁶. The volumes were converted to Z scores.

Cognitive performance

Cognition, as an important function of the human brain, was proven to be associated with the brain aging process¹⁵. In this study, we considered cognitive factors as an objective measurement of the whole-brain aging process. Thus, we obtained performance measures on seven cognitive tasks from the UKB and processed them as previously described^{77,78}. Details of seven measures and calculations for analysis have been provided elsewhere⁷⁸. The scores were first normalized, if not normally distributed, and then converted to Z scores.

The measurement time and number of participants for each exposure variable included in this study were provided in Supplementary Table 17.

Covariates

Covariates were selected based on previous related studies^{79,80}, including age, sex, ethnicity, neighborhood SES (nSES), smoking status, body mass index (BMI), alcohol intake frequency, regular exercise, healthy diet, history of cancer and cardiovascular disease (CVD) at baseline. The definition of covariates was reported in the Supporting Methods.

Statistical analyses

Baseline characteristics of each group included in multidimensional aging metrics were described. Means (standard deviations) and numbers (percentages) were used to describe the continuous and categorical variables. Mann–Whitney *U* and chi-squared tests were used to examine the differences in continuous and categorical variables. A two-sided *P* value of <0.05 was considered statistically significant in this study, unless otherwise stated.

Individual contributions

We used WQS to estimate the relative individual contributions of the multiplexed environmental factors to aging⁸¹. All analyses were adjusted for age, sex, ethnicity, nSES, smoking status, BMI, alcohol intake frequency, regular exercise, healthy diet, and history of cancer and CVD at baseline. WQS is a widely used exposure-index method in epidemiological studies to address high dimensionality and collinearity⁸². Compared to other shrinkage models, WQS has better specificity and sensitivity⁸³. Using a weighted average of factors in quantiles, WQS derives an index and then estimates the overall index effect and the weights of the deriving index by fitting a linear model between the outcome and index⁸⁴. Because WQS assumes that all components are constrained to have the same direction of association (positive or inverse) with the outcome, positive and negative models were used for probable miscarriage outcomes⁸¹. All weights are required to add up to 1, and the empirical weight of each anion, which ranges from 0 to 1, indicates the individual contribution to the WQS index⁸¹. For the volume of specific brain regions and cognitive performance, we used linear regression to preliminarily examine their association with multiplexed environmental factors. The false discovery rate was controlled at 5% across all linear regression models in the main analyses (*n* = 285) using the Benjamini–Hochberg procedure (R package “Stats”, version 4.4.0).

Considering that the overall contribution of multiplexed environmental factors may not be instructive if factors act in different directions⁸⁵ (i.e., one exposure has a protective contribution to the outcome, while another exposure has a harmful contribution), we reversed the green space and blue space exposure in the WQS regression model, with larger values representing less green space. During the model fitting process, the dataset was automatically divided into a 40% training set and a 60% validation set (R package

“gWQS”, version 4.4.0). The training set was utilized for weight estimation, while the validation set was used to test the significance of the WQS index. The final WQS index of this study was averaged from the weights in the 500 bootstrap samples⁸⁶.

Joint contributions

SOMs are unsupervised methods that group observations with comparable exposure profiles to derive low-dimensional projections of class profiles (i.e., subpopulations). In this study, SOM was used to identify subpopulations with the same environmental exposure characteristics^{87,88}, representing the homogenous and heterogeneous environmental exposure patterns among the population. SOM identified the best number of subpopulations by recognizing group structure using the within-subpopulation sum of squares and between-subpopulation sum of squares statistics⁸⁷. In this study, aiming to capture general environmental exposure patterns in a large dimension, we performed SOM in participants with complete data on environmental factors.

Each subpopulation was then matched to the geographic map of the UK based on data from the participants' addresses. Subpopulation's multidimensional aging metrics were compared (vs. the green space subpopulation) using multiple linear regression models, adjusting for all covariates.

