Aus dem
Institut für Epidemiologie (EPI)
Helmholtz-Zentrum München

Air pollution and health — what are the underlying biochemical mediators?

Dissertation
zum Erwerb des Doctor of Philosophy (Ph.D.) an der Medizinischen Fakultät der Ludwig-Maximilians-Universität München

vorgelegt von
Yueli Yao

aus
Weihui / China

Jahr
2026

Mit Genehmigung der Medizinischen Fakultät der
Ludwig-Maximilians-Universität München

Erstes Gutachten:	Prof. Dr. Annette Peters
Zweites Gutachten:	Prof. Dr. Dennis Nowak
Drittes Gutachten:	Prof. Dr. Hermann Ammer
Viertes Gutachten:	Prof. Dr. Eva Rehfuess

Dekan: Prof. Dr. med. Thomas Gudermann

Tag der mündlichen Prüfung: 08.01.2026

谨献给我亲爱的家人

To my beloved family

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Dekanat Medizinische Fakultät
Promotionsbüro

Affidavit

Yao, Yueli

Surname, first name

I hereby declare, that the submitted thesis entitled

Air pollution and health – what are the underlying biochemical mediators?

is my own work. I have only used the sources indicated and have not made unauthorised use of services of a third party. Where the work of others has been quoted or reproduced, the source is always given.

I further declare that the dissertation presented here has not been submitted in the same or similar form to any other institution for the purpose of obtaining an academic degree.

Munich, 27.01.2026
Place, Date

Yueli Yao
Signature doctoral candidate

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Dean's Office Medical Faculty
Doctoral Office

Confirmation of congruency between printed and electronic version of the doctoral thesis

Yao, Yueli

Surname, first name

I hereby declare that the electronic version of the submitted thesis, entitled

Air pollution and health – what are the underlying biochemical mediators?

is congruent with the printed version both in content and format.

Munich, 27.01.2026 — Place, Date

Yueli Yao — Signature doctoral candidate

Table of content

List of abbreviations

AD	Alzheimer's disease
Arg	Arginine
AQG	Air Quality Guidelines
BCAAs	Branched-chain amino acids
BMI	Body-mass index
CAD	Coronary artery disease
CHD	Coronary heart disease
CI	Confidence intervals
CRP	C-reactive protein
COPD	Chronic obstructive pulmonary disease
CpG	Cytosine-phosphate-Guanin
CVD	Cardiovascular disease
DALYs	Disability-adjusted life-years
DNAm	DNA methylation
DNAmHannumAge	Hannum's clock
DNAmHorvathAge	Horvath's clock
DNAmPhenoAge	Levine's clock
DNAmSkinBloodAge	Skin & blood clock
DNAmTL	DNAm-based telomere length
ECG	Electrocardiogram
EPHX2	Soluble epoxide hydrolase
EU	European Union
EWAS	Epigenome-wide association study
GBD	Global burden of disease
hs-CRP	High-sensitivity C-creative protein
IQR	Interquartile range
KORA	German population-based Cooperative Health Research in the Region of Augsburg
LPC	Lysophosphatidylcholine
LUR	Land-use regression

NAS	Normative Aging Study
NO_2	Nitrogen dioxide
NO_x	Nitrogen oxides
O_3	Ozone
PCs	Phosphatidylcholines
PM	Particulate matter
$PM_{2.5}$	Fine particulate matter, aerodynamic diameter less than 2.5 µm
PM_{10}	Particulate matter with an aerodynamic diameter less than 10 µm
PM_{coarse}	Particulate matter with an aerodynamic diameter between 2.5 and 10 µm
PNC	Particle number concentration
QC	Quality control
RR	Relative risk
sEH	Soluble epoxide hydrolase
SLC39A5	Solute carrier family 39 member 5
SM	Sphingomyelin
TL	Telomere length
Trp	Tryptophan
WHO	World Health Organization

List of publications included in this thesis

This thesis consists of the following publications:

Paper I

Yao Y, Schneider A, Wolf K, Zhang S, Wang-Sattler R, Peters A, Breitner S. Longitudinal associations between metabolites and long-term exposure to ambient air pollution: Results from the KORA cohort study. Environ Int. 2022; 170: 107632. doi: 10.1016/j.envint.2022.107632.

Paper II

Yao Y, Schneider A, Wolf K, Zhang S, Wang-Sattler R, Prehn C, Adamski J, Peters A, Breitner S. Longitudinal associations between metabolites and immediate, short- and medium-term exposure to ambient air pollution: Results from the KORA cohort study. Sci Total Environ. 2023; 900:165780. doi: 10.1016/j.scitotenv.2023.165780.

This thesis also includes an unpublished manuscript:

Paper III (Appendix)

Yao Y, Wolf K, Breitner S, Zhang S, Waldenberger M, Winkelmann J, Schneider A, Peters A. Long-term exposure to traffic-related air pollution is associated with epigenetic age acceleration. Available at SSRN: http://dx.doi.org/10.2139/ssrn.5255305.

Contribution to the included publications

The thesis comprises two published manuscripts in *Environment International*, *Science of the Total Environment*, and one unpublished manuscript currently submitted to *Environment International*. All three are peer-reviewed and high-impact scientific journals.

I, Yueli Yao, am the first and corresponding author of all publications included in this thesis. Under the supervision of Dr. Alexandra Schneider, Dr. Susanne Breitner-Busch, Dr. Kathrin Wolf, and Prof. Annette Peters, I independently developed the research questions and statistical analysis plans for each study.

My responsibilities included the preparation of datasets for further analysis, writing and optimizing statistical code, and conducting all statistical analyses. I created visualizations by creating tables and figures to facilitate interpretation. Apart from the data analysis, I drafted all three initial manuscripts and supplementary materials, coordinated communication with co-authors, gathered and implemented their comments. As the first and corresponding author, I managed the submission and publication process, addressed peer-review comments, revised the manuscripts, as well as handled proof-reading and post-production tasks.

Paper I

I investigated the effects of long-term air pollution exposure on serum metabolites using repeated measurements from the German population-based Cooperative Health Research in the Region of Augsburg (KORA), including baseline S4 survey (1999–2001) and two follow-up examinations (F4: 2006–08 and FF4: 2013–14). I applied confounder-adjusted mixed-effects regression and pathway analyses, identifying metabolic alterations—particularly in phosphatidylcholines—associated with three air pollutants, which were further linked to the glycerophospholipid metabolism pathway. I also explored effect modification to identify susceptible subgroups.

Paper II

In this study, I extended the analysis to immediate (2-day), short-term (2-week), and medium-term (8-week) exposures. I applied generalized additive mixed-effects models, adjusting for meteorological variables (ambient temperature and relative humidity) and relevant covariates. Using generalized additive mixed models, I examined associations between air pollutants and targeted metabolites. Significant associations were identified, particularly with NO_2, and the glycerophospholipid metabolism pathway was identified as a key pathway.

Paper III (unpublished)

I explored the association between long-term exposure to ambient air pollution and epigenetic aging biomarkers. I further conducted stratified analyses by smoking status and performed a limited epigenome-wide association study (EWAS). Results showed strong

associations between air pollutants and biological age acceleration, particularly among ever smokers.

Across all three papers, I led the research, analyses, writing, and coordination efforts, ensuring the integrity and coherence of the work presented in this thesis.

In addition to the specific contributions described for Paper I, Paper II, and the unpublished Paper III, I regularly presented the research findings at my Thesis Advisory Committee (TAC) meetings, the Work-In-Progress seminars of the Environmental Risks (EnRi) Research Group, the Monday seminars at the Institute of Epidemiology, Helmholtz Munich (EPI-HMGU), and the PhD Journal Clubs at the Institute for Medical Information Processing, Biometry, and Epidemiology (IBE), Ludwig Maximilian University of Munich (LMU). I also made oral presentations at international conferences, including the International Society for Environmental Epidemiology Young (ISEE Young) 2021 and ISEE 2022.

Introductory Summary

1. Background

1.1 Air pollution and health impacts

Clean air is a basic requirement of human health and well-being. While air quality has been improved in the past decades, the World Health Organization (WHO) reports that 99% of the world's population live in regions where air pollution levels exceed the WHO Air Quality Guidelines (AQG) [1,2]. This widespread exposure to air pollution has severe health impacts, as the global burden of disease (GBD) estimates indicate that more than 9 million preventable deaths are linked to environmental exposures, particularly air pollution [3]. Air pollution is responsible for 6.7 million premature deaths annually, with 4.2 million of these directly attributable to ambient (outdoor) air pollution [2]. Particulate matter air pollution contributed 8.0% of the total disability-adjusted life-years (DALYs) and ranked as leading contributor to the GBD [4].

These health effects are caused by air pollutants such as particulate matters (PM), nitrogen dioxide (NO_2), and ozone (O_3), which were associated with specific health outcomes (see **Table 1** for a detailed overview) [2]. For instance, long-term exposure to fine particulate matter ($PM_{2.5}$) and particulate matter with an aerodynamic diameter less than 10 µm (PM_{10}) were linked to higher relative risks (RRs) of all-cause and cause-specific mortality, including cardiovascular disease (CVD), respiratory disease and lung cancer [5], while NO_2 exposure was associated rather low or moderate with all causes mortality but high with mortality of chronic obstructive pulmonary disease (COPD) [6]. It is important to note that since the publication of the WHO guidelines, additional evidence substantiated the conclusions, especially additional studies on the health impact of NO_2 were published [7-10]. Even in regions such as the United States, Canada, and parts of Europe, where ambient air pollution levels are relatively low, studies have shown that long-term exposure to air pollutants still increases the risk of non-accidental and more specifically cardiovascular mortality [11-14]. These findings highlight the significant health burden caused by air pollution worldwide.

The duration of air pollution exposure over days, so called short-term or acute exposure and over years, so called long-term or chronic exposures both lead to higher risk of adverse health effects [15]. Long-term exposure is usually associated with chronic conditions such as CVD, COPD, and cancer [14]. By contrast, short-term exposure has been linked to higher risks of acute health outcomes and exacerbations of pre-existing conditions, such as asthma or acute bronchitis [16,17]. For example, immediate exposure to

high O_3 levels during a single day can exacerbate respiratory conditions, leading to increased hospital visits [18]. The scientific evidence that duration of air pollution exposure is relevant highlight the importance of reducing daily pollution as well as addressing chronic exposure levels. Furthermore, individuals such as children, the elderly, and those with unhealthy behaviors like smoking or with pre-diseases show higher susceptibility to air pollution exposure and therefore experience more severe health impacts [19].

Table 1. Air pollutants, WHO air quality guidelines (AQG), and associated health impacts

Air Pollutant	2021 WHO AQG level [2]	Effects of exceeding AQG levels
$PM_{2.5}$	5 µg/m³ (annual mean)	Increased CVD and respiratory morbidity; higher all-cause and cause-specific mortality including CVD, non-malignant respiratory and lung cancer [2,14]
PM_{10}	15 µg/m³ (annual mean)	Higher RRs for respiratory and lung cancer mortality [2,5]
NO_2	10 µg/m³ (annual mean)	Low or moderate all-causes mortality but high with mortality of COPD [2,6]
O_3	60 µg/m³ (peak season*)	Low or moderate all-causes mortality [2,6]
$PM_{2.5}$	15 µg/m³ (24-hour mean)	Increased RRs for cardiovascular mortality, non-malignant respiratory mortality and cerebrovascular mortality, but lower than long-term exposure [2,20]
PM_{10}	45 µg/m³ (24-hour mean)	Similar to $PM_{2.5}$ [2,20]
NO_2	25 µg/m³ (annual mean)	Respiratory inflammation and reduced lung function [2,17]
O_3	100 µg/m³ (8-hour mean)	Exacerbation of respiratory conditions [2,17,18]

*Peak season refers to the period of six consecutive months when the average ozone concentration reaches the highest level (calculated as a six-month moving average).

1.2 Underlying mechanisms between air pollution and health impacts

Epidemiological and experimental research has provided strong evidence linking ambient air pollution to chronic diseases, including pulmonary, cardio-metabolic, and neurological disorders [21-24]. Two primary biological pathways have been hypothesized by which air pollutants exert harmful effects. First, ultrafine particles, particle components, and gaseous pollutants (such as ozone) can directly enter the blood stream from the lung, leading to changes in blood parameters [25]. Thereby, they are transported to every

organ of the body. Second, larger inhaled particles can initiate local inflammation in the lungs, which may, in turn, trigger systemic inflammatory responses. Additionally, they can activate pulmonary receptors involved in autonomic regulation, potentially leading to changes in electrocardiogram (ECG) and indirectly affecting health outcomes such as autonomic cardiac function. [22]. Several studies suggested that air pollution-related health outcomes are mediated through mechanisms like inflammatory response, oxidative stress, and alterations in genetic and epigenetic regulation [26-29]. A recent review summarized eight hallmarks of environmental insults (**Figure 1**) providing a comprehensive overview of how underlying biological mechanisms related to environmental exposures drive aging [29].

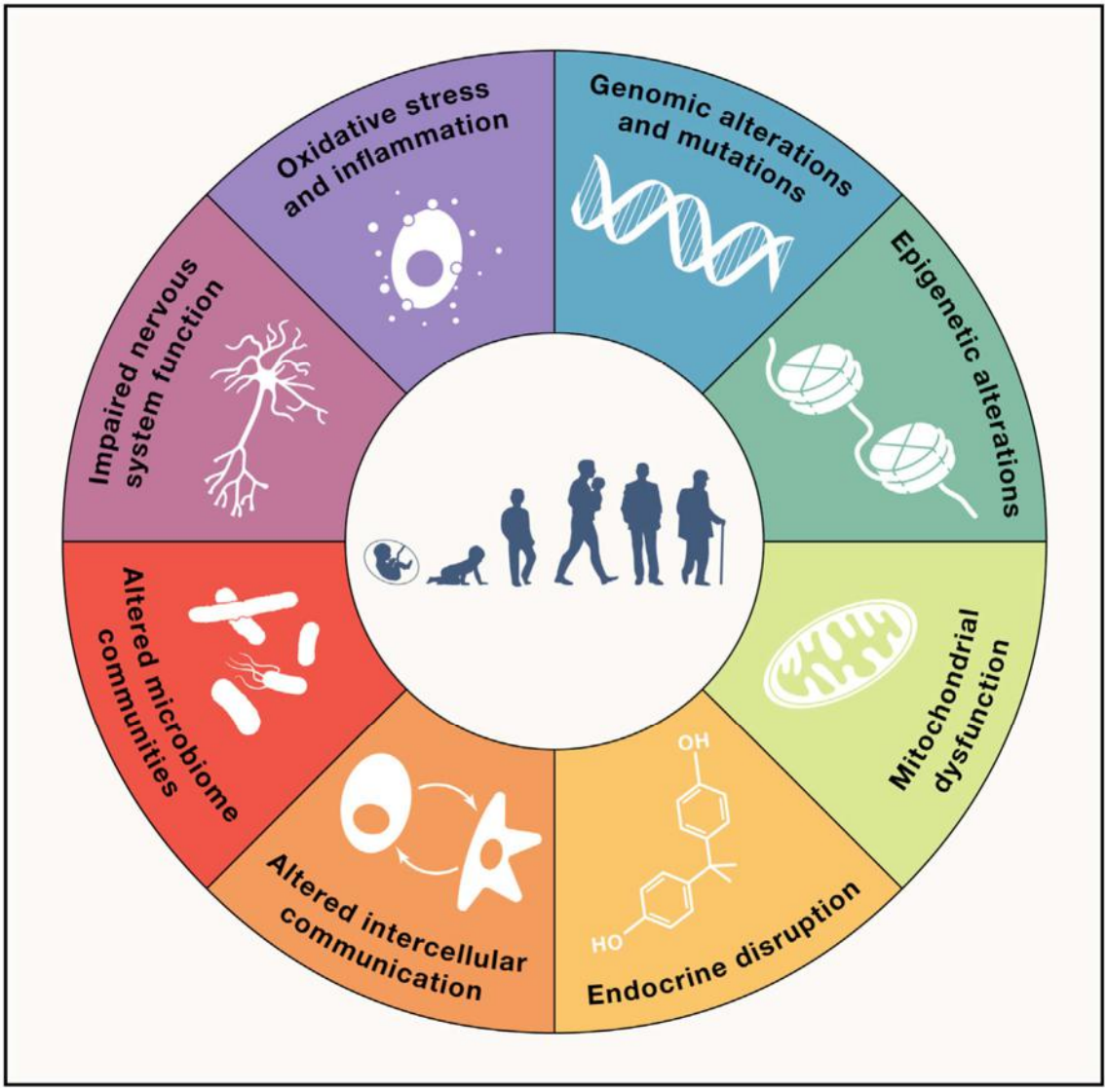


Figure 1. Hallmarks of environmental insults [29].

In this cumulative thesis, I explored the roles of metabolites and DNA methylation-derived epigenetic aging biomarkers in diseases related to ambient air pollution.

1.2.1 Metabolomics

The blood metabolome is considered as a collection of biologically active chemicals in human blood derived from endogenous processes and exogenous exposure to food, medicines, and pollutants [30]. Therefore, the biological perturbations caused by air pollutants might reflect the endogenous metabolites influenced by air pollution and exogenous metabolites originating from air pollutants. Different to genes, which are influenced by epigenetic regulation, or proteins that could be altered by post-translational modifications, metabolites work as direct indicators of biochemical activity. This makes them easier to associate with phenotypes [31].

To deeply and comprehensively understand the role and function of metabolites in biological progresses, metabolomics has become a powerful and robust method to investigate small molecular metabolites presented in the biological systems and corresponding cellular responses perturbed by endogenous or exogenous stimuli [32]. It measures numerous low-molecular weight metabolites including amino acids, sugars, fatty acids, lipids, and steroids. Even minor changes in the chemical structure of metabolites or exposure to external factors like infections and allergens can significantly alter their function. There are two main metabolomics approaches: non-targeted and targeted metabolomics, which are distinguished by the stage of metabolite identification during the data processing and the availability and use of standards [33]. Non-targeted metabolomics is a global and comprehensive analysis that aims to measure all metabolites in one sample, including the identification of known and unknown analyte signals [34]. In contrast, targeted metabolomics measures a set of known metabolites and provides quantification, which is useful to assess the response of an organism to endogenous and xenobiotic exposures, but also important for disease development and diagnosis [35]. Both approaches have their advantages and disadvantages. The targeted metabolomics offers the advantage of focusing and identifying known metabolites, enabling a clearer understanding of biological mechanisms. However, it holds the risk of missing target metabolites, and potentially lose the opportunity to discover novel metabolites underlying diseases [33,36]. In contrast, non-targeted metabolomics promotes the discovery of previously unidentified metabolites or unexpected changes. However, the large volume of data requires extensive processing capacities and complex statistical analyses. Furthermore, the lack of reference standards complicates the identification of unknown metabolites, reducing precision due to the reliance on relative quantification [37,38].

Metabolites are linked with certain diseases such as type 2 diabetes, CVD, and neurodegenerative disorders. For example, branched-chain amino acids (BCAAs) were identified to be associated with a higher risk of type 2 diabetes development [39]. Dysregulated BCAAs and related metabolites were also associated with coronary artery disease (CAD) even when controlling for diabetes [40]. Glycerophospholipids including phosphatidylglycerol and lysylphosphatidylglycerol were linked with the pathophysiology of Alzheimer's disease (AD), as well as an increasing severity of AD pathology [41]. Therefore, the relationship between metabolites and diseases is crucial for utilizing metabolites as

an intermediate phenotype to understand how air pollution impacts human health. Epidemiological studies with environmental measurements and metabolomic analysis can play an important role in this process by identifying metabolites affected by air pollutants. Thereby, we gain insights into the biochemical changes induced by air pollution and establish the mechanisms linking air pollution and its related diseases. With this approach, metabolomics provides a powerful tool to explore the interactions between environmental exposures and health outcomes.

1.2.2 Epigenetic aging

Aging is a complex process that affects most tissues and organs, and is associated with increased morbidity and mortality. In general, the analysis of mortality curves at the population level can assess aging rates, but it cannot predict the risk of morbidity and remaining life span at the individual level [42]. Chronological age refers to the number of years an individual has actually lived since birth, providing a uniform measure across races and genders. However, individuals with the same chronological age can age differently due to their differences in living conditions, lifestyles, and genetic makeup [43,44]. In contrast, biological age reflects an individual's physiological status and functional ability, which is a more accurate measure of their aging and health compared to chronological age [45]. The introduction of markers of "biological age", which can be measured at any stage of life, may enable the early detection of individual's risk for age-related diseases [46].

Recent studies have highlighted DNA methylation (DNAm)-based biomarkers being representative molecular indicators of aging. Since epigenetic changes such as DNAm are reversible, DNAm aging biomarkers may help identify both aging and anti-aging factors [47]. Moreover, DNAm patterns can be influenced by environmental factors such as air pollution, which in turn contribute to epigenetic aging [48]. Epigenetic aging biomarkers are calculated using DNAm data but are restricted to specific subsets of Cytosine-phosphate-Guanine (CpG) sites [47,49-53]. These biomarkers can serve as valuable tools to assess the molecular aging process across the life span, evaluate environmental influences, and predict health outcomes [54]. Depending on their training methods, epigenetic aging biomarkers can either accurately estimate chronological age (e.g., Hannum's and Horvath's epigenetic clocks) or quantify age- or disease-related health outcomes (e.g., Levine's clock) [49,50,52].

The first generation of epigenetic clocks was developed to predict age by leveraging age-associated DNAm changes at specific CpG sites. These clocks aimed to quantify biological age by identifying the most informative CpG sites that strongly correlate with individual's chronological age [49-51]. Horvath's clock (DNAmHorvathAge) was the first epigenetic clock designed as a multi-tissue age predictor, which was trained on DNAm data from a variety of tissues and cell types [50]. It can accurately predict age across the entire duration of the human lifespan not only for adults but also for adolescents and children, and performed very well across a wide range of tissues and cell types [50].

Concurrently, Hannum's clock (DNAmHannumAge) was developed which is a blood-based first-generation epigenetic clock that primarily depends on age-related changes in leukocytes [49]. Therefore, it is particularly suited for studies with blood samples and showed high accuracy for predicting age in adults. Lately, a new DNAm-based biomarker, the skin & blood clock (DNAmSkinBloodAge), was developed to predict age for blood and skin samples, particularly well for fibroblasts and endothelial cells [51]. Its accuracy outperforms previous clocks when applied to these tissues, and therefore will be useful to estimate chronological age in studies with either *in vivo* or *ex vivo* samples.