Additional analysis

General linear regression models were performed as a robust test for the associations of multiplexed environmental factors with multidimensional aging metrics in WQS. We presented the linear effect estimates per an IQR increase in exposure to air pollution, green space, blue space, and road traffic noise by separately including one factor in the models. Given the participant's mobility, which may bias the home location-based measurement of environmental exposure to some extent, we repeated the main analysis after eliminating participants living in the current location for <5 years. Moreover, considering the vital effect of SES on environmental inequality and the complexity of SES indicators (individual SES and nSES), we repeated the main analyses by replacing nSES with iSES to further examine the results. An overall SES variable was created by latent class analysis based on the abovementioned three individual socioeconomic factors (household income, education level, and employment status). According to the item-response probabilities, three latent classes were identified, representing high, medium, and low SES⁷⁹ (details are reported in the “Methods”). Considering the potential heterogeneity among different subpopulations, we performed stratified analyses by sex, age, smoking status, and alcohol drink frequency.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The data used in the present study are available from UKB with restrictions applied. Data were used under license and are thus not publicly available. Access to the UKB data can be requested through a standard protocol (<https://www.ukbiobank.ac.uk/register-apply/>). The use of UK Biobank data was performed under application 61856. Source data are provided with this paper.

Code availability

All analyses were performed in R version 4.2.1 (<http://www.R-project.org>). The code used for WQS and SOM analyses is an adaptation of the R package “gWQS” and “kohonen” (version 3.0.12, https://github.com/blewy/Self-Organizing_maps_R), respectively, and has been made available through the GitHub repository: <https://github.com/SiriusPu/Heterogeneous-associations-of-multiplexed-environmental-factors>

and multidimensional aging metrics. The code to calculate the brain age is available at https://github.com/LCBC-UiO/VidalPineiro_BrainAge.

References

- Kennedy, B. K. et al. Geroscience: linking aging to chronic disease. *Cell* **159**, 709–713 (2014).
- López-Otín, C., Blasco, M. A., Partridge, L., Serrano, M. & Kroemer, G. The hallmarks of aging. *Cell* **153**, 1194–1217 (2013).
- Ferrucci, L., Levine, M. E., Kuo, P. L. & Simonsick, E. M. Time and the metrics of aging. *Circ. Res.* **123**, 740–744 (2018).
- Miller, K. L. et al. Multimodal population brain imaging in the UK Biobank prospective epidemiological study. *Nat. Neurosci.* **19**, 1523–1536 (2016).
- Hanlon, P. et al. Frailty and pre-frailty in middle-aged and older adults and its association with multimorbidity and mortality: a prospective analysis of 493 737 UK Biobank participants. *Lancet Public Health* **3**, e323–e332 (2018).
- Fried, L. P. et al. Frailty in older adults: evidence for a phenotype. *J. Gerontol. A Biol. Sci. Med. Sci.* **56**, M146–M156 (2001).
- Levine, M. E. et al. An epigenetic biomarker of aging for lifespan and healthspan. *Aging* **10**, 573–591 (2018).
- Liu, Z. et al. A new aging measure captures morbidity and mortality risk across diverse subpopulations from NHANES IV: a cohort study. *PLoS Med.* **15**, e1002718 (2018).
- Cole, J. H. Multimodality neuroimaging brain-age in UK biobank: relationship to biomedical, lifestyle, and cognitive factors. *Neurobiol. Aging* **92**, 34–42 (2020).
- Cole, J. H. et al. Brain age predicts mortality. *Mol. Psychiatry* **23**, 1385–1392 (2018).
- Gonneaud, J. et al. Accelerated functional brain aging in pre-clinical familial Alzheimer's disease. *Nat. Commun.* **12**, 5346 (2021).
- Liu, Z. et al. Associations of genetics, behaviors, and life course circumstances with a novel aging and healthspan measure: evidence from the Health and Retirement Study. *PLoS Med.* **16**, e1002827 (2019).
- Cao, X. et al. Contribution of life course circumstances to the acceleration of phenotypic and functional aging: a retrospective study. *EClinicalMedicine* **51**, 101548 (2022).
- Yang, G. et al. Association of unhealthy lifestyle and childhood adversity with acceleration of aging among UK Biobank participants. *JAMA Netw. Open* **5**, e2230690 (2022).
- Tian, Y. E. et al. Heterogeneous aging across multiple organ systems and prediction of chronic disease and mortality. *Nat. Med.* **29**, 1221–1231 (2023).
- Yang, Z. et al. Does healthy lifestyle attenuate the detrimental effects of urinary polycyclic aromatic hydrocarbons on phenotypic aging? An analysis from NHANES 2001–2010. *Ecotoxicol. Environ. Saf.* **237**, 113542 (2022).
- Shi, L. et al. Low-concentration air pollution and mortality in American older adults: a national cohort analysis (2001–2017). *Environ. Sci. Technol.* **56**, 7194–7202 (2022).
- Roscoe, C. et al. Associations of private residential gardens versus other greenspace types with cardiovascular and respiratory disease mortality: observational evidence from UK Biobank. *Environ. Int.* **167**, 107427 (2022).
- Kupcikova, Z., Fecht, D., Ramakrishnan, R., Clark, C. & Cai, Y. S. Road traffic noise and cardiovascular disease risk factors in UK Biobank. *Eur. Heart J.* **42**, 2072–2084 (2021).
- Martens, D. S. et al. Prenatal air pollution and newborns' predisposition to accelerated biological aging. *JAMA Pediatr.* **171**, 1160–1167 (2017).
- Gao, X., Huang, N., Guo, X. & Huang, T. Role of sleep quality in the acceleration of biological aging and its potential for preventive interaction on air pollution insults: findings from the UK Biobank cohort. *Aging Cell* **21**, e13610 (2022).