A major limitation of first-generation epigenetic clocks is that they focus on predicting chronological age rather than age-related health outcomes, resulting in weaker associations with morbidity and age-related diseases [47,52,53]. In contrast, the second generation of epigenetic clocks were developed to capture biological aging by selecting CpG sites that not only correlate with age but also reflect intrinsic and extrinsic factors that influence the aging process [47]. This improvement allows to better predict lifespan, mortality, and morbidity. For example, the Levine's clock (DNAmPhenoAge) firstly includes the generation of weights of chronological age and mortality-associated clinical parameters such as glucose, high-sensitivity C-creative protein (hs-CRP), and subsequently the selection of 513 CpGs by regressing the weights on blood DNAm levels [52]. Compared to the first generation of epigenetic clocks, this two-step approach captures age-related variations as well as variations in risk of diseases and even death. Similarly, the DNAmGrimAge applied a two-stage modeling approach as well. It firstly includes plasma proteins related to aging and inflammation (e.g. hs-CRP) and behavioral risk factors (e.g. DNAm-based estimator of smoking pack-years) firstly, which are then combined into a single composite measure for the lifespan [53]. This assures DNAmGrimAge to predict time-to-death and other aging-related outcomes with much better precision.

Except for these epigenetic aging biomarkers, telomeres — repetitive nucleotide sequences located at the ends of chromosomes — have also been widely studied as an aging biomarker [55]. In general, telomere length (TL) has been considered to be inversely correlated with the number of cell divisions. Therefore, shorter telomeres are always associated with accelerated aging [56,57]. However, due to technical challenges in measuring TL, such as variations in DNA extraction methods or experimental limitations, results were not consistent among studies [58-60]. In contrast, TL estimated by DNAm data with a selection of 140 CpG sites (DNAmTL) is a more robust biomarker [61] and shows a better performance than TL in predicting mortality and age-related diseases [61]. Similar to other epigenetic aging biomarkers, DNAmTL provides insights into biological aging, but with a specific focus on telomere-related aging processes. **Table 2** provides a summary of these epigenetic aging biomarkers discussed above.

Table 2. Overview of the epigenetic aging biomarkers

Epigenetics aging estimators	DNAmHorvathAge	DNAmHannumAge	DNAmSkinBloodAge	DNAmPhenoAge	DNAmGrimAge	DNAmTL
Key Features	Broad tissue applicability; first multi-tissue clock	Blood-specific focus; optimized for adults	Tissue-specific and accurate for skin and blood tissues	Predicts age related health and lifespan outcomes	Integrates risk factors and plasma proteins; mortality and health risk predictor	More robust than traditional TL; predicts mortality and diseases
CpGs (N)	353	71	391	513	1,030	140
Participants (N)[‡]	7,844	482	792	9,926	6,935	---
Age range (years)	0-100	19-101	0-94	0-100	46-78	---
Tissue coverage	51 healthy tissues and cells	Blood	Skin and blood	Blood	Blood	Blood, potentially other tissues
Training phenotypes	Chronological age	Chronological age	Chronological age	Lifespan	Lifespan	Mortality risk, diseases, age-related conditions
Regression Model	Elastic net penalized regression	Elastic net penalized regression	Elastic net penalized regression	Cox penalized regression model	Elastic net Cox penalized regression model	Elastic net regression
Prediction accuracy (R)*	0.960	0.905	0.960	---	---	---

TL: telomere length; CpGs: Cytosine-phosphate-Guanine sites; Participants (N)[‡]: the number of participants used to develop the epigenetic aging biomarkers; Prediction accuracy (R)*: the correlation coefficient between predicted biological age and actual chronological age in the dataset used for validation.

1.2.3 Link between metabolite and epigenetic aging

Omics processes are typically considered as a linear sequence: genome (DNA) → epigenome → transcriptome (RNA) → proteome → metabolome [62]. However, there is interaction and crosstalk between these layers, especially between the metabolome and the epigenome. For example, the homocysteine metabolism provides S-adenosylmethionine, which is the main methyl donor for methylation reactions, and it is central to the methionine cycle. Subsequently, the changes in the methionine cycle can disrupt homocysteine homeostasis and affect DNAm patterns and epigenetic aging [63,64]. This could be one way how metabolites interact with the epigenome, influence DNAm patterns, and contribute to the aging process. Therefore, even though researchers analyzed metabolomics and epigenetic aging data already separately, in-depths pathway analysis to find shared pathways between these two approaches may help identify key biomarkers for air pollution-related diseases.

1.3 Epidemiological evidence

The following section summarizes the evidence on studies of air pollution on metabolomics and epigenetic ageing before the studies conducted as part of this cumulative thesis.

1.3.1 Air pollution and metabolomics

Previous research has investigated the impacts on changes of metabolites by either short-term, medium-, or long-term exposure to ambient air pollution. For example, several studies reported that short-term air pollution exposure could change the serum metabolic alternation mainly on lysophosphatidylcholine (LPC) and phosphatidylcholine (PC) [65], and PC was further identified to be involved in glycerophospholipid metabolism [66]. Other studies also found that short-term air pollution could affect amino acids such as glycine, methionine, and ornithine [67-70]. Also, the perturbation of arginine metabolism was linked with exposure to traffic-related air pollution [71,72]. Furthermore, two analyses based on the Normative Aging Study (NAS, a closed cohort study) reported that long-term exposure to $PM_{2.5}$ and $PM_{2.5}$ species (e.g., ultrafine particle, black carbon) was related to the perturbation of glycerophospholipid, sphingolipid, and biosynthesis of unsaturated fatty acids etc. [66,73]. However, these studies were limited by either a small sample size [72], or conducted on subpopulations like older men [66,73], or participants with specific diseases [72], which hinders the generalization of findings. In addition, due to the cross-sectional design, or variations in air pollutants or exposure windows, or differences of metabolomic approaches, the results were inconsistent across studies. Therefore, studies particularly with longitudinal study design are required to improve our understanding of the underlying biological mechanisms for air pollution-associated adverse outcomes and promote the generalization of findings in different populations.

1.3.2 Air pollution and epigenetic aging

Similar to the investigation of metabolomics, only several epidemiological studies have investigated whether long-term air pollution exposure influences the acceleration of epigenetic aging. With a longitudinal study design, Nwanaji-Enwerem et al. first examined the impact of annual $PM_{2.5}$ and black carbon on DNAmHorvathAge, and increases of both air pollutants were associated with an increased DNAmHorvathAge [74]. A cross-sectional study based on KORA F4 found only weak associations between epigenetic aging biomarkers and long-term air pollution exposure, though pointing to sex-specific associations [75]. Another study from the Sister Study in the U.S. explored the impact of annual ambient air pollution on epigenetic aging biomarker with an additional inclusion of the second generation of epigenetic clocks (Levine's clock, DNAmPhenoAge) [74]. However, the direction of their identified associations varied between deceleration and acceleration of age. Yet, only a few studies included DNAmGrimAge and DNAmTL, while none examined all those epigenetic aging biomarkers introduced in Table 2 [76,77]. There was one study from the Scotland-based Lothian Birth Cohort which covered almost all epigenetic aging biomarkers we introduced before except for DNAmSkinBloodAge [76]. They examined the life-course associations between annual air pollution and these epigenetic aging biomarkers, but with a rather small sample size (N = 525 individuals) and an elder population (ages 70–80). Therefore, more studies covering multiple air pollutants and epigenetic aging biomarkers, particularly with a longitudinal study design, are needed to identify robust results regarding the impact of air pollution on biological aging.

2. Objectives

The main objective of this cumulative thesis was to explore the associations between air pollution (short-, medium-, as well as long-term exposure) and molecular health outcomes including blood metabolites and epigenomic aging biomarkers within longitudinal analysis based on the Augsburg population-based KORA (Cooperative Health Research in the Region of Augsburg) cohort. Additionally, I was also interested in potential modifying effects by characteristics of potential susceptibility such as age, nutrition, lifestyle factors, medication intakes, as well as underlying diseases.

Through this cumulative thesis, I aimed to answer the following questions:

i. Are there any associations between short-, medium-, or long-term air pollution exposure and metabolite level changes or epigenetic aging biomarkers?

ii. Are these potential associations modified by characteristics of susceptibility e.g. age, nutrition, lifestyle factors, medication intake, and diseases?

iii. What are the underlying biological pathways suggested by these exposure-affected metabolites or epigenetic aging biomarkers?

3. Outline of Methods

3.1 Study design and participants

All analyses for this thesis were conducted with datasets from the KORA cohort. In brief, the KORA cohort is a regional research platform for population-based surveys and subsequent follow-up studies in the fields of epidemiology, health economics, and health care research (https://www.helmholtz-munich.de/en/epi/cohort/kora). Participants were recruited from the city of Augsburg and neighbouring administrative districts covering urban and rural areas. The fourth cross-sectional health survey of the KORA cohort (KORA S4) was conducted from October 1999 to April 2001 with an inclusion of 4,261 participants aged 25–74 years. The first follow-up (KORA F4) was conducted between October 2006 and May 2008 including 3,080 participants, and the second follow-up (KORA FF4) consisted of 2,279 participants with examinations between June 2013 and September 2014. Only participants with at least two visits across the entire study period were included in the longitudinal analyses.

3.2 Air pollution

For short-term analyses, air pollutants including $PM_{2.5}$, PM with an aerodynamic diameter between 2.5 and 10 µm (PM_{coarse}), NO_2, and O_3 along with meteorological parameters such as ambient temperature and relative humidity were measured at fixed monitoring sites. Daily 24-h average concentrations of each air pollutant and meteorological parameters were calculated on the day of blood withdrawal, provided that at least 75% of the hourly values were available. Instead for O_3, the daily maximum 8-h average level was used. Furthermore, to assess the effects of immediate, short-, and medium-term exposure, we calculated moving averages for three exposure windows including: 2-day, 2-week, and 8-week moving averages. The 2-day, 2-week and 8-week averages were calculated by averaging daily concentrations over the current day and the previous one day, two or eight weeks before blood withdrawal, respectively.

For long-term analysis, residential annual averages of air pollutants including $PM_{2.5}$, PM_{10}, PM_{coarse}, $PM_{2.5}$ absorbance (a proxy of elemental carbon related to traffic exhaust, $PM_{2.5abs}$), ultrafine particles (PM ≤ 100 nm in aerodynamic diameter, represented by particle number concentration (PNC)), NO_2, nitrogen oxides (NO_x), and O_3 was estimated using land-use regression (LUR) models [78]. The LUR models were developed by regressing measured annual average concentrations from 2014–2015 at 20 selected monitoring sites located in the KORA study area against spatial predictors. Residential exposure levels were estimated by applying participants' home addresses to the fitted models.

For participants who relocated during the entire study period, updated residential addresses were used for exposure assignments, while those who did not move were assigned the same exposure levels across all visits.

3.3 Biomarker measurements

The serum metabolite profiling was examined using the AbsoluteIDQ^{TM} p180 kit (BIOCRATES Life Sciences AG, Innsbruck, Austria) for KORA S4 (March–April 2011) and FF4 (February–October 2019), enabling the simultaneous quantification of 188 metabolites [79]. For KORA F4, 163 serum metabolites were detected by the AbsoluteIDQ^{TM} p150 kit [80]. After the quality control (QC) in each study, we identified that 108 metabolites were overlapping among KORA S4, F4, and FF4, and were included in further analyses, including 12 amino acids, 12 acylcarnitines, 72 glycerophospholipids (including 32 phosphatidylcholines (PC) with acyl-acyl (diacyl) side chains, 33 PC with acyl-alkyl side chains, and seven lysophosphatidylcholines (LPC)), 11 sphingomyelins (SM) and a sum of hexoses.

For the epigenetic aging analyses, we used the Infinium HumanMethylation450K BeadChip to measure the DNA methylation in blood samples of KORA S4 and KORA F4, while the DNA methylation for KORA FF4 participants was examined by the update Infinium HumanMethylationEPIC BeadChip. Subsequently, we used the Horvath's online calculator (http://dnamage.genetics.ucla.edu/) to compute epigenetic aging biomarkers including DANmHorvathAge, DNAmHannumAge, DNAmSkinBloodAge, DNAmPhenoAge, DNAmGrimAge, and DNAmTL [49-53,61]. The age acceleration was calculated by the difference between each epigenetic clock and chronological age, except for DNAmTL.

3.4 Statistical methods

For the long-term analyses in Paper I, linear mixed-effect models with random participant-specific intercepts were performed to examine the associations between repeatedly measured metabolites and air pollutants. We included the covariates and potential confounders that have been used in previous environmental epidemiological studies, including demographics, lifestyle, and medical factors.

Instead, for the immediate, short-, and medium-term analyses in Paper II, we applied covariate-adjusted generalized additive mixed-effects models to examine the associations between each exposure window of air pollution and repeatedly measured metabolite levels. In addition to the covariates we used in paper I, we further adjusted for time trend and meteorological parameters by using regression splines to account for nonlinearity in their relationships with metabolites.

In Paper III, we again applied linear mixed-effect models with random participant-specific intercepts to examine the associations between epigenetic aging biomarkers and long-term exposure to air pollution. In addition to demographic, lifestyle, and medical factors,

we further adjusted for confounders related to the measurement of DNA methylation including houseman-estimated white cell types as fixed effects, and technical batch and chip numbers as random effects to control for technical variation. In addition, health effects of air pollution exposure may share pathways with smoking. Therefore, we performed stratified analyses for ever and never smokers to explore potential differences in their effect estimates.

To identify characteristics of susceptibility, we performed effect modification analyses by incorporating an interaction term between each air pollutant and potential effect modifiers. We considered for example age (<65 years *vs* ≥65 years; as 65 years is the current official retirement age in Germany), sex (male *vs* female), obesity (body-mass index (BMI) ≥30 kg/ m^2 *vs* <30 kg/m^2), smoking status (current/former *vs* never smoker), physical activity (low *vs* medium *vs* high), dietary pattern (adverse *vs* ordinary *vs* favorable), medication intake (yes *vs* no), hypertension (yes *vs* no), and type 2 diabetes (yes *vs* no).

Several sensitivity analyses have been conducted for all three papers to check the robustness of our results: (1) we included participants with less than two visits having complete data on air pollution, covariates, and targeted biomarkers (either metabolites or DNA methylation); (2) we addressed selection bias using inverse probability weighting (IPW); (3) we performed two-pollutant models if correlation coefficients were below 0.7; (4) we investigated the co-effects of short- and long-term air pollution exposure by including them as co-exposures in the models. In addition, we also performed paper-specific sensitivity analyses. For example, in Paper I, we further examined the effect estimates only in fasting participants and non-movers, and additionally adjusted for inflammation (hs-CRP) and storage time of blood sampless. In Paper II, we applied a crude model which did not adjust for individual covariates to investigate the acute effects from air pollutants and meteorological parameter. In Paper III, we performed several sensitivity analyses regarding the characteristics of calculating the epigenetic aging biomarkers. This included imputing missing CpG sites not covered by the Infinium HumanMethylationEPIC BeadChip, and performing a restricted epigenome-wide association study (EWAS) using only the CpG sites included in these biomarkers.

Furthermore, in all papers, we performed pathway analyses to explore the underlying biological mechanisms. In Papers I and II, we used the “Pathway Analysis” module in MetaboAnalyst 5.0 to identify the potential biological processes associated with air pollutant-related metabolites [81]. In Paper III, Ingenuity Pathway Analysis (IPA, QIAGEN Inc.) was used to determine canonical pathways enriched with those genes annotated by exposure-related CpG sites.

In all included papers, the statistical significance was determined at a *p*-value below 0.05, and further corrected by Bonferroni or Benjamini-Hochberg false discovery rate (FDR) except for in Paper III. In Papers I and II, pathways were identified as the most relevant pathways if they had a *p*-value ≤ 0.1, or an impact value > 0.5 with a *p*-value ≤ 0.3. In Paper III, the pathway was determined if the *p*-value was less than 0.05. We used the software R (version 3.6.2 for Paper I, version 4.1.2 for Paper II, and version 4.3.1 for

Paper III) to perform the analyses reported in this cumulative thesis. **Figure 2** summarizes the workflow of our methods.

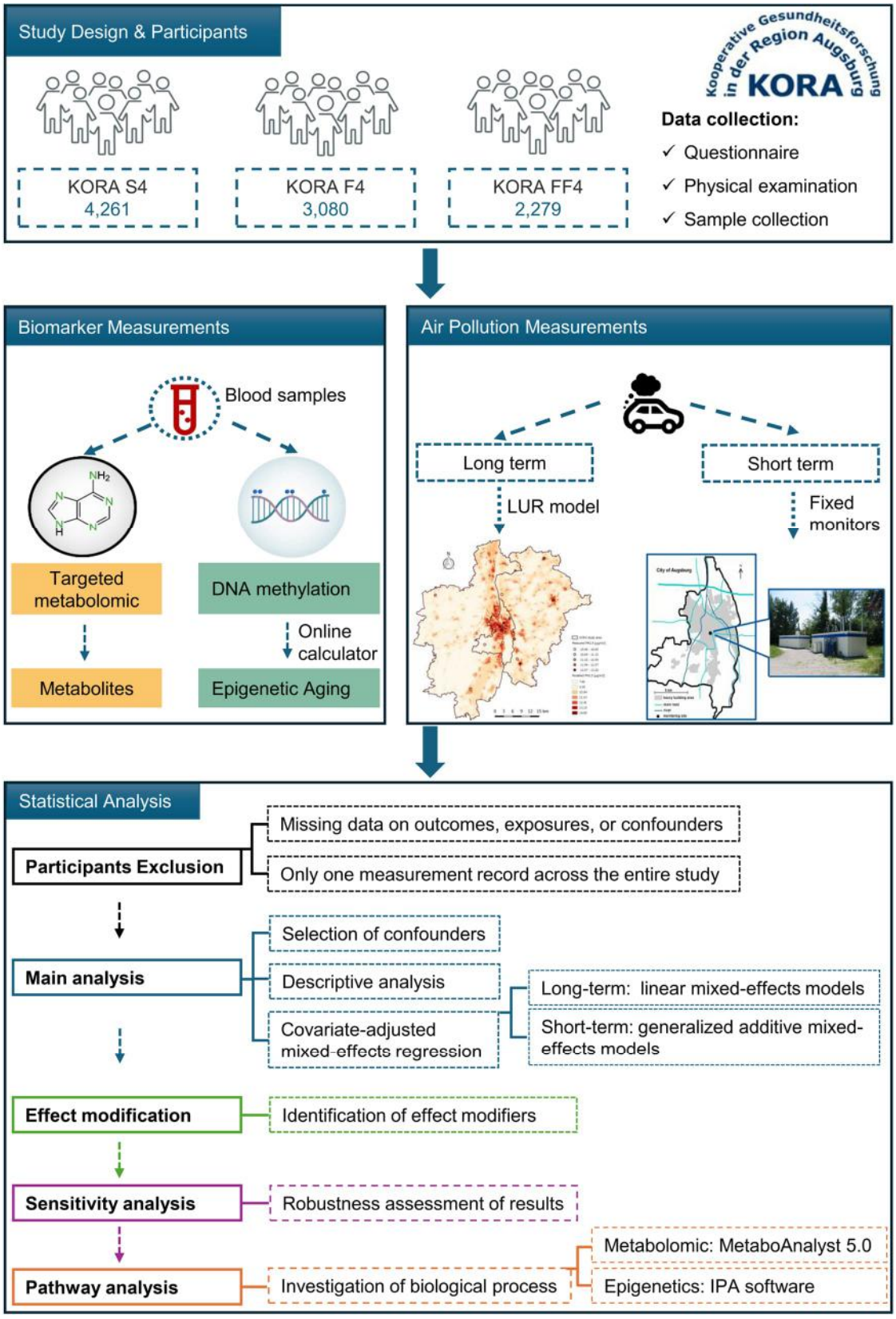


Figure 2. The workflow of methods included in this thesis.

4. Results

The first publication explored the associations between long-term exposure to ambient air pollution and specific serum metabolites, accounting for individual characteristics such as lifestyle factors, diseases, and medication intake in a longitudinal cohort setting (KORA S4, KORA F4, and KORA FF4) [82]. Nine significant associations were identified between air pollutants and metabolites, primarily from the phosphatidylcholine (PC) subgroup.

Specifically, higher exposure to $PM_{2.5abs}$, PM_{coarse}, and NO_2 was associated with lower levels of PC acyl-alkyl C34:2 (PC ae C34:2) and PC ae C36:3. Additionally, $PM_{2.5abs}$ and PM_{coarse} were negatively linked to PC ae C36:4, while $PM_{2.5abs}$ was negatively associated with PC ae C34:3. Pathway analysis indicated that the identified metabolites are involved in glycerophospholipid, linoleic acid, and alpha-linolenic acid metabolism. Stronger associations were observed in individuals with certain characteristics, including older age, obesity, lower education levels, low physical activity, and unhealthy dietary patterns.

- **Key finding 1:** Long-term exposure to air pollution was related to metabolic changes, particularly reductions in specific PC (e.g., PC ae C34:2, PC ae C36:3, PC ae C36:4, and PC ae C34:3), pointing to glycerophospholipid and fatty acid metabolism as underlying biological processes. As potential susceptible subgroups for metabolic impacts of air pollution, we identified individuals with older age, obesity, lower education, low physical activity, and unhealthy dietary habits.

The second publication investigated the effects of immediate (2-day moving average), short- (2-week moving average), and medium-term (8-week moving) exposures to ambient air pollution on serum metabolites using the same KORA survey and follow-up data as in Paper I [83]. Multiple significant associations were identified between air pollutants ($PM_{2.5}$, PM_{coarse}, NO_2, and O_3) across all three exposure windows and metabolites, particularly within the PC metabolite subgroup, which is consistent with our findings from the long-term analyses. Longer exposure windows, such as 8-week moving averages, showed the strongest and most consistent effects, especially for NO_2 with PC (e.g., PC aa C40:4 and PC ae C42:5). In specific:

1) **$PM_{2.5}$**: Medium-term exposure was linked to arginine (Arg), tryptophan (Trp), and eight PC. Short-term exposure (2-week) was associated with Trp and one PC (PC aa C30:0).
2) **PM_{coarse}**: Medium-term exposure showed positive associations with multiple amino acids (e.g., glycine (Gly), methionine (Met), ornithine (Orn), phenylalanine (Phe), serine (Ser) and threonine (Thr)), one lysophosphatidylcholine (LPC a C16:0), one PC ae C40:4, and one sphingomyelin (SM C16:0). Immediate exposure was negatively associated with PC aa C40:4 and positively with PC ae C44:3.

3) **NO_2**: Showed the largest number of associations across metabolites, particularly in medium-term windows with 33 PC, five LPC, seven SM, and three amino acids (e.g., Arg, Trp, tyrosine (Tyr)), and one acylcarnitine (C0).

4) **O_3**: Few metabolites showed significant associations. For example, Trp showed consistent associations across short- and medium-term exposures, and SM (OH) C24:1 was related to immediate exposure.

In summary, Trp, PC aa C40:4, and PC ae C42:5 showed robust and consistent associations. These exposure related metabolites linked to air pollutants were associated with several pathways: 1) glycerophospholipid metabolism in short-term PM_{coarse} exposure; 2) cysteine and methionine metabolism, glyoxylate and dicarboxylate metabolism, and glycine, serine, and threonine metabolism in medium-term PM_{coarse} exposure; 3) glycerophospholipid metabolism (all windows) and phenylalanine, tyrosine, and tryptophan biosynthesis (short-term) in NO_2 exposures. However, these associations became insignificant after FDR correction.

- **Key finding 2:** Glycerophospholipid metabolism, a pathway linked to inflammation and oxidative stress, was consistently associated with NO_2 across all exposure windows, as well as with short-term PM_{coarse} and medium-term $PM_{2.5}$ exposure. This suggested, it might be a key biological mechanism influenced by air pollution.

The third publication revealed associations between epigenetic aging biomarkers and air pollution [84]. First, DNAmTL showed robust negative associations with multiple air pollutants, including particulate air pollutants ($PM_{2.5}$, PM_{10}, PNC, PM_{coarse}, $PM_{2.5abs}$) and nitrogen oxides (NO_2 and NO_x). Second, DNAmGrimAge was positively associated with these pollutants, but associations were weakened when additionally adjusting for smoking status. In stratified analyses, ever smokers showed stronger associations between air pollutants and several epigenetic aging markers such as DNAmHorvathAge, DNAmPhenoAge, and DNAmTL. Moreover, DNAmGrimAge and DNAmSkinBloodAge also showed stronger effects compared to the pooled analyses, though estimates reached not statistically significance. In contrast, never smokers showed fewer significant associations, highlighting greater susceptibility among ever smokers.

Effect modification analysis identified hypertension as a significant modifier for DNAmTL associations with PM ($PM_{2.5}$, PM_{10}, PM_{coarse}, PNC, and $PM_{2.5abs}$) and nitrogen oxides (NO_2 and NO_x), while no consistent patterns were observed for other modifiers such as age, sex, obesity, or education. The limited EWAS indicated no significant associations between air pollutants and 1,253 CpGs in the main analysis or stratified analyses among never smokers. However, in ever smokers, two CpGs (gene annotation: solute carrier family 39 member 5 (*SLC39A5*) and soluble epoxide hydrolase (*EPHX2*)) were significantly associated with PM_{coarse}. Top pathways identified by these annotated genes are involved in triacylglycerol biosynthesis and lipid metabolism.

- **Key finding 3:** Epigenetic age acceleration, particularly DNAmTL and DNAmGrimAge, was strongly associated with air pollution exposure, with ever smokers showing greater susceptibility. Moreover, lipid metabolism was identified as a potential biological pathway in ever smokers.

5. Discussion

This research work advances our understanding of the interplay between air pollution, metabolomics, and epigenetic aging, addressing key gaps in environmental health research. Leveraging comprehensive longitudinal datasets (KORA S4, KORA F4, and KORA FF4), we systematically examined the metabolic and epigenetic effects of short-, medium-, and long-term air pollution exposure.

In relation to our first research question, we found significant associations between various air pollutants and both serum metabolites and epigenetic aging biomarkers across multiple exposure periods. These results provided robust evidence that air pollution exposure impacts health at both the metabolic and epigenomic levels. In addressing the second research question—whether individual characteristics modify the effects of air pollution on molecular outcomes—we observed that individual susceptibility factors, particularly age, smoking status, lifestyle behaviors, and socioeconomic status, modified these associations. Specifically, individuals with older age, obesity, low physical activity, lower educational attainment, and unhealthy dietary patterns may be more susceptible to the metabolic effects of air pollution. Furthermore, ever smokers showed stronger associations between air pollution exposure and epigenetic aging biomarkers, indicating increased susceptibility within this subgroup. Regarding our third research, pathway analyses highlighted glycerophospholipid metabolism and lipid-related pathways as potentially underlying mechanisms linking air pollution exposure with both metabolic changes and accelerated epigenetic aging.

These findings contribute to the growing body of evidence on molecular responses to environmental exposures and support the development of personalized risk assessments and targeted public health interventions. Detailed results of this cumulative thesis have been discussed in each corresponding publication.

In the following sections, we present: (1) interplay of metabolomic and epigenetic aging: uncovering common mechanisms or pathways; (2) susceptible subgroups; (3) health impacts of air pollution across shorter and longer exposure periods under consideration of air quality regulation; (4) strengths and limitations of this work; (5) outlook for future research.

5.1 Interplay of metabolomic and epigenetic aging: uncovering common mechanisms or pathways

To summarize the results from Papers I and II, PC was the only consistently overlapping metabolite subgroup identified in both long- and short-term analyses [82,83]. PC is one of the most abundant phospholipids in all mammalian cell membranes [85]. As a key component of lipoproteins, PC belongs to the glycerophospholipid family and plays a central role in glycerophospholipid metabolism [85]. It is essential for maintaining overall

health, and alterations in its metabolism may indicate an imbalance between anti-inflammatory and pro-inflammatory processes [86]. Our long-term analysis also showed elevated hs-CRP levels following long-term air pollution exposure, supporting the presence of an inflammatory state consistent with the observed metabolic alterations [82]. PC metabolism contributes to disease development, and has been reported to be associated with the development of diseases e.g. type 2 diabetes [87], CVD [88], and AD [89].

Metabolites are intermediates or end products of metabolism that can influence cellular physiology by modulating other “omics” levels and reflecting changes induced by various exposures [90]. In our findings from Paper III, we identified a positive association between air pollution exposure and epigenetic age acceleration, particularly in DNAmTL [84]. From the limited EWAS analysis, the annotated genes *SLC39A5* and *EPHX2* were identified as the key targeted genes. *EPHX2*, encoding soluble epoxide hydrolase (*sEH*), plays a crucial role in lipid metabolism by converting bioactive epoxides into their less active diols, influencing inflammation, cardiovascular health, and metabolic disorders [91]. It has been reported that variation in *EPHX2* gene is associated with an increased risk of incident coronary heart disease (CHD) [92] and ischemic stroke [93]. Due to its involvement in various pathophysiological processes, *EPHX2* inhibition has emerged as a potential therapeutic strategy for inflammatory, cardiovascular, and pulmonary diseases [94-96]. From our pathway analysis, *EPHX2* has been implicated in lipid metabolism regulation, which may also underlie the observed metabolic effects.

Lipid metabolism plays a crucial role in aging and aging-related diseases, and specific phospholipid blood profiles have been shown to change with age and may be associated with exceptional human longevity [97]. Considering the PC alternations in our metabolomic analysis and the *EPHX2*-regulated lipid metabolism from our epigenetic aging analyses, the lipid metabolic perturbation particularly the PC metabolism was suggested as the representative biomarkers linking air pollution exposure to adverse health outcomes. Further supporting this, metabolomics-related studies using the KORA cohort data have reported that a decrease in certain acyl-alkyl PC was associated with aging [98].

These findings emphasize the potential of PC metabolism alterations as a biomarker of metabolic and inflammatory disturbances. Given that both PC metabolism and *EPHX2* function implicated in lipid regulation from our epigenetic aging analysis, their interplay may represent a mechanistic link between air pollution exposure and its related diseases.

5.2 Susceptible subgroups

Previous epidemiological studies showed that participants can respond differently to air pollution exposure and reported specific subgroups to be more susceptible than others [19,99]. In general, elder adults, children, and individuals with pre-existing diseases, unhealthy lifestyle factors (e.g. inactive physical activity, smoking, and nutrition status), and low socioeconomic status might be more susceptible to air pollution [100-103]. The variability in individual responses to air pollution exposure highlights the necessity of identifying these susceptibility factors. A deeper understanding of population susceptibility can

strengthen the scientific foundation for risk assessment, and support the development of policies to reduce the effects of air pollution particularly for the susceptible population.

In summary, the findings of our effect modification analyses were not consistent. Our findings from the long-term metabolomics analysis underscored the significant role of glycerophospholipid metabolism in mediating the health effects of air pollution. These metabolic pathways are particularly relevant when considering susceptibility factors such as obesity, lower physical activity and disadvantageous socioeconomic status (e.g. low educational attainment). Furthermore, the epigenetic aging analysis indicated that ever smokers had accelerated epigenetic aging compared to never smokers. Therefore, our findings support that obesity, smoking, and low socioeconomic status may enhance the adverse impacts from air pollution exposure, while active physical activity and a healthier dietary pattern in the long run attenuate air pollution-induced harmful effects through anti-inflammatory effects and reductions in oxidative stress [104,105]. In summary, our results suggested that elder people and ever smokers are more susceptible to air pollution, while healthy lifestyle may exert anti-inflammatory effects and reduce the risk of inflammation-associated diseases. Even though reducing pollution levels is the final goal, fully protecting the population remains challenging. Prioritizing the identification of individual susceptibility characteristics and developing targeted interventions would be an important approach for future research and policy.

5.3 Health impacts of air pollution across shorter and longer exposure periods under consideration of air quality regulation

Results from this cumulative dissertation highlight the significant health impacts of air pollution, with long-term exposure showing the most robust effects. In Paper I, we identified four metabolites from the PC subgroup with unsaturated long-chain fatty acids that were negatively influenced by long-term exposure to ambient air pollution [82]. In Paper II, we observed that a broader range of metabolites and metabolic pathways were affected by immediate, short-, and medium-term exposure, with medium-term exposure exhibiting the strongest and most consistent effects [83]. Given the varying exposure windows of immediate, short-, medium-, and long-term air pollution, dynamic metabolic alterations may serve as the underlying mechanism.

Despite these variations, our findings confirm the adverse health effects of air pollution across all exposure windows. During the study period in Augsburg, annual average concentrations of $PM_{2.5}$, PM_{10} and NO_2 at participant's residences remained below the European Union (EU) air quality standard values but exceeded WHO AQG values [2,106]. Similar for the short-term exposure, daily averages of $PM_{2.5}$ from all three surveys (KORA S4, KORA F4, KORA FF4) were slightly higher or similar to WHO AQG values, while daily averages of NO_2 consistently exceeded WHO AQG values.

These findings emphasize the urgent need for stricter air pollution control measures to protect public health. Aligning regulations with the latest WHO AQG is essential to reduce air pollution levels and safeguard public health. Recognizing this urgency, the EU Commission released the revised Ambient Air Quality Directive in October 2024, bringing the 2030 EU air quality standards closer to WHO recommendation levels [107]. More details are presented in **Table 3**. However, continued efforts are necessary to ensure effective implementation to achieve cleaner air and better protection for public health.

Table 3. Limit values of air pollutants in WHO Air Quality Guidelines, previous and new EU Ambient Air Quality Directives [2,106,107].

Air pollutants ($\mu g/m^3$)	WHO Guidelines (from 2021)	Previous EU Directives (from 2005)	New EU Directives 2024 (from 2030 on)
$PM_{2.5}$ (annual)	5	25	10
$PM_{2.5}$ (24-hour)	15	---	25
PM_{10} (annual)	15	40	20
PM_{10} (day)	45	50	45
NO_2 (annual)	10	40	20
NO_2 (24-hour)	25	---	50
O_3 (peak season)*	60	120	120
O_3 (24-hour)	100	---	100

*Peak season refers to the period of six consecutive months when the average ozone concentration reaches the highest level (calculated as a six-month moving average).

5.4 Strengths and limitations

Given the methods and data used in this cumulative thesis, our analyses have several strengths and limitations in assessing the health effects of ambient air pollution exposure using targeted metabolomics and epigenetic aging biomarkers. The three analyses share several main strengths. Firstly, the use of repeated biomarker measurements enhanced statistical power and reduced residual confounding compared to a cross-sectional design. Secondly, our results are based on the well-characterized KORA cohort, ensuring reliable and standardized data collection. Thirdly, the targeted metabolomics approach used in the first two publications provided precise metabolite annotation and quantification, minimizing the risk of false identifications compared to untargeted analysis. Additionally, the longitudinal design allowed a robust evaluation of long-term health effects. Furthermore, our analyses accounted for both external and intrinsic factors, including lifestyle and dietary influences, which are known to affect the human metabolome and epigenome, and provided insights into individual susceptibility.

Despite these strengths, the analyses also share some common limitations. Even though the targeted metabolomics is precise to define metabolites, it limits the discovery of new biomarkers and does not capture the entire metabolome compared to untargeted metabolomics. Also, exposure misclassification remains a challenge, as air pollution estimates were based on spatial models or fixed-site monitoring, which may not reflect true individual exposure. Potential measurement errors due to spatial and temporal misalignments could also impact health effect estimates. Furthermore, all findings were based on the KORA cohort. Therefore, the lack of replication in other populations limits the generalizability. Lastly, epigenetic aging markers based on blood samples may skew pathway identification toward inflammatory responses, making it impossible to assess methylation changes in other organs or tissues.

5.5 Outlook: Future research

In summary, the findings of this cumulative thesis indicate that both short- and long-term exposure to air pollutants are associated with changes in metabolomic and epigenetic aging signatures. The thesis highlights the ability to strengthen our understanding of the interplay between air pollution and health effects by using omics biomarkers and calls for future investigation in this research area.

While this study performed targeted metabolomics with precise annotation and quantification, integrating untargeted approaches could enhance biomarker discovery and provide a more comprehensive assessment of metabolic alterations linked to air pollution exposure. In a recent study based on the UK Airwave cohort, researchers developed a model to predict biological age using untargeted metabolic profiling [108]. Their findings suggested that metabolomics could serve as a promising tool for assessing biological age, providing a complementary insight into current epigenetic clocks. Future research could explore how air pollution influences biological aging by combining metabolomics with epigenetic clocks.

On the one hand, future studies should apply multi-omics approaches by integrating metabolomics, epigenetics, and proteomics to gain a comprehensive understanding of how air pollution influences biological systems at different levels. The predictive models of biological age built by multi-omics integration would improve risk assessments for long-term health effects of air pollution exposure. Moreover, to perform analyses within several cohorts with diverse populations across different regions would highly improve the generalizability of findings.

6. Conclusions

This dissertation provides comprehensive insights into metabolic and epigenetic biomarkers associated with air pollution exposure. The findings revealed significant alterations in serum metabolites, particularly phosphatidylcholines linked to glycerophospholipid metabolism, under both short- and long-term exposure to air pollution. These metabolic changes were more pronounced in susceptible populations, such as the elderly, individuals with obesity, and those with unhealthy lifestyles. In addition, the study underscores the role of epigenetic age acceleration biomarkers in mediating the impacts of air pollution, with ever smokers showing greater susceptibility. The identified underlying pathway related to lipid metabolism further emphasizes the biological progresses in the health impacts of air pollution.

Within this thesis, we identified reliable biomarkers of air pollution-related diseases which are crucial for understanding the health impacts of air pollution and developing targeted interventions. Identifying air pollution associated biomarkers is critical for early disease detection and risk assessment. To reduce these impacts, well-planned control measures on air pollution aligned with WHO air quality guideline are essential. By integrating biomarker research with pollution reduction efforts, we can promote disease prevention and environmental health strategies, and further alleviate the health burden associated with air pollution.

References

1. Peters, A, Künzli, N, Forastiere, F, and Hoffmann, B. 2019. Promoting clean air: combating fake news and denial. The Lancet Respiratory Medicine. 7(8): 650-652.
2. World Health Organization (WHO), WHO global air quality guidelines: particulate matter ($PM_{2.5}$ and PM_{10}), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. 2021, Geneva: World Health Organization.
3. Landrigan, PJ, Fuller, R, Acosta, NJR, Adeyi, O, Arnold, R, Basu, N, et al. 2018. The Lancet Commission on pollution and health. The Lancet. 391(10119): 462-512.
4. Brauer, M, Roth, GA, Aravkin, AY, Zheng, P, Abate, KH, Abate, YH, et al. 2024. Global burden and strength of evidence for 88 risk factors in 204 countries and 811 subnational locations, 1990–2021: a systematic analysis for the Global Burden of Disease Study 2021. The Lancet. 403(10440): 2162-2203.
5. Chen, J and Hoek, G. 2020. Long-term exposure to PM and all-cause and cause-specific mortality: A systematic review and meta-analysis. Environment International. 143: 105974.
6. Huangfu, P and Atkinson, R. 2020. Long-term exposure to NO2 and O3 and all-cause and respiratory mortality: A systematic review and meta-analysis. Environment International. 144: 105998.
7. Xue, T, Tong, M, Wang, M, Yang, X, Wang, Y, Lin, H, et al. 2023. Health Impacts of Long-Term NO2 Exposure and Inequalities among the Chinese Population from 2013 to 2020. Environmental Science & Technology. 57(13): 5349-5357.
8. Meng, X, Liu, C, Chen, R, Sera, F, Vicedo-Cabrera, AM, Milojevic, A, et al. 2021. Short term associations of ambient nitrogen dioxide with daily total, cardiovascular, and respiratory mortality: multilocation analysis in 398 cities. BMJ. 372: n534.
9. Wang, K, Yuan, Y, Wang, Q, Yang, Z, Zhan, Y, Wang, Y, et al. 2023. Incident risk and burden of cardiovascular diseases attributable to long-term NO2 exposure in Chinese adults. Environment International. 178: 108060.
10. Sun, W, Lu, K, and Li, R. 2024. Global estimates of ambient NO2 concentrations and long-term health effects during 2000–2019. Environmental Pollution. 359: 124562.
11. Crouse, DL, Peters, PA, van Donkelaar, A, Goldberg, MS, Villeneuve, PJ, Brion, O, et al. 2012. Risk of nonaccidental and cardiovascular mortality in relation to long-term exposure to low concentrations of fine particulate matter: a Canadian national-level cohort study. Environ Health Perspect. 120(5): 708-14.
12. Yazdi, MD, Wang, Y, Di, Q, Requia, WJ, Wei, Y, Shi, L, et al. 2021. Long-term effect of exposure to lower concentrations of air pollution on mortality among US Medicare participants and vulnerable subgroups: a doubly-robust approach. The Lancet Planetary Health. 5(10): e689-e697..
13. Stafoggia, M, Oftedal, B, Chen, J, Rodopoulou, S, Renzi, M, Atkinson, RW, et al. 2022. Long-term exposure to low ambient air pollution concentrations and mortality among 28 million people: results from seven large European cohorts within the ELAPSE project. The Lancet Planetary Health. 6(1): e9-e18.

14. Brunekreef, B, Strak, M, Chen, J, Andersen, ZJ, Atkinson, R, Bauwelinck, M, et al. 2021. Mortality and Morbidity Effects of Long-Term Exposure to Low-Level PM(2.5), BC, NO(2), and O(3): An Analysis of European Cohorts in the ELAPSE Project. Res Rep Health Eff Inst. 2021(208): 1-127.

15. Beverland, IJ, Cohen, GR, Heal, MR, Carder, M, Yap, C, Robertson, C, et al. 2012. A comparison of short-term and long-term air pollution exposure associations with mortality in two cohorts in Scotland. Environ Health Perspect. 120(9): 1280-5.

16. Mann, JK, Balmes, JR, Bruckner, TA, Mortimer, KM, Margolis, HG, Pratt, B, et al. 2010. Short-term effects of air pollution on wheeze in asthmatic children in Fresno, California. Environ Health Perspect. 118(10): 1497-502.

17. Dauchet, L, Hulo, S, Cherot-Kornobis, N, Matran, R, Amouyel, P, Edmé, J-L, et al. 2018. Short-term exposure to air pollution: Associations with lung function and inflammatory markers in non-smoking, healthy adults. Environment International. 121: 610-619.

18. Gao, H, Wang, K, W. Au, W, Zhao, W, and Xia, Z-l. 2020. A Systematic Review and Meta-Analysis of Short-Term Ambient Ozone Exposure and COPD Hospitalizations. International Journal of Environmental Research and Public Health. 17(6): 2130.

19. Makri, A and Stilianakis, NI. 2008. Vulnerability to air pollution health effects. International Journal of Hygiene and Environmental Health. 211(3): 326-336.

20. Orellano, P, Reynoso, J, Quaranta, N, Bardach, A, and Ciapponi, A. 2020. Short-term exposure to particulate matter (PM10 and PM2.5), nitrogen dioxide (NO2), and ozone (O3) and all-cause and cause-specific mortality: Systematic review and meta-analysis. Environment International. 142: 105876.

21. Guarnieri, M and Balmes, JR. 2014. Outdoor air pollution and asthma. Lancet. 383(9928): 1581-92.

22. Brook, RD, Rajagopalan, S, Pope, CA, Brook, JR, Bhatnagar, A, Diez-Roux, AV, et al. 2010. Particulate Matter Air Pollution and Cardiovascular Disease. Circulation. 121(21): 2331-2378.

23. Rajagopalan, S and Brook, RD. 2012. Air Pollution and Type 2 Diabetes: Mechanistic Insights. Diabetes. 61(12): 3037-3045.

24. Peters, R, Ee, N, Peters, J, Booth, A, Mudway, I, and Anstey, KJ. 2019. Air Pollution and Dementia: A Systematic Review. J Alzheimers Dis. 70(s1): S145-s163.

25. Nemmar, A, Hoet, PHM, Vanquickenborne, B, Dinsdale, D, Thomeer, M, Hoylaerts, MF, et al. 2002. Passage of Inhaled Particles Into the Blood Circulation in Humans. Circulation. 105(4): 411-414.