22. Ward-Caviness, C. K. et al. Long-term exposure to air pollution is associated with biological aging. *Oncotarget* **7**, 74510–74525 (2016).
23. Miri, M. et al. Association of greenspace exposure with telomere length in preschool children. *Environ. Pollut.* **266**, 115228 (2020).
24. Sudlow, C. et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* **12**, e1001779 (2015).
25. Affairs, DFEFR Strategic noise mapping: explaining which noise sources were included in 2017 noise maps. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/902825/strategic-noise-mapping-round3.pdf (2019).
26. Archive, N. R. F. Centre for ecology & hydrology rivers. <https://nrfa.ceh.ac.uk/data/search> (2023).
27. Hou, L. et al. Air pollution exposure and telomere length in highly exposed subjects in Beijing, China: a repeated-measure study. *Environ. Int.* **48**, 71–77 (2012).
28. Pieters, N. et al. Biomolecular markers within the core axis of aging and particulate air pollution exposure in the elderly: a cross-sectional study. *Environ. Health Perspect.* **124**, 943–950 (2016).
29. Marques, I. et al. Associations of green and blue space exposure in pregnancy with epigenetic gestational age acceleration. *Epigenetics* 1–11. <https://doi.org/10.1080/15592294.2023.2165321> (2023).
30. Jeong, A. et al. Residential greenness-related DNA methylation changes. *Environ. Int.* **158**, 106945 (2022).
31. Xu, R. et al. Residential surrounding greenness and DNA methylation: an epigenome-wide association study. *Environ. Int.* **154**, 106556 (2021).
32. Xu, R. et al. Surrounding greenness and biological aging based on DNA methylation: a Twin and Family Study in Australia. *Environ. Health Perspect.* **129**, 87007 (2021).
33. Markevych, I. et al. Exploring pathways linking greenspace to health: theoretical and methodological guidance. *Environ. Res.* **158**, 301–317 (2017).
34. Tétreault, L. F., Perron, S. & Smargiassi, A. Cardiovascular health, traffic-related air pollution and noise: are associations mutually confounded? A systematic review. *Int. J. Public Health* **58**, 649–666 (2013).
35. Hajat, A., Hsia, C. & O'Neill, M. S. Socioeconomic disparities and air pollution exposure: a global review. *Curr. Environ. Health Rep.* **2**, 440–450 (2015).
36. Tonne, C. et al. Socioeconomic and ethnic inequalities in exposure to air and noise pollution in London. *Environ. Int.* **115**, 170–179 (2018).
37. Christensen, G. M. et al. The complex relationship of air pollution and neighborhood socioeconomic status and their association with cognitive decline. *Environ. Int.* **167**, 107416 (2022).
38. Nie, C. et al. Distinct biological ages of organs and systems identified from a multi-omics study. *Cell Rep.* **38**, 110459 (2022).
39. Schaum, N. et al. Ageing hallmarks exhibit organ-specific temporal signatures. *Nature* **583**, 596–602 (2020).
40. Arnsten, A. F. & Goldman-Rakic, P. S. Noise stress impairs prefrontal cortical cognitive function in monkeys: evidence for a hyperdopaminergic mechanism. *Arch. Gen. Psychiatry* **55**, 362–368 (1998).
41. McEwen, B. S., Weiss, J. M. & Schwartz, L. S. Selective retention of corticosterone by limbic structures in rat brain. *Nature* **220**, 911–912 (1968).
42. Tantoh, D. M. et al. SOX2 promoter hypermethylation in non-smoking Taiwanese adults residing in air pollution areas. *Clin. Epigenetics* **11**, 46 (2019).
43. Su, C. L. et al. Blood-based SOX2-promoter methylation in relation to exercise and PM_{2.5} exposure among Taiwanese adults. *Cancers* **12**, 504 (2020).
44. Tarantini, L. et al. Effects of particulate matter on genomic DNA methylation content and iNOS promoter methylation. *Environ. Health Perspect.* **117**, 217–222 (2009).