26. Zanobetti, A, Baccarelli, A, and Schwartz, J. 2011. Gene–Air Pollution Interaction and Cardiovascular Disease: A Review. Progress in Cardiovascular Diseases. 53(5): 344-352.

27. Panni, T, Mehta, AJ, Schwartz, JD, Baccarelli, AA, Just, AC, Wolf, K, et al. 2016. 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. Environmental Health Perspectives. 124(7): 983-990.

28. Wolf, K, Popp, A, Schneider, A, Breitner, S, Hampel, R, Rathmann, W, et al. 2016. Association Between Long-term Exposure to Air Pollution and Biomarkers Related to Insulin Resistance, Subclinical Inflammation, and Adipokines. Diabetes. 65(11): 3314-3326.

29. Peters, A, Nawrot, TS, and Baccarelli, AA. 2021. Hallmarks of environmental insults. Cell. 184(6): 1455-1468.

30. Rappaport, SM, Barupal, DK, Wishart, D, Vineis, P, and Scalbert, A. 2014. The Blood Exposome and Its Role in Discovering Causes of Disease. Environmental Health Perspectives. 122(8): 769-774.

31. Patti, GJ, Yanes, O, and Siuzdak, G. 2012. Metabolomics: the apogee of the omics trilogy. Nature Reviews Molecular Cell Biology. 13(4): 263-269.

32. Holmes, E, Wilson, ID, and Nicholson, JK. 2008. Metabolic Phenotyping in Health and Disease. Cell. 134(5): 714-717.

33. Ribbenstedt, A, Ziarrusta, H, and Benskin, JP. 2018. Development, characterization and comparisons of targeted and non-targeted metabolomics methods. PLOS ONE. 13(11): e0207082.

34. Griffiths, WJ, Koal, T, Wang, Y, Kohl, M, Enot, DP, and Deigner, H-P. 2010. Targeted Metabolomics for Biomarker Discovery. Angewandte Chemie International Edition. 49(32): 5426-5445.

35. Gonzalez-Covarrubias, V, Martínez-Martínez, E, and Del Bosque-Plata, L. 2022. The Potential of Metabolomics in Biomedical Applications. Metabolites. 12(2).

36. Roberts, LD, Souza, AL, Gerszten, RE, and Clish, CB. 2012. Targeted metabolomics. Curr Protoc Mol Biol. Chapter 30: Unit 30.2.1-24.

37. Schrimpe-Rutledge, AC, Codreanu, SG, Sherrod, SD, and McLean, JA. 2016. Untargeted Metabolomics Strategies—Challenges and Emerging Directions. Journal of the American Society for Mass Spectrometry. 27(12): 1897-1905.

38. Vinayavekhin, N and Saghatelian, A. 2010. Untargeted metabolomics. Curr Protoc Mol Biol. Chapter 30: Unit 30.1.1-24.

39. Wang, TJ, Larson, MG, Vasan, RS, Cheng, S, Rhee, EP, McCabe, E, et al. 2011. Metabolite profiles and the risk of developing diabetes. Nature Medicine. 17(4):

40. Bhattacharya, S, Granger, CB, Craig, D, Haynes, C, Bain, J, Stevens, RD, et al. 2014. Validation of the association between a branched chain amino acid metabolite profile and extremes of coronary artery disease in patients referred for cardiac catheterization. Atherosclerosis. 232(1): 191-196.

41. Akyol, S, Ugur, Z, Yilmaz, A, Ustun, I, Gorti, SKK, Oh, K, et al. 2021. Lipid Profiling of Alzheimer's Disease Brain Highlights Enrichment in Glycerol(phospho)lipid, and Sphingolipid Metabolism. Cells. 10(10): 2591.

42. Bürkle, A, Moreno-Villanueva, M, Bernhard, J, Blasco, M, Zondag, G, Hoeijmakers, JHJ, et al. 2015. MARK-AGE biomarkers of ageing. Mechanisms of Ageing and Development. 151: 2-12.

43. Crimmins, EM and Saito, Y. 2001. Trends in healthy life expectancy in the United States, 1970–1990: gender, racial, and educational differences. Social Science & Medicine. 52(11): 1629-1641.

44. Lubetkin, EI, Jia, H, Franks, P, and Gold, MR. 2005. Relationship Among Sociodemographic Factors, Clinical Conditions, and Health-related Quality of Life: Examining the EQ-5D in the U.S. General Population. Quality of Life Research. 14(10): 2187-2196.

45. Jones, MJ, Goodman, SJ, and Kobor, MS. 2015. DNA methylation and healthy human aging. Aging Cell. 14(6): 924-932.

46. Ferrucci, L, Levine, ME, Kuo, PL, and Simonsick, EM. 2018. Time and the Metrics of Aging. Circ Res. 123(7): 740-744.

47. Horvath, S and Raj, K. 2018. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. Nature Reviews Genetics. 19(6): 371-384.

48. Rider, CF and Carlsten, C. 2019. Air pollution and DNA methylation: effects of exposure in humans. Clin Epigenetics. 11(1): 131.

49. Hannum, G, Guinney, J, Zhao, L, Zhang, L, Hughes, G, Sadda, S, et al. 2013. Genome-wide methylation profiles reveal quantitative views of human aging rates. Mol Cell. 49(2): 359-367.

50. Horvath, S. 2013. DNA methylation age of human tissues and cell types. Genome Biology. 14(10): 3156.

51. Horvath, S, Oshima, J, Martin, GM, Lu, AT, Quach, A, Cohen, H, et al. 2018. Epigenetic clock for skin and blood cells applied to Hutchinson Gilford Progeria Syndrome and ex vivo studies. Aging (Albany NY). 10(7): 1758-1775.

52. Levine, ME, Lu, AT, Quach, A, Chen, BH, Assimes, TL, Bandinelli, S, et al. 2018. An epigenetic biomarker of aging for lifespan and healthspan. Aging (Albany NY). 10(4): 573-591.

53. Lu, AT, Quach, A, Wilson, JG, Reiner, AP, Aviv, A, Raj, K, et al. 2019. DNA methylation GrimAge strongly predicts lifespan and healthspan. Aging (Albany NY). 11(2): 303-327.

54. Dhingra, R, Nwanaji-Enwerem, JC, Samet, M, and Ward-Caviness, CK. 2018. DNA Methylation Age-Environmental Influences, Health Impacts, and Its Role in Environmental Epidemiology. Curr Environ Health Rep. 5(3): 317-327.

55. Mather, KA, Jorm, AF, Parslow, RA, and Christensen, H. 2010. Is Telomere Length a Biomarker of Aging? A Review. The Journals of Gerontology: Series A. 66A(2): 202-213.

56. Chen, BH, Carty, CL, Kimura, M, Kark, JD, Chen, W, Li, S, et al. 2017. Leukocyte telomere length, T cell composition and DNA methylation age. Aging (Albany NY). 9(9): 1983-1995.

57. Blackburn, EH, Epel, ES, and Lin, J. 2015. Human telomere biology: A contributory and interactive factor in aging, disease risks, and protection. Science. 350(6265): 1193-1198.

58. Kimura, M, Stone, RC, Hunt, SC, Skurnick, J, Lu, X, Cao, X, et al. 2010. Measurement of telomere length by the Southern blot analysis of terminal restriction fragment lengths. Nature Protocols. 5(9): 1596-1607.

59. Denham, J, Marques, FZ, and Charchar, FJ. 2014. Leukocyte telomere length variation due to DNA extraction method. BMC Research Notes. 7(1): 877.

60. Verhulst, S, Susser, E, Factor-Litvak, PR, Simons, MJP, Benetos, A, Steenstrup, T, et al. 2015. Commentary: The reliability of telomere length measurements. International Journal of Epidemiology. 44(5): 1683-1686.

61. Lu, AT, Seeboth, A, Tsai, PC, Sun, D, Quach, A, Reiner, AP, et al. 2019. DNA methylation-based estimator of telomere length. Aging (Albany NY). 11(16): 5895-5923.

62. Haas, R, Zelezniak, A, Iacovacci, J, Kamrad, S, Townsend, S, and Ralser, M. 2017. Designing and interpreting 'multi-omic' experiments that may change our understanding of biology. Current Opinion in Systems Biology. 6: 37-45.

63. Sturla, SJ. 2007. DNA adduct profiles: chemical approaches to addressing the biological impact of DNA damage from small molecules. Current Opinion in Chemical Biology. 11(3): 293-299.

64. Ulrey, CL, Liu, L, Andrews, LG, and Tollefsbol, TO. 2005. The impact of metabolism on DNA methylation. Human Molecular Genetics. 14(suppl_1): R139-R147.

65. Ward-Caviness, CK, Breitner, S, Wolf, K, Cyrys, J, Kastenmüller, G, Wang-Sattler, R, et al. 2016. Short-term NO2 exposure is associated with long-chain fatty acids in prospective cohorts from Augsburg, Germany: results from an analysis of 138 metabolites and three exposures. International Journal of Epidemiology. 45(5): 1528-1538.

66. Nassan, FL, Wang, C, Kelly, RS, Lasky-Su, JA, Vokonas, PS, Koutrakis, P, et al. 2021. Ambient PM2.5 species and ultrafine particle exposure and their differential metabolomic signatures. Environment International. 151: 106447.

67. Breitner, S, Schneider, A, Devlin, RB, Ward-Caviness, CK, Diaz-Sanchez, D, Neas, LM, et al. 2016. Associations among plasma metabolite levels and short-term exposure to PM2.5 and ozone in a cardiac catheterization cohort. Environment International. 97: 76-84.

68. Mu, L, Niu, Z, Blair, RH, Yu, H, Browne, RW, Bonner, MR, et al. 2019. Metabolomics Profiling before, during, and after the Beijing Olympics: A Panel Study of Within-Individual Differences during Periods of High and Low Air Pollution. Environmental Health Perspectives. 127(5): 057010.

69. Feng, B, Liu, C, Yi, T, Song, X, Wang, Y, Liu, S, et al. 2021. Perturbation of amino acid metabolism mediates air pollution associated vascular dysfunction in healthy adults. Environmental Research. 201: 111512.

70. Hood, RB, Liang, D, Tang, Z, Kloog, I, Schwartz, J, Laden, F, et al. 2022. Length of PM2.5 exposure and alterations in the serum metabolome among women undergoing infertility treatment. Environmental Epidemiology. 6(1): e191.

71. Liang, D, Moutinho, JL, Golan, R, Yu, T, Ladva, CN, Niedzwiecki, M, et al. 2018. Use of high-resolution metabolomics for the identification of metabolic signals associated with traffic-related air pollution. Environment International. 120: 145-154.

72. Liang, D, Ladva, CN, Golan, R, Yu, T, Walker, DI, Sarnat, SE, et al. 2019. Perturbations of the arginine metabolome following exposures to traffic-related air pollution in a panel of commuters with and without asthma. Environment International. 127: 503-513.

73. Nassan, FL, Kelly, RS, Kosheleva, A, Koutrakis, P, Vokonas, PS, Lasky-Su, JA, et al. 2021. Metabolomic signatures of the long-term exposure to air pollution and temperature. Environmental Health. 20(1): 3. 10.1186/s12940-020-00683-x.

74. White, AJ, Kresovich, JK, Keller, JP, Xu, Z, Kaufman, JD, Weinberg, CR, et al. 2019. Air pollution, particulate matter composition and methylation-based biologic age. Environment International. 132: 105071.

75. Ward-Caviness, CK, Nwanaji-Enwerem, JC, Wolf, K, Wahl, S, Colicino, E, Trevisi, L, et al. 2016. Long-term exposure to air pollution is associated with biological aging. Oncotarget. 7(46): 74510-74525.

76. Baranyi, G, Deary, IJ, McCartney, DL, Harris, SE, Shortt, N, Reis, S, et al. 2022. Life-course exposure to air pollution and biological ageing in the Lothian Birth Cohort 1936. Environment International. 169: 107501.

77. Koenigsberg, SH, Chang, CJ, Ish, J, Xu, Z, Kresovich, JK, Lawrence, KG, et al. 2023. Air pollution and epigenetic aging among Black and White women in the US. Environ Int. 181: 108270.

78. Wolf, K, Cyrys, J, Harciníková, T, Gu, J, Kusch, T, Hampel, R, et al. 2017. Land use regression modeling of ultrafine particles, ozone, nitrogen oxides and markers of particulate matter pollution in Augsburg, Germany. Science of The Total Environment. 579: 1531-1540.

79. Illig, T, Gieger, C, Zhai, G, Römisch-Margl, W, Wang-Sattler, R, Prehn, C, et al. 2010. A genome-wide perspective of genetic variation in human metabolism. Nat Genet. 42(2): 137-41.

80. Mittelstrass, K, Ried, JS, Yu, Z, Krumsiek, J, Gieger, C, Prehn, C, et al. 2011. Discovery of sexual dimorphisms in metabolic and genetic biomarkers. PLoS Genet. 7(8): e1002215.

81. Pang, Z, Chong, J, Zhou, G, de Lima Morais, DA, Chang, L, Barrette, M, et al. 2021. MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. Nucleic Acids Research. 49(W1): W388-W396.

82. Yao, Y, Schneider, A, Wolf, K, Zhang, S, Wang-Sattler, R, Peters, A, et al. 2022. Longitudinal associations between metabolites and long-term exposure to ambient air pollution: Results from the KORA cohort study. Environment International. 170: 107632.

83. Yao, Y, Schneider, A, Wolf, K, Zhang, S, Wang-Sattler, R, Prehn, C, et al. 2023. Corrigendum to “Longitudinal associations between metabolites and immediate, short- and medium-term exposure to ambient air pollution: Results from the KORA cohort study” [Sci. Total Environ. 900 (2023) 165780]. Science of The Total Environment. 905: 167050.

84. Yao, Y, Wolf, K, Breitner, S, Zhang, S, Waldenberger, M, Winkelmann, J, et al. 2025. Long-Term Exposure to Ambient Air Pollution is Associated with Epigenetic Age Acceleration. Available at SSRN: http://dx.doi.org/10.2139/ssrn.5255305.

85. van der Veen, JN, Kennelly, JP, Wan, S, Vance, JE, Vance, DE, and Jacobs, RL. 2017. The critical role of phosphatidylcholine and phosphatidylethanolamine metabolism in health and disease. Biochimica et Biophysica Acta (BBA) - Biomembranes. 1859(9, Part B): 1558-1572.

86. Treede, I, Braun, A, Sparla, R, Kühnel, M, Giese, T, Turner, JR, et al. 2007. Anti-inflammatory effects of phosphatidylcholine. J Biol Chem. 282(37): 27155-27164.

87. Li, Y, Wang, DD, Chiuve, SE, Manson, JE, Willett, WC, Hu, FB, et al. 2015. Dietary Phosphatidylcholine Intake and Type 2 Diabetes in Men and Women. Diabetes Care. 38(2): e13-e14.

88. Tang, WHW, Wang, Z, Levison, BS, Koeth, RA, Britt, EB, Fu, X, et al. 2013. Intestinal Microbial Metabolism of Phosphatidylcholine and Cardiovascular Risk. New England Journal of Medicine. 368(17): 1575-1584.

89. Whiley, L, Sen, A, Heaton, J, Proitsi, P, García-Gómez, D, Leung, R, et al. 2014. Evidence of altered phosphatidylcholine metabolism in Alzheimer's disease. Neurobiol Aging. 35(2): 271-8.

90. Rinschen, MM, Ivanisevic, J, Giera, M, and Siuzdak, G. 2019. Identification of bioactive metabolites using activity metabolomics. Nat Rev Mol Cell Biol. 20(6): 353-367.

91. Gautheron, J and Jéru, I. 2021. The Multifaceted Role of Epoxide Hydrolases in Human Health and Disease. International Journal of Molecular Sciences. 22(1): 13.

92. Lee, CR, North, KE, Bray, MS, Fornage, M, Seubert, JM, Newman, JW, et al. 2006. Genetic variation in soluble epoxide hydrolase (EPHX2) and risk of

coronary heart disease: The Atherosclerosis Risk in Communities (ARIC) study. Human Molecular Genetics. 15(10): 1640-1649.

93. Fornage, M, Lee, CR, Doris, PA, Bray, MS, Heiss, G, Zeldin, DC, et al. 2005. The soluble epoxide hydrolase gene harbors sequence variation associated with susceptibility to and protection from incident ischemic stroke. Human Molecular Genetics. 14(19): 2829-2837.

94. Simpkins, AN, Rudic, RD, Roy, S, Tsai, HJ, Hammock, BD, and Imig, JD. 2010. Soluble epoxide hydrolase inhibition modulates vascular remodeling. American Journal of Physiology-Heart and Circulatory Physiology. 298(3): H795-H806.

95. Wang, L, Yang, J, Guo, L, Uyeminami, D, Dong, H, Hammock, BD, et al. 2012. Use of a Soluble Epoxide Hydrolase Inhibitor in Smoke-Induced Chronic Obstructive Pulmonary Disease. American Journal of Respiratory Cell and Molecular Biology. 46(5): 614-622.

96. Inceoglu, B, Jinks, SL, Schmelzer, KR, Waite, T, Kim, IH, and Hammock, BD. 2006. Inhibition of soluble epoxide hydrolase reduces LPS-induced thermal hyperalgesia and mechanical allodynia in a rat model of inflammatory pain. Life Sciences. 79(24): 2311-2319.

97. Johnson, AA and Stolzing, A. 2019. The role of lipid metabolism in aging, lifespan regulation, and age-related disease. Aging Cell. 18(6): e13048.

98. Chak, CM, Lacruz, ME, Adam, J, Brandmaier, S, Covic, M, Huang, J, et al. 2019. Ageing Investigation Using Two-Time-Point Metabolomics Data from KORA and CARLA Studies. Metabolites. 9(3): 44.

99. Bell, ML, Zanobetti, A, and Dominici, F. 2013. Evidence on Vulnerability and Susceptibility to Health Risks Associated With Short-Term Exposure to Particulate Matter: A Systematic Review and Meta-Analysis. American Journal of Epidemiology. 178(6): 865-876.

100. Sacks, JD, Stanek, LW, Luben, TJ, Johns, DO, Buckley, BJ, Brown, JS, et al. 2011. Particulate Matter–Induced Health Effects: Who Is Susceptible? Environmental Health Perspectives. 119(4): 446-454.

101. Sun, S, Cao, W, Qiu, H, Ran, J, Lin, H, Shen, C, et al. 2019. Benefits of physical activity not affected by air pollution: a prospective cohort study. International Journal of Epidemiology. 49(1): 142-152.

102. Tainio, M, Jovanovic Andersen, Z, Nieuwenhuijsen, MJ, Hu, L, de Nazelle, A, An, R, et al. 2021. Air pollution, physical activity and health: A mapping review of the evidence. Environment International. 147: 105954.

103. Lim, CC, Hayes, RB, Ahn, J, Shao, Y, Silverman, DT, Jones, RR, et al. 2019. Mediterranean Diet and the Association Between Air Pollution and Cardiovascular Disease Mortality Risk. Circulation. 139(15): 1766-1775.

104. Gleeson, M, Bishop, NC, Stensel, DJ, Lindley, MR, Mastana, SS, and Nimmo, MA. 2011. The anti-inflammatory effects of exercise: mechanisms and implications for the prevention and treatment of disease. Nature Reviews Immunology. 11(9): 607-615.

105. Péter, S, Holguin, F, Wood, LG, Clougherty, JE, Raederstorff, D, Antal, M, et al. 2015. Nutritional Solutions to Reduce Risks of Negative Health Impacts of Air Pollution. Nutrients. 7(12): 10398-10416.

106. European Union (EU), Directive 2008/50/EC of the European Parliament and of the Council of 21 May 2008 on ambient air quality and cleaner air for Europe. 2008.

107. European Union (EU). DIRECTIVE (EU) 2024/2881 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 23 October 2024 on ambient air quality and cleaner air for Europe. 2024.

108. Robinson, O, Chadeau Hyam, M, Karaman, I, Climaco Pinto, R, Ala-Korpela, M, Handakas, E, et al. 2020. Determinants of accelerated metabolomic and epigenetic aging in a UK cohort. Aging Cell. 19(6): e13149.

Publications

Paper I

Title:	Longitudinal associations between metabolites and long-term expo sure to ambient air pollution: Results from the KORA cohort study
Authors:	Yueli Yao, Alexandra Schneider, Kathrin Wolf, Siqi Zhang, Rui Wang-Sattler, Annette Peters, Susanne Breitner
Status:	Published
Journal:	*Environment International*
Year:	2022
Volume:	170
DOI:	https://doi.org/10.1007/s00125-023-05943-2
Supplements:	https://www.sciencedirect.com/science/article/pii/S0160412022005591#s0125
Impact factor:	11.8 (Journal Citation Reports®, year 2022)
Rank:	17/275 in Category Environmental Sciences Journals (Journal Citation Reports®, year 2022)

Environment International 170 (2022) 107632

Contents lists available at ScienceDirect

Environment International

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

Full length article

Longitudinal associations between metabolites and long-term exposure to ambient air pollution: Results from the KORA cohort study

Yueli Yao [a,b,*], Alexandra Schneider [b], Kathrin Wolf [b], Siqi Zhang [b], Rui Wang-Sattler [b,c], Annette Peters [a,b,c,d,1], Susanne Breitner [a,b,1]

[a] Institute for Medical Information Processing, Biometry and Epidemiology – IBE, Ludwig-Maximilians-Universität München, Munich, Germany
[b] Institute of Epidemiology, Helmholtz Zentrum München – German Research Center for Environmental Health, Neuherberg, Germany
[c] German Center for Diabetes Research, DZD, Munich-Neuherberg, Germany
[d] German Centre for Cardiovascular Research, DZHK, Partner Site Munich, Munich, Germany

ARTICLE INFO

Handling Editor: Shoji Nakayama

Keywords:
Long-term air pollution
Targeted metabolomics
Phosphatidylcholine
Susceptibility

ABSTRACT

Background: Long-term exposure to air pollution has been associated with cardiopulmonary diseases, while the underlying mechanisms remain unclear.
Objectives: To investigate changes in serum metabolites associated with long-term exposure to air pollution and explore the susceptibility characteristics.
Methods: We used data from the German population-based Cooperative Health Research in the Region of Augsburg (KORA) S4 survey (1999–2001) and two follow-up examinations (F4: 2006–08 and FF4: 2013–14). Mass-spectrometry-based targeted metabolomics was used to quantify metabolites among serum samples. Only participants with repeated metabolites measurements were included in the current analysis. Land-use regression (LUR) models were used to estimate annual average concentrations of ultrafine particles, particulate matter (PM) with an aerodynamic diameter less than 10 μm (PM_{10}), coarse particles (PM_{coarse}), fine particles, $PM_{2.5}$ absorbance (a proxy of elemental carbon related to traffic exhaust, $PM_{2.5abs}$), nitrogen oxides (NO_2, NO_x), and ozone at individuals' residences. We applied confounder-adjusted mixed-effects regression models to examine the associations between long-term exposure to air pollution and metabolites.
Results: Among 9,620 observations from 4,261 KORA participants, we included 5,772 (60.0%) observations from 2,583 (60.6%) participants in this analysis. Out of 108 metabolites that passed stringent quality control across three study points in time, we identified nine significant negative associations between phosphatidylcholines (PCs) and ambient pollutants at a Benjamini-Hochberg false discovery rate (FDR) corrected p-value < 0.05. The strongest association was seen for an increase of 0.27 μg/m^3 (interquartile range) in $PM_{2.5abs}$ and decreased phosphatidylcholine acyl-alkyl C36:3 (PC ae C36:3) concentrations [percent change in the geometric mean: −2.5% (95% confidence interval: −3.6%, −1.5%)].
Conclusions: Our study suggested that long-term exposure to air pollution is associated with metabolic alterations, particularly in PCs with unsaturated long-chain fatty acids. These findings might provide new insights into potential mechanisms for air pollution-related adverse outcomes.