45. Panni, T. et al. Genome-wide analysis of DNA methylation and fine particulate matter air pollution in three study populations: KORA F3, KORA F4, and the Normative Aging Study. *Environ. Health Perspect.* **124**, 983–990 (2016).
46. Eze, I. C. et al. Genome-wide DNA methylation in peripheral blood and long-term exposure to source-specific transportation noise and air pollution: the SAPALDIA Study. *Environ. Health Perspect.* **128**, 67003 (2020).
47. Wu, Y. et al. Air pollution and DNA methylation in adults: a systematic review and meta-analysis of observational studies. *Environ. Pollut.* **284**, 117152 (2021).
48. Ahadi, S. et al. Personal aging markers and ageotypes revealed by deep longitudinal profiling. *Nat. Med.* **26**, 83–90 (2020).
49. Xu, Q. et al. CHIMGEN: a Chinese imaging genetics cohort to enhance cross-ethnic and cross-geographic brain research. *Mol. Psychiatry* **25**, 517–529 (2020).
50. Jernigan, T. L. & Brown, S. A. Introduction. *Dev. Cogn. Neurosci.* **32**, 1–3 (2018).
51. Zhang, Y. et al. The Consortium on Vulnerability to Externalizing Disorders and Addictions (c-VEDA): an accelerated longitudinal cohort of children and adolescents in India. *Mol. Psychiatry* **25**, 1618–1630 (2020).
52. Kooijman, M. N. et al. The Generation R Study: design and cohort update 2017. *Eur. J. Epidemiol.* **31**, 1243–1264 (2016).
53. Gateway to global aging data. <https://g2aging.org> (2020).
54. Hao, M. et al. Phenotype correlations reveal the relationships of physiological systems underlying human ageing. *Aging Cell* **20**, e13519 (2021).
55. Cohen, A. A. et al. A complex systems approach to aging biology. *Nat. Aging* **2**, 580–591 (2022).
56. Fried, L. P. et al. The physical frailty syndrome as a transition from homeostatic symphony to cacophony. *Nat. Aging* **1**, 36–46 (2021).
57. Fry, A. et al. Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. *Am. J. Epidemiol.* **186**, 1026–1034 (2017).
58. de Hoogh, K. et al. Development of land use regression models for particle composition in twenty study areas in Europe. *Environ. Sci. Technol.* **47**, 5778–5786 (2013).
59. Beelen, R. et al. Development of NO₂ and NO_x land use regression models for estimating air pollution exposure in 36 study areas in Europe – the ESCAPE project. *Atmos. Environ.* **72**, 10–23 (2013).
60. Aung, N. et al. Association between ambient air pollution and cardiac morpho-functional phenotypes: insights from the UK Biobank Population Imaging Study. *Circulation* **138**, 2175–2186 (2018).
61. Doiron, D. et al. Air pollution, lung function and COPD: results from the population-based UK Biobank study. *Eur. Respir. J.* **54** <https://doi.org/10.1183/13993003.02140-2018> (2019).
62. Wang, M. et al. Ambient air pollution, healthy diet and vegetable intakes, and mortality: a prospective UK Biobank study. *Int. J. Epidemiol.* **51**, 1243–1253 (2022).
63. Government, D. f. C. a. L. Generalised land use database statistics for England 2005. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/825473/GLUD_Statistics_for_England_2005.pdf (2007).
64. Kefalopoulos, S. et al. Advances in the development of common noise assessment methods in Europe: the CNOSSOS-EU framework for strategic environmental noise mapping. *Sci. Total Environ.* **482–483**, 400–410 (2014).
65. Morley, D. W. et al. International scale implementation of the CNOSSOS-EU road traffic noise prediction model for epidemiological studies. *Environ. Pollut.* **206**, 332–341 (2015).
66. Cai, Y. et al. Long-term exposure to road traffic noise, ambient air pollution, and cardiovascular risk factors in the HUNT and lifelines cohorts. *Eur. Heart J.* **38**, 2290–2296 (2017).