1. Introduction

Epidemiological studies have shown associations between chronic exposure to ambient air pollution and pulmonary, cardio-metabolic, and neurological disease, and even mortality (Bae et al. 2021; Cao et al. 2020; Hales et al. 2021; Kasdagli et al. 2022; Liu et al. 2021; Mortamais et al. 2021; Park et al. 2021; Wolf et al. 2021). However, the underlying biological mechanisms are not yet fully elucidated. Hypothesized pathways linking air pollution exposure and health include the direct translocation of ambient particles with a smaller aerodynamic diameter (e.g., ultrafine particles) and gaseous air pollutants (e.g., nitrogen dioxide and ozone) from the lung into the blood leading to alternations of

* Corresponding author at: Institute of Epidemiology, Helmholtz Centre Munich – German Research Center for Environmental Health, Neuherberg D-85764, Germany.
E-mail address: yueli.yao@helmholtz-muenchen.de (Y. Yao).
[1] Authors share last authorship.

https://doi.org/10.1016/j.envint.2022.107632
Received 29 July 2022; Received in revised form 11 November 2022; Accepted 12 November 2022
Available online 13 November 2022

blood parameters (Nemmar et al. 2002). Another possible pathway is the induction of local inflammatory responses in the lung by larger inhaled ambient particles leading to autonomic cardiac, systemic inflammatory, and haemostatic activities (Brook et al. 2010).

The blood metabolome is a collection of biologically active chemicals in the human blood, derived from endogenous processes and exogenous exposure to food, medicines, and pollutants (Rappaport et al. 2014). Metabolomics has become a well-developed tool to investigate small molecular metabolites presented in the biological systems and corresponding cellular responses perturbed by endogenous or exogenous stimuli (Holmes et al. 2008). Recent epidemiological studies have provided evidence of adverse air pollution-induced effects on metabolomic biomarkers (Chen et al. 2019a; Gaskins et al. 2021; Hood et al. 2022; Li et al. 2017; Li et al. 2021; Liang et al. 2018; Ritz et al. 2022; van Veldhoven et al. 2019; Vlaanderen et al. 2017; Ward-Caviness et al. 2016). However, these studies mainly focused on short-term and intermediate exposures (day-to-day changes) to air pollution. Only a few studies have examined the metabolomics signatures in response to long-term air pollution exposures (e.g., annual averages) within cohort studies (Jeong et al. 2018; Nassan et al. 2021a; Nassan et al. 2021b; Walker et al. 2019). These studies were either limited to small sample sizes or focused on specific individuals, for example, older men or participants with adult-onset asthma or cardio-cerebrovascular diseases.

Given the limited evidence, especially within a general population cohort study, we aimed to determine the associations between long-term ambient air pollution and targeted metabolomics within the population-based Cooperative Health Research in the Region of Augsburg (KORA) cohort, conducted in the area of Augsburg, Germany. Additionally, we explored the role of potential individual characteristics in modifying the effects of air pollution effects, including body mass index (BMI), lifestyle (e.g., smoking status, alcohol consumption, physical activity, and dietary patterns), pre-existing diseases (e.g., hypertension and diabetes), and medication intakes (e.g., anti-hypertensive, anti-diabetic, and lipid-lowering medication). We hypothesized that long-term exposure to air pollution is associated with the perturbation of serum metabolite concentrations involved in some metabolic pathways related to adverse health effects from ambient air pollution and that individuals' characteristics can modify these health effects.

2. Methods

2.1. Study design and participants

In this longitudinal study, we used data from the KORA cohort. The fourth cross-sectional health survey of the KORA cohort (KORA S4) was conducted from October 1999 to April 2001. It involved 4,261 participants aged 25–74 years with German citizenship in the city of Augsburg, Germany, and two adjacent counties. Two follow-up examinations were carried out: within the first follow-up (KORA F4), 3,080 participants were examined between Oct 2006 and May 2008, whereas the second follow-up (KORA FF4) consisted of 2,279 participants with examinations between June 2013 and Sept 2014.

A computer-assisted personal interview, a self-administered questionnaire, and physical examinations were performed at each visit by trained investigators at the study centre. Physical activity was categorized based on the time spent on physical exercise into low (no or almost no physical exercise), medium (regular or irregular approx. one hour per week), and high (more than two hours per week) levels. Alcohol consumption was categorized into no (0 g/day), moderate (men 0.1–39.9 g/day and women 0.1–19.9 g/day), and high (men $\geq$ 40 g/day and women $\geq$ 20 g/day) consumption. Smoking status was categorized into current (regular or irregular smokers), former (ex-smokers), and never (never-smokers) smokers. A diet questionnaire with a qualitative food frequency list was performed to collect the dietary intake; a continuous dietary score and categorical dietary patterns were defined based on participants' answers. Briefly, the individuals' dietary intake was collected using a food-frequency questionnaire investigating 24 food groups. An index was built rating the frequency with which each food was consumed by assigning either 0, 1, or 2 points based on recommendations of the German Nutrition Society (DGE). Higher scores reflect better compliance with DGE recommendations. A sum dietary score ranging from 0 to 27 was calculated according to DGE guidelines and subsequently grouped into three categories: adverse ($\leq$13 points), ordinary (14 $\sim$ 15 points), and favourable ($\geq$16 points) dietary patterns. This approach was established in earlier KORA studies and was validated against a weighed 7-day dietary protocol (Rabel et al. 2018; Winkler and Döring 1998).

Only participants who attended at least two visits across the entire study period were included in this longitudinal analysis. Additionally, we excluded participants with missing data on covariates used in our main analysis (**Fig. S1**). Written informed content was obtained from all participants. The KORA study was approved by the ethics committee of the Bavarian Chamber of Physicians (Munich, Germany).

2.2. Biomarker measurements

Blood samples were drawn into serum gel tubes between 8:00 am and 10:30 am after at least 8 h of overnight fasting. The blood samples were kept at 4 °C up to six hours after blood withdrawal for further procedure. Serum was collected and filled into synthetic straws, and stored in liquid nitrogen (−80 °C) until the further analyses were conducted.

2.3. High-sensitivity C-reactive protein (hs-CRP)

The high-sensitivity C-reactive protein (hs-CRP) assay was performed shortly after the blood withdrawal for each study wave (KORA S4 (September-December 2001), KORA F4 (July-October 2008), and for KORA FF4 (December 2015-March 2016)). hs-CRP was measured in serum by a BN nephelometer (Siemens Healthcare Diagnostics Product GmbH, Marburg, Germany) in the collaborating Biomarker Laboratory at the University of Ulm, Germany.

2.4. Targeted metabolomics

The metabolite profiling in serum samples was done with the AbsoluteIDQ^{TM} p180 kit (BIOCRATES Life Sciences AG, Innsbruck, Austria) for KORA S4 (March–April 2011) and FF4 (February–October 2019), allowing for the simultaneous quantification of 188 metabolites. KORA F4 samples were measured with the AbsoluteIDQ^{TM} p150 kit to detect 163 metabolites in August 2008–March 2009. The assay procedures have been described previously in detail (Römisch-Margl et al. 2012).

Identical quality control (QC) procedures were used in each of the three study points in time. Each metabolite should meet the following three criteria: (1) The average value of the coefficient of variance (CV) in the five/six reference samples or three quality control samples should be less than 25%; (2) 50% of all measured sample concentrations for the metabolite should be above the limit of detection (LOD), which was defined as three times the median of zero samples; (3) The rate of missing value of metabolite should be less than 5%. The non-detectable values of each metabolite were randomly imputed by values ranging from 75% to 125% of half of the lowest measured value of the corresponding metabolite in each plate. In order to minimize the plate effects in each visit, plate normalization factors were calculated by dividing the mean of reference sample values (QC samples in KORA F4) in each plate by the mean of all reference sample values in all plates, and then used to normalize each metabolite (Han et al. 2022; Huang et al. 2020).

Additionally, to control for the effects of the different kits between KORA F4 and KORA S4/FF4, up to eight participants' samples were randomly selected from each of the 36 kit plates in KORA F4 and re-measured using the same AbsoluteIDQt p180 kit used in KORA S4/FF4

in September–October 2019 (Han et al. 2022). The difference in each metabolite between the corresponding participants in KORA F4 and re-measured KORA F4, and a further mean difference of each metabolite were calculated. The kit normalization factor was calculated by dividing the mean of each metabolite in KORA F4 by the mean of each metabolite in KORA F4 minus the mean difference between KORA F4 and re-measured KORA F4, and used to correct KORA F4 metabolite data. Extreme outliers of each metabolite were defined as a value beyond the range of mean ± 5 × standard deviations and imputed by the K-nearest neighbors algorithm (KNN).

In total, 135 metabolites in KORA S4, 114 in KORA F4, and 145 in KORA FF4 passed the quality control. Out of these, 108 metabolites were overlapped among KORA S4, F4, and FF4 and were used in the subsequent analysis. Metabolites covered the following compound classes: 12 amino acids, 12 acylcarnitines, 72 glycerophospholipids (including 32 phosphatidylcholines with acyl-acyl (diacyl) side chains, 33 phosphatidylcholines with acyl-alkyl side chains, and seven lysophosphatidylcholines), 11 sphingomyelins (SM) and a sum of hexoses (including glucose). The complete list of metabolites is presented in the supplementary material (**Table S1**).

2.5. Exposure assessment

Residential annual mean exposure to air pollution including ultrafine particles (particulate matter (PM) ≤ 100 nm in aerodynamic diameter, represented by particle number concentration (PNC)), PM with an aerodynamic diameter less than 10 μm (PM_{10}), between 2.5 and 10 μm (PM_{coarse}), and less than 2.5 μm ($PM_{2.5}$), $PM_{2.5}$ absorbance (a proxy of elemental carbon related to traffic exhaust, $PM_{2.5abs}$), nitrogen oxides (NO_2, NO_x), and ozone (O_3) was estimated using land-use regression (LUR) models. The performance of LUR models was evaluated by leave-one-out cross-validation (LOOCV) (Wolf et al. 2017). Briefly, three bi-weekly measurements at 20 locations within the KORA study area were carried out between March 2014 and April 2015 to cover the warm, cold, and intermediate seasons. Simultaneously, measurements were obtained at a reference site throughout the whole period to adjust for temporal variation. Annual average air pollutant concentrations were then calculated at those sites. The LUR model was built by regressing the measured annual average concentrations in 2014–15 against geographic information system-based spatial predictors including local land use (e.g. residential land, industrial, commercial and transport units, urban green, and water bodies), building density, population density, household density, topography, coordinates, and traffic variables (e.g. total traffic load of all (major) roads in a buffer, traffic intensity on nearest (major) road, and heavy-duty traffic intensity on nearest (major) road) (Wolf et al. 2017). Participants' home addresses were applied to the fitted models to determine residential exposure levels. The adjusted model-explained variance (R^2) of the LUR models ranged from 68% (PM_{coarse}) to 94% (NO_2), and the adjusted LOOCV R^2 was between 55% (PM_{coarse}) and 89% (NO_2), which indicated a good model fit. The process has been described in detail elsewhere (Wolf et al. 2017). For participants who moved during the study period, the updated residential addresses were used for exposure assignment; otherwise, the same exposure levels were assigned across different visits.

3. Statistical analyses

3.1. Statistical methods

Basic descriptive analyses were performed for participant characteristics, air pollutants, and meteorological parameters. Kruskal-Wallis test (one-way ANOVA) and Pearson's Chi-squared test were applied for continuous and categorical variables, respectively. Spearman's rank correlation coefficient was used to calculate correlations between air pollutants.

We applied linear mixed-effects models with random participant-specific intercepts to examine the associations between repeatedly measured metabolite levels and air pollutants. In addition, linear mixed-effects models were also performed between hs-CRP and air pollutants to investigate the systemic inflammatory response. All outcomes (metabolites and hs-CRP) were natural-log transformed to increase the conformity to normal distributions of residuals. Covariates included in the models were selected a priori based on previous studies and the Bayesian Information Criterion (BIC) (Holmes et al. 2008; Lacruz et al. 2016; Nassan et al. 2021a; Sun et al. 2020b; Ward-Caviness et al. 2016). Minimum models adjusted for age, sex, body-mass index (BMI), an indicator of each visit (KORA S4, KORA F4, or KORA FF4), and season of blood withdrawal (winter: December–February, spring: March–May, summer: June–August, and autumn: September–November). Main models additionally included smoking status (never/former/current), alcohol consumption (g/day), physical activity (low/medium/high), educational attainment (primary school/high school/college), fasting status (overnight fasting of 8 h or not) and diet score (continuous). Extended models further added hypertension, diabetes, medication intake (anti-hypertensive, anti-diabetic, and lipid-lowering medication), high-density lipoproteins (HDL), and total cholesterol. Effect estimates are presented as percent changes in the geometric mean (together with 95% confidence intervals [95% CI]) of the repeatedly assessed outcomes per interquartile range (IQR) increase in air pollutant concentrations.

Single – Pollutant models

$$\log(Y_{ij}) = \beta_0 + \mu_i + \beta_1 \times AP_{ij} + \beta_{2-n} \times Covariates_{ij} + e_{ij}$$

In the formula, Y_{ij} is the metabolite concentration of participant i at visit j. β_0 denotes the fixed intercept, and μ_i represents the random intercept for subject i. β_1 is the estimate of each air pollutant and AP_{ij} indicates the annual averages of the air pollutants (PM_{10}, PM_{coarse}, $PM_{2.5}$, $PM_{2.5abs}$, PNC, NO_2, NO_x, and O_3) for participant i at visit j. $\beta_{2\text{-}n}$ is estimate for each covariate, and $Covariates_{ij}$ represents the measurement of covariates for participant i at visit j. e_{ij} is the residual normal error.

Effect modification was investigated by including an interaction term between each air pollutant and the potential effect modifier assessed at each visit. The examined modifiers included age (<65 years *vs* ≥ 65 years; the age 65 years is the current official retirement age in Germany)), sex (male *vs* female), obesity (BMI < 30 kg/m^2 *vs* ≥ 30 kg/m^2), smoking status (current *vs* never/former smoker), alcohol consumption (low *vs* medium *vs* high), education (low *vs* high (high school/college)), physical activity (low *vs* medium *vs* high), dietary pattern (adverse *vs* ordinary *vs* favourable), hypertension (no *vs* yes), diabetes (no *vs* yes), and medication intakes (no *vs* yes). The effect modification analyses were only conducted for those metabolites significantly associated with air pollutants.

We performed several sensitivity analyses in this study: 1) We included all participants with data on air pollution, phenotypes, and metabolites in the analysis. 2) We restricted our analyses to participants who did fasting eight hours before the blood withdrawal throughout the entire study period. 3) Additionally, we restricted our main analysis to participants who did not move within the study period. 4) To control for selection bias introduced by selecting participants with more than one measurement, we estimated weights for those included using the inverse probability weighting (IPW) method (Weuve et al. 2012). Briefly, the probability of being included in our main analysis among all study participants in KORA S4 was calculated using logistic regression. We used individual characteristics of our main analysis as possible predictors. Then, we applied the inverse of the predicted probability determined from the logistic regression as the weight in our main model. 5) Given the temporal variation of each air pollutant exposure, we used back-extrapolated annual average air pollutant concentrations from the respective years of KORAS4, F4 and FF4 instead of using annual average air pollutant concentration estimated by the LUR models in 2014–2015 **(Text S1)**. Briefly, the absolute differences between the LUR model and the air pollutants data from monitors in the period of each visit were

calculated. They were then used to correct each visit's air pollutant concentrations, respectively. 6) To examine the influence of air pollution-associated systemic inflammation, we further included high-sensitivity C-reactive protein (hs-CRP) in our main models. 7) We performed two-pollutant models by including two air pollutants simultaneously if their Spearman correlation was smaller than 0.7. 8) We also performed a mixed-effects quantile regression to assess the association between air pollution exposure and metabolites at deciles of the metabolites. 9) To investigate the co-effects between long-term and short-term air pollution exposure, we simultaneously included short-term exposures (at the day of blood withdrawal, one day, two days, three days, four days, as well as two-day, five-day and two-week moving averages before the blood withdrawal) to each air pollutant in the corresponding long-term exposure model. The short-term exposure included $PM_{2.5}$, PM_{10}, PM_{coarse}, NO_2, NO_x and O_3, was measured consecutively by local monitors and the daily average exposure concentration of each air pollutant was assigned to each participant based on the date of blood withdrawal in each visit (**Text S1**). 10) To assess the effect of the storage time (Haid et al. 2018), we performed an additional sensitivity analysis, including the storage year in our main models. Briefly, we calculated the storage time between the collection date of the blood sample for each participant and the detection time (middle date in the whole measurement period). Then, we included this storage year in our main models. We assumed a non-linear relationship between the change of metabolites concentrations and the storage time, so we used a spline for the storage years to account for non-linearity in these relationships. 11) In the main models, we assessed the exposure–response relationships between all metabolites and air pollutants for deviations from linearity using penalized splines with the degree of freedom selected by generalized cross-validation, and restricted our analyses to the linear section of the relationship.

All statistical analyses were done with R (version 3.6.2), and the *p*-value cut-off was set as 5.8×10^{-5} to account for multiple testing introduced by assessing eight air pollutants and 108 metabolites in this study (0.05/(108*8)). We also report all associations with *p*-values<0.05 after Benjamini-Hochberg false discovery rate (FDR) correction since the Bonferroni method for adjusting *p*-values is more conservative.

3.2. Pathway analysis for metabolites

For metabolites showing significant associations with air pollutants after correcting for multiple testing, we performed pathway analysis using the "Pathway Analysis" module in MetaboAnalyst 5.0, a web-based software for metabolomics data analysis (Pang et al. 2021). This module supports pathway analysis by integrating two parts, enrichment analysis and topology analysis, based on the Kyoto Encyclopedia of Genes and Genomes (KEGG) database, which is a collection of manually drawn pathway maps representing the knowledge of molecular interaction, reaction, and relation networks. In the enrichment analysis, the *p*-value is calculated by the one-tailed Fisher's exact test, which represents the probability of observing at least *k* metabolites in a pathway, if there is no association with air pollution:

$$p(X \geq k) = 1 - \sum_{i=0}^{k-1} \frac{\binom{M}{i}\binom{N-M}{n-i}}{\binom{N}{n}}$$

where *N* represents the number of the metabolites detected by the platform, *M* indicates the metabolites in the pathway of interest (i^{th}), *n* is the metabolites significantly associated with air pollution, and *k* means the number of metabolites overlapped between *M* and *n* mapping to the i^{th} pathway (Wieder et al. 2021). The pathway topology analysis uses two well-established node centrality measures to estimate node importance. Furthermore, to take into account the comparison among different pathways, the node importance values calculated from centrality measures are further normalized by the sum of the importance of the pathway. Therefore, the total/maximum importance of each pathway is one. The importance measure of each metabolite node reflects the percentage with regard to the total pathway importance, and the pathway impact value is the cumulative percentage from the matched metabolite nodes. Pathways with a *p*-value ≤ 0.1, or with an impact value > 0.5 while *p*-value ≤ 0.3 were considered the most relevant pathways.

4. Results

4.1. Characteristics of study participants

Participant characteristics are summarized in Table 1. Only participants attending at least two visits during the entire study period with no missing information in the main confounders were included in our main analyses. Therefore, among 9,620 observations from 4,261 study participants in the KORA cohort, we included 5,772 (60.0%) observations from 2,583 (60.6%) participants in this analysis. Specifically, 1,977 (76.5%) out of the 2,583 participants attended two examinations, and 606 (23.5%) attended all three examinations (Table 1).

Due to the fasting status restriction in KORA S4, only 1,601 from overall 4,261 participants had data on metabolite levels, mainly elderly individuals. Therefore, on average, KORA S4 participants were older than those of KORA F4 and KORA FF4 (Table 1). Meanwhile, the average educational attainment, the percentages of 8 h overnight fasting before blood withdrawal, current smoker, unhealthy dietary pattern, and medium and high levels of physical activity of KORA S4 were lower (*p*-value < 0.01). In contrast, the mean BMI, alcohol consumption, cholesterol, HDL, and hs-CRP and the percentage of hypertension were higher in KORA S4 (*p*-value < 0.01).