67. Biobank, U. Category 10080: blood assays—biological samples. <https://biobank.ndph.ox.ac.uk/showcase/label.cgi?id=100080> (2012).
68. Biobank, U. Haematology data companion document. <https://biobank.ndph.ox.ac.uk/showcase/ukb/docs/haematology.pdf> (2017).
69. Biobank, U. Companion document to accompany serum biomarker data. https://biobank.ndph.ox.ac.uk/showcase/ukb/docs/serum_biochemistry.pdf (2019).
70. Buendía-Roldán, I. et al. Determination of the phenotypic age in residents of Mexico City: effect of accelerated ageing on lung function and structure. *ERJ Open Res.* **6**, 00084–02020 (2020).
71. Hoogendijk, E. O. et al. Frailty: implications for clinical practice and public health. *Lancet* **394**, 1365–1375 (2019).
72. Clegg, A., Young, J., Iliffe, S., Rikkert, M. O. & Rockwood, K. Frailty in elderly people. *Lancet* **381**, 752–762 (2013).
73. Alfaro-Almagro, F. et al. Image processing and quality control for the first 10,000 brain imaging datasets from UK Biobank. *Neuroimage* **166**, 400–424 (2018).
74. Zhang, J. et al. Associations of midlife dietary patterns with incident dementia and brain structure: findings from the UK Biobank Study. *Am. J. Clin. Nutr.* **118**, 218–227 (2023).
75. Tian, Q., Pilling, L. C., Atkins, J. L., Melzer, D. & Ferrucci, L. The relationship of parental longevity with the aging brain—results from UK Biobank. *Geroscience* **42**, 1377–1385 (2020).
76. Vidal-Pineiro, D. et al. Individual variations in ‘brain age’ relate to early-life factors more than to longitudinal brain change. *Elife* **10**, e69995 (2021).
77. Kendall, K. M. et al. Cognitive performance and functional outcomes of carriers of pathogenic copy number variants: analysis of the UK Biobank. *Br. J. Psychiatry* **214**, 297–304 (2019).
78. Kendall, K. M. et al. Cognitive performance among carriers of pathogenic copy number variants: analysis of 152,000 UK Biobank subjects. *Biol. Psychiatry* **82**, 103–110 (2017).
79. Zhang, Y. B. et al. Associations of healthy lifestyle and socioeconomic status with mortality and incident cardiovascular disease: two prospective cohort studies. *BMJ* **373**, n604 (2021).
80. Klompmaker, J. O. et al. Associations of combined exposures to surrounding green, air pollution, and road traffic noise with cardiometabolic diseases. *Environ. Health Perspect.* **127**, 87003 (2019).
81. Carrico, C., Gennings, C., Wheeler, D. C. & Factor-Litvak, P. Characterization of weighted quantile sum regression for highly correlated data in a risk analysis setting. *J. Agric. Biol. Environ. Stat.* **20**, 100–120 (2015).
82. Araki, A. et al. Combined exposure to phthalate esters and phosphate flame retardants and plasticizers and their associations with wheeze and allergy symptoms among school children. *Environ. Res.* **183**, 109212 (2020).
83. Czarnota, J., Gennings, C. & Wheeler, D. C. Assessment of weighted quantile sum regression for modeling chemical mixtures and cancer risk. *Cancer Inf.* **14**, 159–171 (2015).
84. Hu, H. et al. Methodological challenges in spatial and contextual exposome-health studies. *Crit. Rev. Environ. Sci. Technol.* **53**, 1–20 (2022).
85. Keil, A. P. et al. A quantile-based g-computation approach to addressing the effects of exposure mixtures. *Environ. Health Perspect.* **128**, 47004 (2020).
86. Zhang, S., Tang, H. & Zhou, M. Sex-specific associations between nine metal mixtures in urine and urine flow rate in US adults: NHANES 2009–2018. *Front. Public Health* **11**, 1241971 (2023).
87. Pearce, J. L. et al. Using self-organizing maps to develop ambient air quality classifications: a time series example. *Environ. Health* **13**, 56 (2014).
88. Pearce, J. L. et al. Characterizing the spatial distribution of multiple pollutants and populations at risk in Atlanta, Georgia. *Spat. Spatiotemporal Epidemiol.* **18**, 13–23 (2016).