4.2. Characteristics of air pollutants

Annual average concentrations of $PM_{2.5}$, PM_{10} and NO_2 at participant's residences were below the EU air quality standard values of 25 $\mu g/m^3$ for $PM_{2.5}$, and 40 $\mu g/m^3$ for PM_{10} and NO_2, respectively. While they were all higher than the WHO air quality guideline values of 5 $\mu g/m^3$, 10 $\mu g/m^3$ and 10 $\mu g/m^3$ for $PM_{2.5}$, PM_{10} and NO_2, respectively. The maximum annual O_3 concentration (45.9 $\mu g/m^3$) was also below the WHO air quality guideline values calculated from peak season (60 $\mu g/m^3$) (Table 2). All pollutants showed a strong positive relationship, except for O_3, which showed weak or negative correlations with other air pollutants (Table 2).

4.3. Association between metabolites and long-term air pollution

In our main models, several metabolites from the phosphatidylcholines group showed significant negative associations with PM_{coarse}, $PM_{2.5abs}$ and NO_2, respectively (Fig. 1). Specifically, PC ae C34:2 and PC ae C36:3 were negatively associated with PM_{coarse} and $PM_{2.5abs}$ (at a *p*-value $< 5.8 \times 10^{-5}$). Additionally, at an FDR-corrected *p*-value < 0.05, we observed decreases in PC ae C34:2 and PC ae C36:3 in association with NO_2. Moreover, PC ae C36:4 showed negative associations with $PM_{2.5abs}$ and PM_{coarse}, and PC ae C34:3 with $PM_{2.5abs}$, respectively. These results were robust in our minimum and extended models (Fig. 2). In addition, we observed positive associations between hs-CRP and PM_{coarse}, PM_{10}, PNC and NO_x (uncorrected *p*-value < 0.05) (**Fig. S2**). While the four identified metabolites showed moderate to high correlations with each other, hs-CRP was not associated with them at all (**Fig. S3**).

4.4. Pathway analysis

In the pathway analysis, we uploaded the four metabolites significantly associated with at least one of the long-term exposures to $PM_{2.5abs}$, PM_{coarse}, or NO_2. We identified four metabolic pathways,

Table 1
Descriptive statistics of participant characteristics for KORA S4, F4 and FF4 (N = 5,772).

Variable	S4 (N = 1,129) Mean ± SD / N (%)	F4 (N = 2,556) Mean ± SD / N (%)	FF4 (N = 2,087) Mean ± SD / N (%)	*p*–value
Age (years)	63.3 ± 5.4	57.5 ± 13.3	60.7 ± 12.3	< 0.001
Sex (male)	570 (50.6)	1,240 (48.5)	1,012 (48.5)	0.46
Education				< 0.001
Primary school	753 (66.7)	1,357 (53.1)	1034 (49.5)	
High school	221 (19.6)	621 (24.3)	530 (25.4)	
College	155 (13.7)	578 (22.6)	523 (25.1)	
BMI (kg/m^2)	28.4 ± 4.2	27.7 ± 4.7	27.8 ± 4.9	< 0.001
Alcohol consumption (g/day)	16.2 ± 20.9	14.4 ± 19.5	14.9 ± 20.1	0.025
Dietary score	16.2 ± 3.6	15.3 ± 3.6	15.1 ± 3.6	< 0.001
Dietary patterns				< 0.001
Adverse	271 (24.0)	817 (31.9)	715 (34.3)	
Ordinary	212 (18.8)	541 (21.2)	451 (21.6)	
Favorable	646 (57.2)	1,198 (46.9)	921 (44.1)	
Fasting (8 h) (% yes)	1,016 (90.0)	2,543 (99.5)	2,074 (99.4)	< 0.001
Smoking status				0.002
Current smoker	137 (12.1)	384 (15.0)	307 (14.7)	
Former smoker	437 (38.7)	1,066 (41.7)	902 (43.2)	
Never smoker	555 (49.2)	1,106 (43.3)	878 (42.1)	
Physical activity				< 0.001
Low	444 (39.3)	818 (32.0)	589 (28.2)	
Medium	479 (42.4)	1,115 (43.6)	952 (45.6)	
High	206 (18.3)	623 (24.4)	546 (26.2)	
Hypertension (% yes)	609 (53.9)	1,016 (39.8)	825 (39.5)	< 0.001
Diabetes (% yes)	92 (8.2)	224 (8.8)	215 (10.3)	0.11
Medication intake (% yes)				
Anti–hypertension medication	397 (35.2)	861 (33.7)	782 (37.5)	0.03
Anti–diabetes medication	53 (4.7)	154 (6.0)	174 (8.3)	< 0.001
Lipid lowering medication	128 (11.3)	351 (13.7)	342 (16.4)	< 0.001
Cholesterol (mg/dL)*	243.6 ± 40.8	216.1 ± 38.8	216.7 ± 39.5	< 0.001
HDL (mg/dL)*	58.1 ± 16.5	56.1 ± 14.4	65.9 ± 18.8	< 0.001
hs-CRP (mg/L)*	3.1 ± 4.9	2.4 ± 4.8	2.5 ± 4.6	< 0.001

KORA = Cooperative Health Research in the Region of Augsburg; S4 = fourth cross-sectional health survey of the KORA cohort; F4 = first follow-up examination of KORA S4; FF4 = second follow-up examination of KORA S4; BMI = body mass index; HDL = high density lipoprotein; hs-CRP = high sensitivity C-reactive protein; S4 participants were selected based on whether they did fasting or not. Dietary patterns was classified by the dietary score basing on the assessment of individual's dietary intake (questionnaire): Adverse = ≤13 points, Ordinary = 14 ~ 15 points, Favourable = ≥16 points. Physical activity was defined according to the exercise time per week: Low = almost or no sporting activity, Medium = regular/ irregular approx. 1 h per week, High = regularly>2 h in the week. *Cholesterol was missing for one (0.09%) participant in KORA S4, and one (0.05%) in KORA FF4; HDL was missing for one (0.09%) participant in KORA S4, one (0.04%) in KORA F4, and one (0.05%) in KORA FF4; hs-CRP was missing for 12 (1.06%) participants in KORA S4, five (0.22%) in KORA F4, and two (0.10%) in KORA FF4. 1,977 participants attended two examinations, and 606 attended three examinations. *p*-value was based on the Kruskal-Wallis test for continuous variables, and Pearson's Chi-squared test for categorical variables.

including the arachidonic acid (p-value = 0.003, impact value = 0), linoleic acid (p-value = 0.008, impact value = 0), alpha-linolenic acid (p-value = 0.02, impact value = 0), and glycerophospholipid (p-value = 0.02, impact value = 0.1) metabolisms that were related to long-term exposure to $PM_{2.5abs}$, PM_{coarse}, and NO_2 exposure, where p-value was from enrichment analysis and pathway impact value was from the topology analysis. However, they were insignificant after using the FDR method to correct the raw p-value (Fig. 3 and **Table S2**).

4.5. Effect modification

Effect modification analyses were conducted for the four metabolites significantly associated with long-term exposure to $PM_{2.5abs}$, PM_{coarse} and NO_2. Results are presented in Fig. 4 showing that the associations between PC ae C34:3 and $PM_{2.5abs}$, PM_{coarse} and NO_2 were significantly modified by physical activity (Bonferroni-corrected $p < 0.004$). Participants with low physical activity showed the strongest effects. A similar pattern was seen for PC ae 34:2, PC ae 36:3 and $PM_{2.5abs}$ and NO_2 (uncorrected $p < 0.05$). Moreover, results indicated a consistent modification of the air pollutant effects on PC ae 34:2 by education. Participants with a lower education showed stronger effects compared to those with a higher education. Results also suggested a consistent modification of the air pollutant effects on PC ae 36:4 by obesity - obese individuals showed stronger associations between metabolites and air pollutants. We did not find consistent differences between smokers and participants who never smoked and participants with medium or high alcohol consumption versus those without alcohol intake. Additionally, results suggested effect modification by disease status (hypertension and diabetes) and medication intake, while the differences were not statistically significant (**Fig. S4**, **Fig. S5**). This might be due to the large difference in the sample sizes of the different groups since much fewer participants had diabetes or intake of anti-hypertension, anti-diabetes, or lowering-lipid medicines. We also did not find significant differences between males and females in most metabolites except that a few PCs decreased more in females than males when exposed to O_3 (**Fig. S6**).

4.6. Sensitivity analyses

The associations between air pollution and the four metabolites were generally robust in different sensitivity analyses. Results remained stable when restricting the participants to fasting individuals or those who did not move their residences during the whole study period (Fig. 5). Additionally, including all participants, using predicted inverse probabilities, or using back-extrapolated air pollutant exposures to adjust for measurement error did not change the results. The results were still robust after further including hs-CPR in the main and extended models (**Fig. S7**).

The associations between metabolites and particulate air pollutants (PM_{10}, PM_{coarse}, $PM_{2.5abs}$, and PNC) were robust after additionally adjusting for $PM_{2.5}$ except for PNC where associations were attenuated (**Fig. S8**). The associations between metabolites and particle metrics ($PM_{2.5}$, PM_{10}, PM_{coarse}, $PM_{2.5abs}$, and PNC) and NO_2 were also stable after adjusting for O_3 (**Fig. S9**). After additionally adjusted by the storage year of blood samples into the main model, the effect estimates keep stable (**Fig. S10**). The additional adjustment of short-term air pollution exposure slightly strengthened the effect estimates of long-

Table 2
Descriptive statistics and Spearman correlation coefficients of air pollution concentrations in long-term analysis (N = 2,583).

Pollutant	Mean ± SD	Range	IQR	Spearman correlation coefficients							
				$PM_{2.5}$	PM_{10}	$PM_{2.5abs}$	PM_{Coarse}	PNC	O_3	NO_2	NO_x
$PM_{2.5}$ ($\mu g/m^3$)	11.8 ± 1.0	8.2–14.3	1.4	1							
PM_{10} ($\mu g/m^3$)	16.6 ± 1.5	12.3–22.3	2.1	0.52	1						
$PM_{2.5abs}$ ($10^{-5}/m$)	1.2 ± 0.18	0.8–1.8	0.3	0.61	0.78	1					
PM_{coarse} ($\mu g/m^3$)	4.9 ± 1.0	2.6–8.7	1.4	0.57	0.78	0.81	1				
PNC ($10^3/cm^3$)	7.3 ± 1.8	3.2–15.0	2.0	0.65	0.80	0.78	0.76	1			
O_3 ($\mu g/m^3$)	39.1 ± 2.4	31.3–45.9	3.4	–0.18	0.05	–0.10	0.14	–0.03	1		
NO_2 ($\mu g/m^3$)	14.1 ± 4.4	6.9–27.5	6.9	0.72	0.72	0.86	0.83	0.78	–0.16	1	
NO_x ($\mu g/m^3$)	21.8 ± 7.4	4.0–50.5	8.8	0.75	0.73	0.72	0.75	0.90	–0.06	0.83	1

*Exposure levels were estimated at participants' residences in KORA S4. In total, 2153 participants didn't move since S4, and 430 participants moved between S4 and F4, or F4 to FF4. For participants who changed residence among S4, F4 and FF4, the updated residential addresses were used for exposure assignment to the respective study. Otherwise, the same exposure levels from KORA S4 were assigned across different visits. $PM_{2.5}$ = particulate matter with an aerodynamic diameter less than or equal to 2.5 μm; PM_{coarse} = particulate matter with an aerodynamic diameter of 2.5–10 μm; PM_{10} = particulate matter with an aerodynamic diameter less than or equal to 10 μm; $PM_{2.5abs}$ = $PM_{2.5}$ absorbance; PNC = particle number concentration; NO_2 = nitrogen dioxide; NO_x = nitrogen oxide; O_3 = ozone.

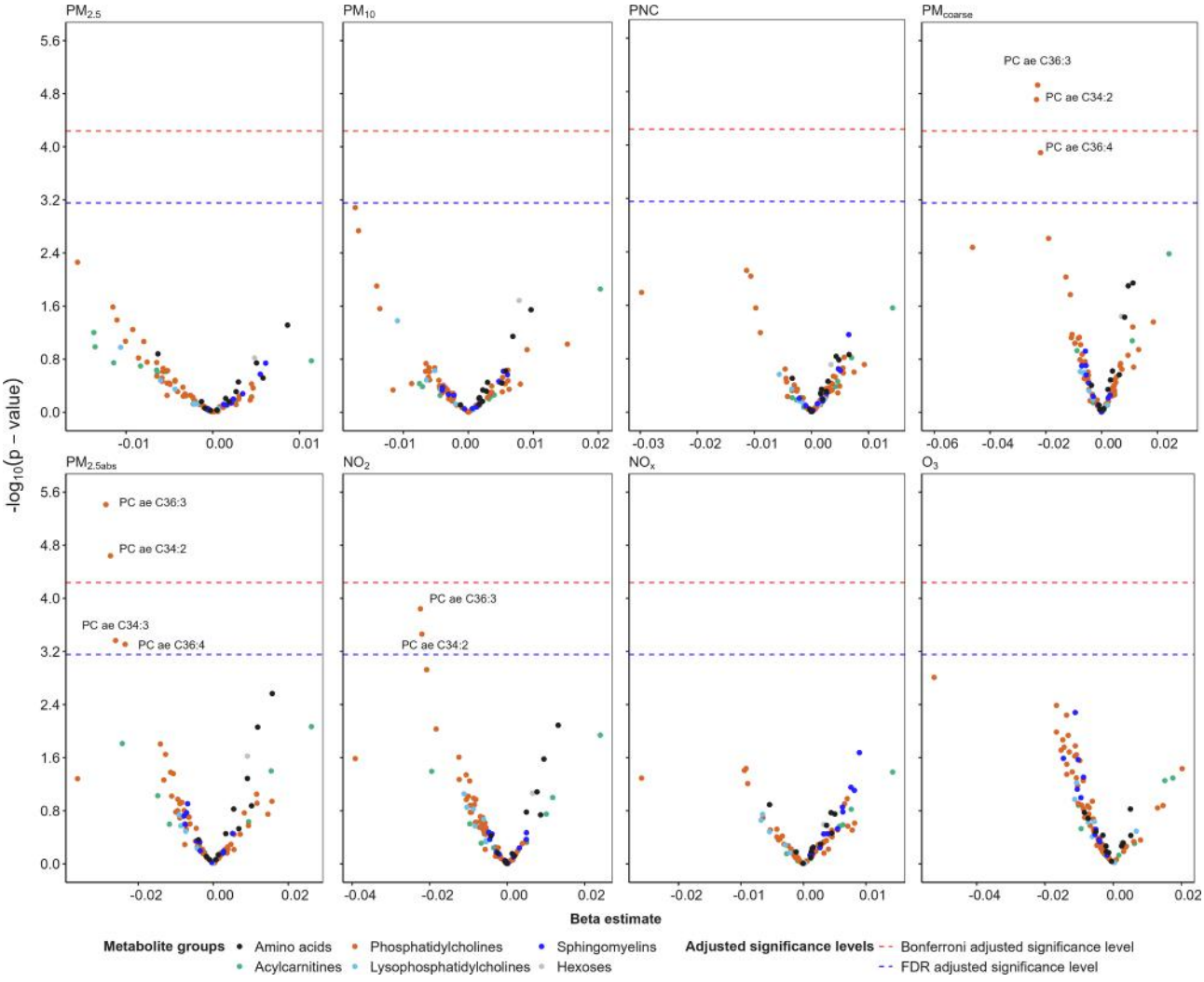


Fig. 1. Volcano plots presenting the associations between long-term air pollutant exposure and metabolites. The results were derived from the main models adjusted for age, sex, body-mass index (BMI), an indicator for each visit(KORA S4, KORA F4, or KORA FF4), season of blood withdrawal, smoking status, alcohol consumption, physical activity, educational attainment, fasting status, and dietary score. The Y axis shows the negative logarithm of the *p*-value (logarithmic base of 10). The X axis indicates the association between air pollutants and metabolites. The red and blue dashed lines represent adjusted statistical significance levels according to Bonferroni and FDR methods, respectively. The points with six different colors represent six metabolite groups involved in this study including amino acids (black), acylcarnitines (green), phosphatidylcholines (orange), lysophosphatidylcholines (light blue), sphingomyelins (blue), and hexoses (grey). PC ae: acyl-alkyl phosphatidylcholine. $PM_{2.5}$ = particulate matter with an aerodynamic diameter less than or equal to 2.5 μm; PM_{coarse} = particulate matter with an aerodynamic diameter of 2.5–10 μm; PM_{10} = particulate matter with an aerodynamic diameter less than or equal to 10 μm; $PM_{2.5abs}$ = $PM_{2.5}$ absorbance; PNC = particle number concentration; NO_2 = nitrogen dioxide; NO_x = nitrogen oxide; O_3 = ozone.

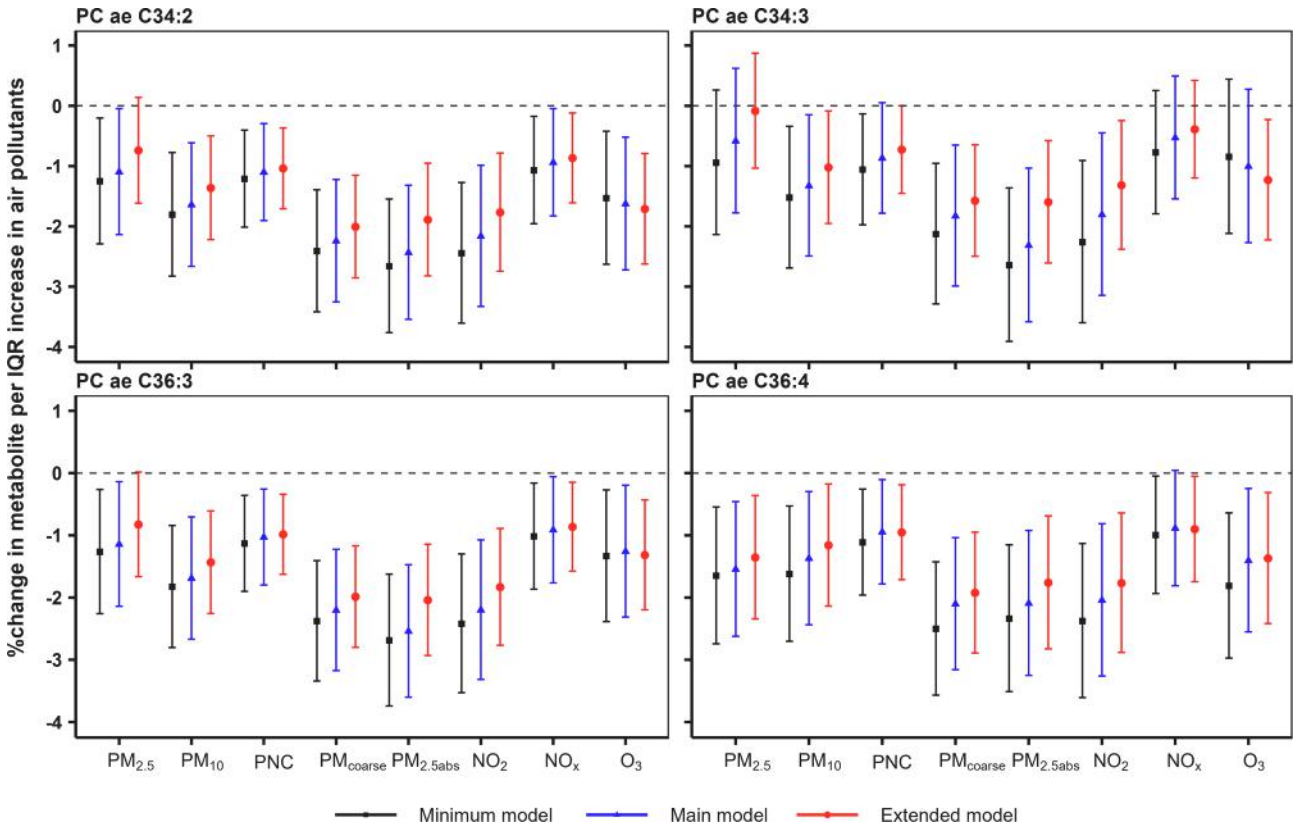


Fig. 2. Comparisons of percent changes (95% CIs) of metabolites per IQR increase in air pollutant concentrations between the results from minimum, main and extended models. Minimum model: minimum models were adjusted for age, sex, BMI, season, an indicator of each visit (KORA S4, KORA F4, or KORA FF4); Main model: further adjusted for educational attainment, smoking status, fasting status, alcohol consumption, physical activity, and dietary score; Extended model: additionally included hypertension, diabetes, and medication intake (anti-hypertension, anti-diabetes, and lipid lowering medications), HDL, and total cholesterol into the main models. An IQR increase was 1.40 μg/m^3 for $PM_{2.5}$, 2.06 μg/m^3 for PM_{10}, 1.36 μg/m^3 for PM_{coarse}, 1.95 × 10^3/cm^3 for PNC, 0.27 × 10^{-5}/m for $PM_{2.5abs}$, 6.86 μg/m^3 for NO_2, 8.69 μg/m^3 for NO_x, and 3.45 μg/m^3 for O_3. PC ae: acyl-alkyl phosphatidylcholine. $PM_{2.5}$ = particulate matter with an aerodynamic diameter less than or equal to 2.5 μm; PM_{coarse} = particulate matter with an aerodynamic diameter of 2.5–10 μm; PM_{10} = particulate matter with an aerodynamic diameter less than or equal to 10 μm; $PM_{2.5abs}$ = $PM_{2.5}$ absorbance; PNC = particle number concentration; NO_2 = nitrogen dioxide; NO_x = nitrogen oxide; O_3 = ozone.

term exposure but showed consistent estimates across the different exposure windows (Fig. S11).

Mixed-effects quantile regression showed similar associations for $PM_{2.5abs}$ exposure across deciles (Fig. S12). In contrast, there were stronger associations between PM_{coarse} exposure and the four metabolites from the 10th percentile up to the fifth decile (Fig. S13). Similarly, stronger associations were seen between the four metabolites and NO_2 exposure from the second to the sixth decile (Fig. S14).

We also checked the exposure–response relationships of all metabolites and $PM_{2.5abs}$, PM_{coarse} and NO_2 exposure. There was no deviation from linearity for NO_2 with the four metabolites (Fig. S15), while slight deviations were observed for $PM_{2.5abs}$ and PM_{coarse} (Fig. S16, Fig. S17). We excluded the extreme values for $PM_{2.5abs}$ (>99% of total $PM_{2.5bs}$) and PM_{coarse} (<5% of total PM_{coarse} and > 95% of total PM_{coarse}) from the dataset to ensure a linear exposure–response relationship for $PM_{2.5abs}$, the results kept robust with our main analysis results (Fig. S18-S21).

5. Discussion

This longitudinal study identified nine associations between long-term exposure to air pollution and targeted serum metabolites, mainly from the phosphatidylcholine subgroup. In particular, we observed that participants exposed to higher $PM_{2.5abs}$, PM_{coarse} and NO_2 had lower levels of PC ae C34:2 and PC ae C36:3. In addition, PC ae C36:4 showed a negative association with $PM_{2.5abs}$ and PM_{coarse}, and PC ae C34:3 was negatively associated with $PM_{2.5abs}$. In the subsequent pathway analysis, they were identified as related to glycerophospholipid, linoleic acid and alpha-linolenic acid metabolism. Moreover, we found effect modifications for several individual characteristics: participants with older age, obesity, lower educational attainment, low physical activity levels, or adverse dietary patterns showed stronger associations than their counterparts. In addition, we could confirm positive associations between several air pollutants (PM_{coarse}, PM_{10}, PNC, and NO_x) and hs-CRP as previously reported cross-sectionally for FF4 (Pilz et al. 2018), where we saw positive but non-significant associations with PNC, PM_{10}, PM_{coarse}, $PM_{2.5abs}$, NO_2, and NO_x.

Metabolites are the intermediates or end products of metabolism, and could affect cellular physiology through modulation of other "omics" levels and represent changes induced by exposures (Rinschen et al. 2019). Alterations in the lipid metabolism due to the unbalance of anti- and pro-inflammatory biomarkers and oxidative stress levels could be one of the underlying mechanisms linking air pollution exposure to adverse health effects. Only a few studies explored the associations between long-term exposure to air pollution and metabolites in a cohort setting. A cross-sectional study based on the TwinsUK cohort reported eight inflammation and oxidative stress-related metabolites out of 280 untargeted metabolomics profiling. For example, α-tocopherol, glycine, and benzoate were associated with long-term $PM_{2.5}$. Moreover, CRP was negatively associated with seven of these eight metabolites (Menni et al. 2015). A study including cohorts from Italy and Switzerland reported that long-term exposure to air pollution on adult asthma and cardiovascular disease was related to unsaturated fatty acids e.g., linolenic acid metabolism (Jeong et al. 2018). Another cross-sectional cohort study based on 79 metabolites indicated that annual ultrafine particles

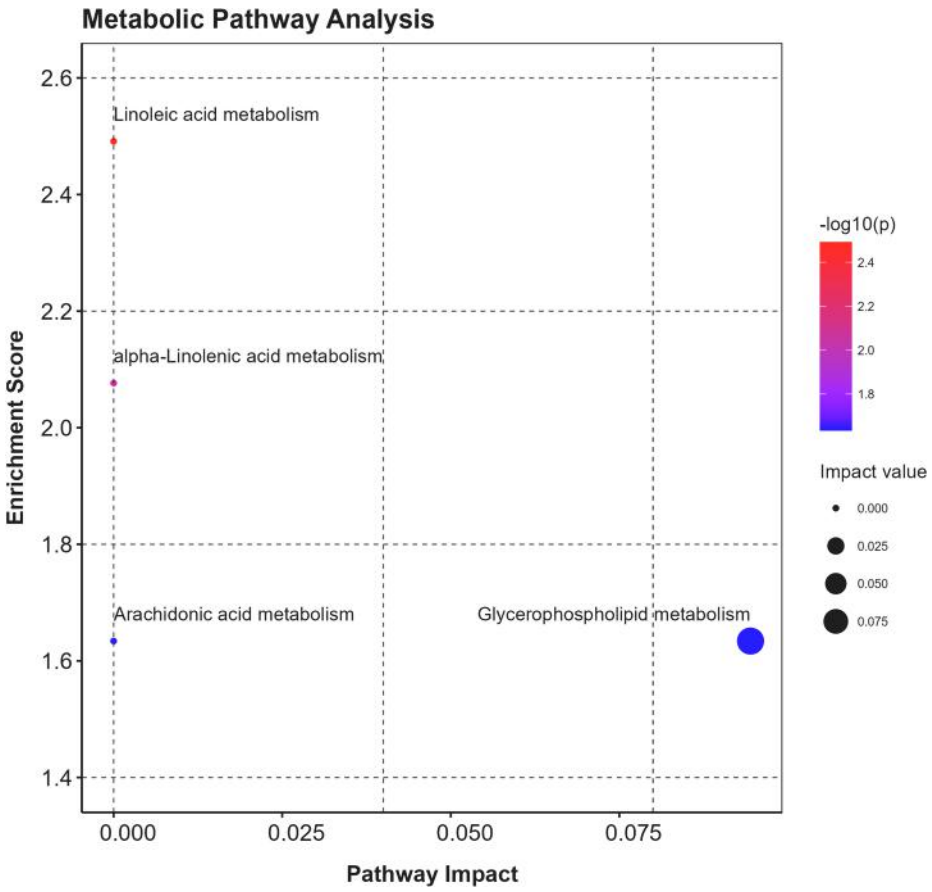


Fig. 3. Metabolic pathways identified for long-term exposure to $PM_{2.5abs}$, PM_{coarse} and NO_2. The plot is the same for $PM_{2.5abs}$, PM_{coarse} and NO_2, since the corresponding ID in the KEGG database for these four metabolites that were significantly associated with $PM_{2.5abs}$, PM_{coarse} or NO_2 exposure is identical. The pathway analysis is based on both enrichment analysis and pathway topology analysis. The Y-axis is the negative logarithm of the *p*-value (logarithmic base of 10) from the enrichment test. The X-axis indicates the structural impact of $PM_{2.5abs}$, PM_{coarse} or NO_2 related metabolites in the enriched pathways, which is based on the cumulative importance of all the significant metabolites within the pathway. The size of each bubble represents the impact value. The colour of each bubble represents the significance of the enrichment. PM_{coarse} = particulate matter with an aerodynamic diameter of 2.5–10 μm; $PM_{2.5abs}$ = $PM_{2.5}$ absorbance; NO_2 = nitrogen dioxide.

(UFP) exposure was associated with metabolites that might increase oxidative stress and affect inflammatory processes and endothelial function (Walker et al. 2019). Two studies within the Normative Aging Study (NAS, a closed cohort study) reported that long-term exposure to $PM_{2.5}$ species (e.g., UFP, black carbon), $PM_{2.5}$, and air temperature was associated with perturbed metabolic pathways, including glycerophospholipid, sphingolipid, and biosynthesis of unsaturated fatty acids etc. (Nassan et al. 2021a; Nassan et al. 2021b). They also reported that long-term NO_2 exposure was positively associated with four lipid metabolites, while these metabolites were not significantly associated with any metabolomics pathway.

Since untargeted metabolomics was used in all these studies, comparing our results with those from single metabolite levels is difficult. Nevertheless, they are mostly consistent in identifying metabolic pathways related to inflammation, unsaturated fatty acids, and glycerophospholipid associated with long-term exposure to air pollution. Potential differences in results compared with our findings might also be due to the small sample size (less than 1,000 participants) of some of these studies, differences in study designs (e.g., case-control study), and selected study population (e.g., older men).

Several studies reported the associations between metabolomics and short-term and intermediate exposure to air pollution. A longitudinal study on the effects of high-level $PM_{2.5}$ exposure on serum metabolomics reported that metabolites related to phospholipid metabolism (lysophosphatidic acid, phospholipid acid, and lysophosphatidylethanolamine) were decreased for a 10 $\mu g/m^3$ increase in $PM_{2.5}$ (Huan et al. 2021), which supports our findings to some extent where four phosphatidylcholine metabolites were decreased in association with $PM_{2.5abs}$, PM_{coarse} and NO_2. In a previous cross-sectional analysis based on KORA S4, F4 and the follow-up of survey 3 (KORA F3), Ward-Caviness et al. observed a significant positive association between one lysophosphatidylcholine (LPC) and short-term NO_2 exposures (Ward-Caviness et al. 2016). This longitudinal analysis did not find any associations between LPC and long-term exposures to NO_2 or other air pollutants. However, we observed decreased levels of four PCs in association with long-term $PM_{2.5abs}$, PM_{coarse} and NO_2 exposure. A perturbation between LPC and PC was also reported in two other studies investigating the associations between short-term exposure to air pollution and untargeted metabolomics profiling (Chen et al. 2019a; Yan et al. 2019). Chen et al. indicated that two fatty acids, five phospholipids (phosphatidylserine, PEs, phosphatidic acid), and one sphingosine in urine significantly decreased with a higher short-term $PM_{2.5}$ exposure; these metabolites were related to energy metabolism, oxidative stress and inflammation (Chen et al. 2019a). In the second study, Yan et al. investigated the associations between exposure to traffic-related air pollution (NO_x, $PM_{2.5}$) during the first trimester and serum metabolites measured in mid-pregnancy. They observed that higher exposure to air pollution was related to alterations in several oxidative stress and inflammatory pathways, including fatty acid, phospholipid, linoleate, and eicosanoid metabolism (Yan et al. 2019).

Phosphatidylcholine (PC) is the representative and important component of lipoproteins that belongs to glycerophospholipid. It has a polar phosphocholine head group, which is connected via a glycerol backbone to two fatty acid side chains of varying lengths and degrees of saturation. The fatty acids are bound to the sn1 and sn2 positions of the glycerol backbone, either via two esters (acyl) bonds (diacyl-PC, PC aa) or by one ester and one ether (alkyl) bond (acyl-alkyl-PC, PC ae). It is the most abundant phospholipid in all mammalian cell membranes and subcellular organelles and could be attacked by reactive oxygen species (ROS) and lead to lipid peroxidation, especially the polyunsaturated fatty acids (Ayala et al. 2014; Cole et al. 2012; van der Veen et al. 2017). PC and LPC serve as reservoirs and transporters of glycerophospholipid components: fatty acids, phosphate, glycerol, and choline, which could regulate homeostatic and inflammatory processes.

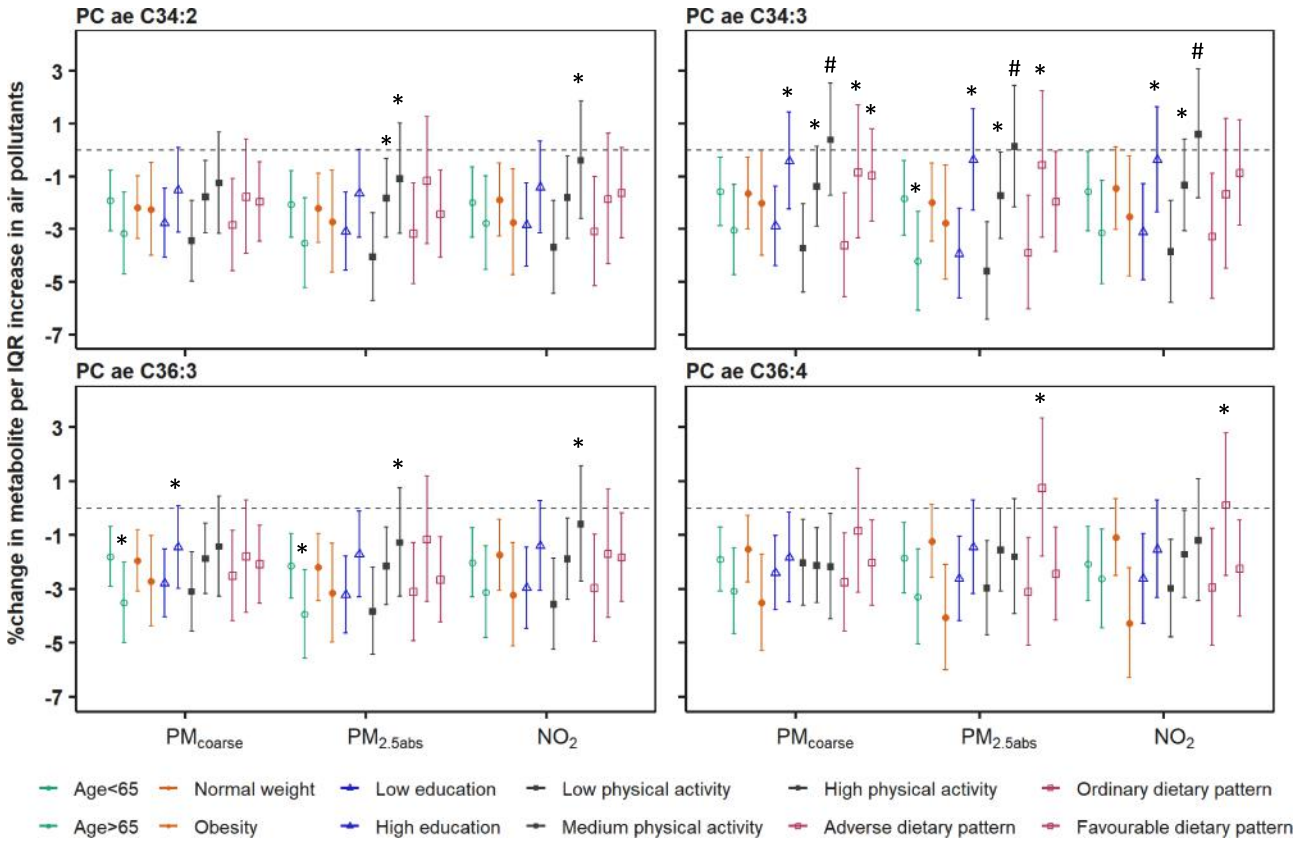


Fig. 4. Percent changes (95% CIs) of metabolites per IQR increase in air pollutant concentrations stratified by age, BMI, educational attainment, physical activity level and dietary patterns. Results were from our main models adjusted for covariates including age, sex, BMI, season, an indicator of each visit (KORA S4, KORA F4, or KORA FF4), educational attainment, smoking status, fasting status, alcohol consumption, physical activity, and dietary score, while the continuous variable will be replaced by each corresponding effect modifier. An IQR increase was 1.36 μg/m^3 for PM_{coarse}, 0.27 × 10^{-5}/m for $PM_{2.5abs}$, and 6.86 μg/m^3 for NO_2. PC ae: acyl-alkyl phosphatidylcholine. PM_{coarse} = particulate matter with an aerodynamic diameter of 2.5–10 μm; $PM_{2.5abs}$ = $PM_{2.5}$ absorbance; NO_2 = nitrogen dioxide. * $p < 0.05$ # $p < 0.004$ (0.05/(3*4)).

Lysophosphatidylcholine (LPC) could be hydrolyzed by PC via Phospholipase A2 (PLA2). The decreased concentrations of PCs, as well as increased PLA2, might indicate increased turnover of PCs for the synthesis of pro- and anti-inflammatory factors (Kertys et al. 2020). In an *in vivo* study, significant reductions in LPC and PC concentrations were observed after chronically exposing to ambient $PM_{2.5}$, which might result from repeated inflammation (Chen et al. 2014). This might explain the negative associations between PCs and long-term NO_2 exposure in this study compared to the study of Ward-Caviness and colleagues, which observed a positive association between LPCs and short-term NO_2 exposure. Furthermore, given the positive associations between CRP and long-term exposure to air pollution (PM_{coarse}, PM_{10}, PNC, and NO_x), these findings suggested a perturbation of anti-inflammation and pro-inflammation after long-term exposure to air pollution.

The higher abundance of PC in human tissues compared to other phospholipid classes has been shown to play an important role in health and diseases (van der Veen et al. 2017). The inhibition of hepatic PC synthesis and changes in hepatic phospholipid composition were related to fatty liver disease and impaired liver regeneration after surgery (van der Veen et al. 2017). The altered PC metabolism may also promote the development of Alzheimer's and cardiovascular diseases (CVD) (Tang et al. 2013; Whiley et al. 2014). In other metabolomics-related studies using the KORA cohort, a decrease in a few acyl-alkyl-PCs was associated with smoking (including an overlapped PC ae C34:3) and ageing (Chak et al. 2019; Xu et al. 2013). Plasmalogens are a subclass of alkyl-PCs with antioxidant properties (Engelmann 2004). A decreased serum concentration of acyl-alkyl-PCs and alkenyl-PC (plasmalogen) lipids were found in stable coronary artery disease (CAD) and acute myocardial infarction (MI) (Moxon et al. 2017; Sutter et al. 2016). Our findings, therefore, might indicate that acyl-alkyl-PCs could be the underlying biomarkers involved in the biological mechanisms of chronic air pollution exposure-associated diseases.

Apart from environmental impacts, pathological stimuli and normal physiological variations can also lead to differences in metabolic profiles (Lacruz et al. 2016; Soininen et al. 2015; Suhre et al. 2010). Lifestyle factors, including obesity, smoking, alcohol consumption, physical activity and dietary patterns, were considered risk factors for metabolism (Lacruz et al. 2016). Previous studies reported that male, older, obese, smoking, and unhealthy individuals are more susceptible to air pollution exposures (Chen et al. 2019b; Hou et al. 2020; Sun et al. 2020a; Yazdi et al. 2021; Zhang et al. 2021). We observed that older participants (>65 years old) showed a stronger association than the younger ones, which could be explained by a greater susceptibility in older adults to oxidative stress and also a higher prevalence of pre-existing diseases in the older group (Peters et al. 2021; Sacks et al. 2011). We did not find significant differences between males and females in most metabolites except that a few PCs decreased more in females than males when exposed to O_3. In contrast, males were more susceptible when exposed to the other air pollutants. The obese subgroup was more susceptible to long-term exposure to air pollution in this study, which could be hypothesized that altered PCs facilitate inflammation in obese participants. Physical activity has immediate beneficial effects, accumulating over time. In the long run, it reduces the risk of developing cardiovascular and respiratory diseases, type 2 diabetes, and certain types of cancers and reduces the risk of all-cause and cause-specific mortality (Tainio

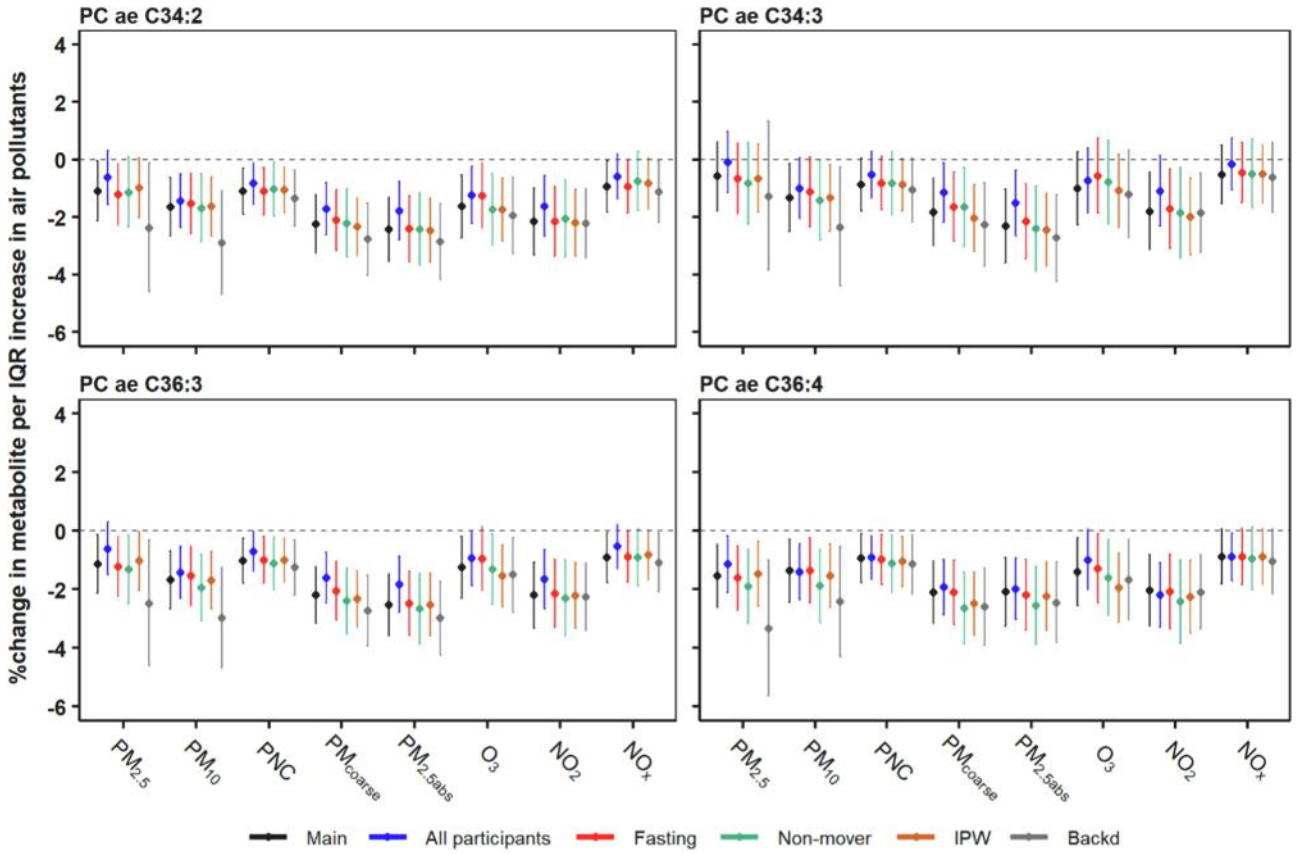


Fig. 5. Percent changes (95% CIs) of metabolites per IQR increase in air pollutant concentrations in different sensitivity analyses. The confounders used in different sensitivity analyses were the same as used in our main model including age, sex, BMI, season, indicator of each visit, educational attainment, smoking status, fasting status, alcohol consumption, physical activity, and dietary score. An IQR increase was 1.40 μg/m^3 for $PM_{2.5}$, 2.06 μg/m^3 for PM_{10}, 1.36 μg/m^3 for PM_{coarse}, 1.95 × 10^3/cm^3 for PNC, 0.27 × 10^{-5}/m for $PM_{2.5abs}$, 6.86 μg/m^3 for NO_2, 8.69 μg/m^3 for NO_x, and 3.45 μg/m^3 for O_3. Main: results from the main models (participants with repeated measurements); All participants: all participants with at least one visit in KORA S4, KORA F4 or KORA FF4; Fasting: participants who did overnight fasting; No-mover: participants who never change their residences during the entire study; IPW: further add predicted inverse probability of each participant into the main models; Backd: back-extrapolated annual average air pollutant concentrations were used rather the LUR estimated air pollutant concentrations. PC ae: acyl-alkyl phosphatidylcholine. $PM_{2.5}$ = particulate matter with an aerodynamic diameter less than or equal to 2.5 μm; PM_{coarse} = particulate matter with an aerodynamic diameter of 2.5–10 μm; PM_{10} = particulate matter with an aerodynamic diameter less than or equal to 10 μm; $PM_{2.5abs}$ = $PM_{2.5}$ absorbance; PNC = particle number concentration; NO_2 = nitrogen dioxide; NO_x = nitrogen oxide; O_3 = ozone.

et al. 2021). Our results suggested that higher physical activity attenuated the adverse effects of air pollution, which suggests that the long-term beneficial effects of physical activity might outweigh the harmful effects of air pollution, as previous studies reported (Fuertes et al. 2018; Sun et al. 2020b). Participants with low educational attainment were more vulnerable to air pollution exposure. Low education has been associated with increased susceptibility to adverse health effects of air pollution due to a higher prevalence of pre-existing diseases and limited access to medical care and fresh foods (Sacks et al. 2011). We observed that individuals with healthier dietary patterns showed weaker associations with air pollution, which might follow the findings that sufficient intakes of essential micronutrients (e.g., vitamins and long-chain polyunsaturated fatty acids) could modulate air pollution-induced harmful effects by reducing the oxidative stress and inflammatory response (Lim et al. 2019; Péter et al. 2015).