Acknowledgements

This research was supported by a grant from the National Natural Science Foundation of China (72374180, 72074221), Research Center of Prevention and Treatment of Senescence Syndrome, School of Medicine Zhejiang University (2022010002), the Fundamental Research Funds for the Central Universities, Key Laboratory of Intelligent Preventive Medicine of Zhejiang Province (2020E10004), the Zhejiang Provincial Natural Science Foundation of China (Y23D050006), the National Natural Science Foundation of China (42001013), and Zhejiang University Global Partnership Fund. The funders had no role in the study design; data collection, analysis, or interpretation; in the writing of the report; or in the decision to submit the article for publication.

Author contributions

Z.Y.L. designed the study. F.P., W.R.C., C.X.L., C.M., L.Z.Z., and M.H. conducted the data analysis. F.P. drafted the manuscript. Z.Y.L., K.J.H., Y.N.M., J.Q.F., X.Q.C., J.Z., and R.H. critically revised the manuscript for important intellectual content. All authors approved the final version of the manuscript. Z.Y.L. is the guarantor. The corresponding author attests that all the listed authors meet the authorship criteria and that no others meeting the criteria have been omitted. Ph.D. candidate Mengying Wang from Zhejiang University and Ph.D. candidate Guoao Li and Prof. Fen Huang from Anhui Medical University helped deal with some questions about data analysis. They did not fully meet the authorship criteria, and thus, we heartedly acknowledged them for their potential contribution to this work.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-49283-0>.

Correspondence and requests for materials should be addressed to Yanan Ma, Kejia Hu or Zuyun Liu.

Peer review information *Nature Communications* thanks Xu Gao, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

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

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

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, 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 changes were made. 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/4.0/>.

© The Author(s) 2024