The targeted metabolomics approach used in our study has the strength to give an annotation of all metabolites compared to untargeted metabolomics analysis (unknown metabolites were also quantified), which might mislead false annotation for metabolites. To our knowledge, this is the first study using repeated measurements of targeted metabolomics to explore the health effects of long-term exposure to ambient air pollution within a population-representative cohort study of adults, and also with the largest number of study participants. We further assessed multiple air pollutants, including different particle matters and gaseous air pollutants. Moreover, the KORA cohort is a well-characterized study with standardized and comprehensive methods to collect individual information, enhancing our results' reliability. The longitudinal study design with repeated measurements of biomarkers strengthened statistical power and reduced potential residual confounding from unmeasured factors. It might also provide analytical improvement to previous cross-sectional analyses despite the lack of replication by other cohorts. Furthermore, the residential air pollutant concentrations were estimated using well-defined LUR models, which captured the spatial variation in exposure and enabled us to conclude consistent patterns across various air pollutants, reducing the risk of chance findings. This study also has the strength to assess the susceptibility from both external and intrinsic factors, especially dietary intake and lifestyle, which are known to affect the human metabolome. However, targeted metabolomics lowered the opportunity for new biomarkers discovery and could not fully represent the whole metabolome. Another limitation of our study is that the annual average concentrations of air pollutants were estimated using spatial models for 2014–15. We believe these exposure estimates are valid for the historical spatial contrasts because previous studies have shown that the spatial variation in exposure remained stable over time (de Hoogh et al. 2018; Eeftens et al. 2011; Gulliver et al. 2013; Wang et al. 2013). Using the air pollution concentrations obtained with a back-extrapolation approach, we investigated the potential effects of temporal variation. In addition,

we restricted our study to non-movers (participants who did not move within the study period) to reduce the exposure misclassification. The robust results from both analyses validated our exposure assessment approach. Nevertheless, we cannot rule out the potential impact of measurement error and note that exposure measurement error driven by spatial and/or temporal misalignments could lead to biases in either direction, as well as incorrect standard errors of health effect estimates.

6. Conclusions

Our study suggested that long-term air pollution exposure is associated with metabolic alterations, particularly in PCs with unsaturated long-chain fatty acids. These findings could provide new insights into potential mechanisms for air pollution-associated adverse outcomes in the general adult population.

CRediT authorship contribution statement

Yueli Yao: Conceptualization, Formal analysis, Investigation, Methodology, Visualization, Writing - original draft, Writing - review & editing. **Alexandra Schneider:** Conceptualization, Supervision, Writing - review & editing. **Kathrin Wolf:** Conceptualization, Writing - review & editing. **Siqi Zhang:** Conceptualization, Writing - review & editing. **Rui Wang-Sattler:** Conceptualization, Writing - review & editing. **Annette Peters:** Conceptualization, Data curation, Funding acquisition, Supervision, Writing - review & editing. **Susanne Breitner:** Conceptualization, Supervision, Writing - review & editing.

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.

Acknowledgments

We thank all participants for their long-term commitment to the KORA study, the staff for data collection and research data management and the members of the KORA Study Group (https://www.helmholtz-munich.de/en/epi/cohort/kora) who are responsible for the design and conduct of the study.

Funding

This work was supported by a scholarship under the State Scholarship Fund by the China Scholarship Council (File No. 201906180003). The KORA study was initiated and financed by the Helmholtz Zentrum München – German Research Center for Environmental Health, which is funded by the German Federal Ministry of Education and Research (BMBF) and by the State of Bavaria. Data collection in the KORA study is done in cooperation with the University Hospital of Augsburg. Furthermore, this study was supported by the Helmholtz Alliance "Aging and Metabolic Programming" (AMPro).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envint.2022.107632.

References

Ayala, A., Muñoz, M.F., Argüelles, S., 2014. Lipid peroxidation: production, metabolism, and signaling mechanisms of malondialdehyde and 4-hydroxy-2-nonenal. Oxid. Med. Cellular Longevity 2014, 360438–360438. http://doi.org/10.1155/2014/360438.

Bae, H.R., Chandy, M., Aguilera, J., Smith, E.M., Nadeau, K.C., Wu, J.C., Paik, D.T., 2021. Adverse effects of air pollution-derived fine particulate matter on cardiovascular homeostasis and disease. Trends Cardiovasc. Med. https://doi.org/10.1016/j.tcm.2021.09.010.

Brook, R.D., Rajagopalan, S., Pope, C.A., Brook, J.R., Bhatnagar, A., Diez-Roux, A.V., Holguin, F., Hong, Y., Luepker, R.V., Mittleman, M.A., Peters, A., Siscovick, D., Smith, S.C., Whitsel, L., Kaufman, J.D., 2010. Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the American Heart Association. Circulation 121 (21), 2331–2378.

Cao, H., Li, B., Peng, W., Pan, L.i., Cui, Z.e., Zhao, W., Zhang, H., Tang, N., Niu, K., Sun, J., Han, X., Wang, Z., Liu, K., He, H., Cao, Y., Xu, Z., Shan, A., Meng, G.e., Sun, Y., Guo, C., Liu, X., Xie, Y., Wen, F., Shan, G., Zhang, L., 2020. Associations of long-term exposure to ambient air pollution with cardiac conduction abnormalities in Chinese adults: The CHCN-BTH cohort study. Environ. Int. 143, 105981.

Chak, C., Lacruz, M., Adam, J., Brandmaier, S., Covic, M., Huang, J., Meisinger, C., Tiller, D., Prehn, C., Adamski, J., Berger, U., Gieger, C., Peters, A., Kluttig, A., Wang-Sattler, R., 2019. Ageing Investigation Using Two-Time-Point Metabolomics Data from KORA and CARLA Studies. Metabolites 9 (3), 44.

Chen, C., Li, H., Niu, Y., Liu, C., Lin, Z., Cai, J., Li, W., Ge, W., Chen, R., Kan, H., 2019a. Impact of short-term exposure to fine particulate matter air pollution on urinary metabolome: A randomized, double-blind, crossover trial. Environ. Int. 130, 104878.

Chen, W., Lin, C., Yan, Y., Cheng, K.T., Cheng, T.-J., 2014. Alterations in rat pulmonary phosphatidylcholines after chronic exposure to ambient fine particulate matter. Mol. BioSyst. 10 (12), 3163–3169. https://doi.org/10.1039/C4MB00435C.

Chen, Z., Newgard, C.B., Kim, J.S., Ilkayeva, O., Alderete, T.L., Thomas, D.C., Berhane, K., Breton, C., Chatzi, L., Bastain, T.M., McConnell, R., Avol, E., Lurmann, F., Muehlbauer, M.J., Hauser, E.R., Gilliland, F.D., 2019b. Near-roadway air pollution exposure and altered fatty acid oxidation among adolescents and young adults – The interplay with obesity. Environ. Int. 130, 104935.

Cole, L.K., Vance, J.E., Vance, D.E., 2012. Phosphatidylcholine biosynthesis and lipoprotein metabolism. Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of. Lipids 1821 (5), 754–761. https://doi.org/10.1016/j.bbalip.2011.09.009.

de Hoogh, K., Chen, J., Gulliver, J., Hoffmann, B., Hertel, O., Ketzel, M., Bauwelinck, M., van Donkelaar, A., Hvidtfeldt, U.A., Katsouyanni, K., Klompmaker, J., Martin, R.V., Samoli, E., Schwartz, P.E., Stafoggia, M., Bellander, T., Strak, M., Wolf, K., Vienneau, D., Brunekreef, B., Hoek, G., 2018. Spatial PM2.5, NO2, O3 and BC models for Western Europe – Evaluation of spatiotemporal stability. Environ. Int. 120, 81–92.

Eeftens, M., Beelen, R., Fischer, P., Brunekreef, B., Meliefste, K., Hoek, G., 2011. Stability of measured and modelled spatial contrasts in NO2 over time. Occup. Environ. Med. 68 (10), 765–770.

Engelmann, B., 2004. Plasmalogens: targets for oxidants and major lipophilic antioxidants. Biochem. Soc. Trans. 32 (1), 147–150. https://doi.org/10.1042/bst0320147.

Fuertes, E., Markevych, I., Jarvis, D., Vienneau, D., de Hoogh, K., Antó, J.M., Bowatte, G., Bono, R., Corsico, A.G., Emtner, M., Gislason, T., Gullón, J.A., Heinrich, J., Henderson, J., Holm, M., Johannessen, A., Leynaert, B., Marcon, A., Marchetti, P., Moratalla, J.M., Pascual, S., Probst-Hensch, N., Sánchez-Ramos, J.L., Siroux, V., Sommar, J., Weyler, J., Kuenzli, N., Jacquemin, B., Garcia-Aymerich, J., 2018. Residential air pollution does not modify the positive association between physical activity and lung function in current smokers in the ECRHS study. Environ. Int. 120, 364–372.

Gaskins, A.J., Tang, Z., Hood, R.B., Ford, J., Schwartz, J.D., Jones, D.P., Laden, F., Liang, D., 2021. Periconception air pollution, metabolomic biomarkers, and fertility among women undergoing assisted reproduction. Environ. Int. 155, 106666.

Gulliver, J., de Hoogh, K., Hansell, A., Vienneau, D., 2013. Development and back-extrapolation of NO2 land use regression models for historic exposure assessment in Great Britain. Environ. Sci. Technol. 47 (14), 7804–7811. https://doi.org/10.1021/es4008849.

Haid, M., Muschet, C., Wahl, S., Romisch-Margl, W., Prehn, C., et al., 2018. Long-Term Stability of Human Plasma Metabolites during Storage at -80 degrees C. J. Proteome Res. 17 (1), 203–211. https://doi.org/10.1021/acs.jproteome.7b00518.

Hales, S., Atkinson, J., Metcalfe, J., Kuschel, G., Woodward, A., 2021. Long term exposure to air pollution, mortality and morbidity in New Zealand: Cohort study. Sci. Total Environ. 801, 149660 https://doi.org/10.1016/j.scitotenv.2021.149660.

Han, S., Huang, J., Foppiano, F., Prehn, C., Adamski, J., et al., 2022. TIGER: technical variation elimination for metabolomics data using ensemble learning architecture. Brief Bioinform 23, 2. https://doi.org/10.1093/bib/bbab535.

Holmes, E., Wilson, I.D., Nicholson, J.K., 2008. Metabolic Phenotyping in Health and Disease. Cell 134 (5), 714–717. https://doi.org/10.1016/j.cell.2008.08.026.

Hood, R.B., Liang, D., Tang, Z., Kloog, I., Schwartz, J., et al., 2022. Length of PM2.5 exposure and alterations in the serum metabolome among women undergoing infertility treatment. Environ. Epidemiol, 6, 1, e191. https://doi.org/10.1097/EE9.0000000000000191.

Hou, J., Liu, X., Tu, R., Dong, X., Zhai, Z., Mao, Z., Huo, W., Chen, G., Xiang, H., Guo, Y., Li, S., Wang, C., 2020. Long-term exposure to ambient air pollution attenuated the association of physical activity with metabolic syndrome in rural Chinese adults: A cross-sectional study. Environ. Int. 136, 105459.

Huan, S., Jin, S., Liu, H., Xia, W., Liang, G., Xu, S., Fang, X., Li, C., Wang, Q., Sun, X., Li, Y., 2021. Fine particulate matter exposure and perturbation of serum metabolome: A longitudinal study in Baoding, China. Chemosphere 276, 130102.

Huang, J., Huth, C., Covic, M., Troll, M., Adam, J., Zukunft, S., Prehn, C., Wang, L.i., Nano, J., Scheerer, M.F., Neschen, S., Kastenmüller, G., Suhre, K., Laxy, M., Schliess, F., Gieger, C., Adamski, J., Hrabe de Angelis, M., Peters, A., Wang-Sattler, R., 2020. Machine Learning Approaches Reveal Metabolic Signatures of Incident Chronic Kidney Disease in Individuals With Prediabetes and Type 2 Diabetes. Diabetes 69 (12), 2756–2765.

Jeong, A., Fiorito, G., Keski-Rahkonen, P., Imboden, M., Kiss, A., Robinot, N., Gmuender, H., Vlaanderen, J., Vermeulen, R., Kyrtopoulos, S., Herceg, Z., Ghantous, A., Lovison, G., Galassi, C., Ranzi, A., Krogh, V., Grioni, S., Agnoli, C., Sacerdote, C., Mostafavi, N., Naccarati, A., Scalbert, A., Vineis, P., Probst-Hensch, N., 2018. Perturbation of metabolic pathways mediates the association of air pollutants with asthma and cardiovascular diseases. Environ. Int. 119, 334–345.

Kasdagli, M.-I., Katsouyanni, K., de Hoogh, K., Lagiou, P., Samoli, E., 2022. Investigating the association between long-term exposure to air pollution and greenness with mortality from neurological, cardio-metabolic and chronic obstructive pulmonary diseases in Greece. Environ. Pollut. 292, 118372 https://doi.org/10.1016/j.envpol.2021.118372.

Kertys, M., Grendar, M., Kosutova, P., Mokra, D., Mokry, J., 2020. Plasma based targeted metabolomic analysis reveals alterations of phosphatidylcholines and oxidative stress markers in guinea pig model of allergic asthma. Biochim. Biophys. Acta Mol. Basis Dis. 1866 (1), 165572 https://doi.org/10.1016/j.bbadis.2019.165572.

Lacruz, M.E., Kluttig, A., Tiller, D., Medenwald, D., Giegling, I., Rujescu, D., Prehn, C., Adamski, J., Frantz, S., Greiser, K.H., Emeny, R.T., Kastenmüller, G., Haerting, J., 2016. Cardiovascular Risk Factors Associated With Blood Metabolite Concentrations and Their Alterations During a 4-Year Period in a Population-Based Cohort. Circ. Cardiovasc. Genet. 9 (6), 487–494.

Li, H., Cai, J., Chen, R., Zhao, Z., Ying, Z., Wang, L., Chen, J., Hao, K.e., Kinney, P.L., Chen, H., Kan, H., 2017. Particulate Matter Exposure and Stress Hormone Levels: A Randomized, Double-Blind, Crossover Trial of Air Purification. Circulation 136 (7), 618–627.

Li, Z., Liang, D., Ye, D., Chang, H.H., Ziegler, T.R., Jones, D.P., Ebelt, S.T., 2021. Application of high-resolution metabolomics to identify biological pathways perturbed by traffic-related air pollution. Environ. Res. 193, 110506.

Liang, D., Moutinho, J.L., Golan, R., Yu, T., Ladva, C.N., Niedzwiecki, M., Walker, D.I., Sarnat, S.E., Chang, H.H., Greenwald, R., Jones, D.P., Russell, A.G., Sarnat, J.A., 2018. Use of high-resolution metabolomics for the identification of metabolic signals associated with traffic-related air pollution. Environ. Int. 120, 145–154.

Lim, C.C., Hayes, R.B., Ahn, J., Shao, Y., Silverman, D.T., Jones, R.R., Thurston, G.D., 2019. Mediterranean Diet and the Association Between Air Pollution and Cardiovascular Disease Mortality Risk. Circulation 139 (15), 1766–1775.

Liu, S., Jørgensen, J.T., Ljungman, P., Pershagen, G., Bellander, T., Leander, K., Magnusson, P.K.E., Rizzuto, D., Hvidtfeldt, U.A., Raaschou-Nielsen, O., Wolf, K., Hoffmann, B., Brunekreef, B., Strak, M., Chen, J., Mehta, A., Atkinson, R.W., Bauwelinck, M., Varraso, R., Boutron-Ruault, M.-C., Brandt, J., Cesaroni, G., Forastiere, F., Fecht, D., Gulliver, J., Hertel, O., de Hoogh, K., Janssen, N.A.H., Katsouyanni, K., Ketzel, M., Klompmaker, J.O., Nagel, G., Oftedal, B., Peters, A., Tjønneland, A., Rodopoulou, S.P., Samoli, E., Bekkevold, T., Sigsgaard, T., Stafoggia, M., Vienneau, D., Weinmayr, G., Hoek, G., Andersen, Z.J., 2021. Long-term exposure to low-level air pollution and incidence of chronic obstructive pulmonary disease: The ELAPSE project. Environ. Int. 146, 106267.

Menni, C., Metrustry, S.J., Mohney, R.P., Beevers, S., Barratt, B., Spector, T.D., Kelly, F. J., Valdes, A.M., 2015. Circulating levels of antioxidant vitamins correlate with better lung function and reduced exposure to ambient pollution. Am. J. Respir. Crit. Care Med. 191 (10), 1203–1207.

Mortamais, M., Gutierrez, L.-A., de Hoogh, K., Chen, J., Vienneau, D., Carrière, I., Letellier, N., Helmer, C., Gabelle, A., Mura, T., Sunyer, J., Benmarhnia, T., Jacquemin, B., Berr, C., 2021. Long-term exposure to ambient air pollution and risk of dementia: Results of the prospective Three-City Study. Environ. Int. 148, 106376.

Moxon, J.V., Jones, R.E., Wong, G., Weir, J.M., Mellett, N.A., Kingwell, B.A., Meikle, P.J., Golledge, J., 2017. Baseline serum phosphatidylcholine plasmalogen concentrations are inversely associated with incident myocardial infarction in patients with mixed peripheral artery disease presentations. Atherosclerosis 263, 301–308.

Nassan, F.L., Kelly, R.S., Kosheleva, A., Koutrakis, P., Vokonas, P.S., Lasky-Su, J.A., Schwartz, J.D., 2021a. Metabolomic signatures of the long-term exposure to air pollution and temperature. Environ. Health 20 (1). https://doi.org/10.1186/s12940-020-00683-x.

Nassan, F.L., Wang, C., Kelly, R.S., Lasky-Su, J.A., Vokonas, P.S., Koutrakis, P., Schwartz, J.D., 2021b. Ambient PM2.5 species and ultrafine particle exposure and their differential metabolomic signatures. Environ. Int. 151, 106447.

Nemmar, A., Hoet, P.H.M., Vanquickenborne, B., Dinsdale, D., Thomeer, M., Hoylaerts, M.F., Vanbilloen, H., Mortelmans, L., Nemery, B., 2002. Passage of inhaled particles into the blood circulation in humans. Circulation 105 (4), 411–414.

Pang, Z., Chong, J., Zhou, G., de Lima Morais, D.A., Chang, L., et al. 2021. MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. Nucl. Acids Res., 49, W1, W388–W396. https://doi.org/10.1093/nar/gkab382.

Park, J., Kim, H.-J., Lee, C.-H., Lee, C.H., Lee, H.W., 2021. Impact of long-term exposure to ambient air pollution on the incidence of chronic obstructive pulmonary disease: A systematic review and meta-analysis. Environ. Res. 194, 110703 https://doi.org/10.1016/j.envres.2020.110703.

Péter, S., Holguin, F., Wood, L., Clougherty, J., Raederstorff, D., Antal, M., Weber, P., Eggersdorfer, M., 2015. Nutritional Solutions to Reduce Risks of Negative Health Impacts of Air Pollution. Nutrients 7 (12), 10398–10416.

Peters, A., Nawrot, T.S., Baccarelli, A.A., 2021. Hallmarks of environmental insults. Cell 184 (6), 1455–1468. https://doi.org/10.1016/j.cell.2021.01.043.

Pilz, V., Wolf, K., Breitner, S., Rückerl, R., Koenig, W., Rathmann, W., Cyrys, J., Peters, A., Schneider, A., 2018. C-reactive protein (CRP) and long-term air pollution with a focus on ultrafine particles. Int. J. Hyg. Environ. Health 221 (3), 510–518.

Rabel, M., Laxy, M., Thorand, B., Peters, A., Schwettmann, L., et al., 2018. Clustering of Health-Related Behavior Patterns and Demographics. Results From the Population-Based KORA S4/F4 Cohort Study. Front. Public Health 6, 387. https://doi.org/10.3389/fpubh.2018.00387.

Rappaport, S.M., Barupal, D.K., Wishart, D., Vineis, P., Scalbert, A., 2014. The Blood Exposome and Its Role in Discovering Causes of Disease. Environ. Health Perspect. 122 (8), 769–774.

Rinschen, M.M., Ivanisevic, J., Giera, M., Siuzdak, G., 2019. Identification of bioactive metabolites using activity metabolomics. Nat. Rev. Mol. Cell Biol. 20 (6), 353–367. https://doi.org/10.1038/s41580-019-0108-4.

Ritz, B., Yan, Q.i., He, D.i., Wu, J., Walker, D.I., Uppal, K., Jones, D.P., Heck, J.E., 2022. Child serum metabolome and traffic-related air pollution exposure in pregnancy. Environ. Res. 203, 111907.

Römisch-Margl, W., Prehn, C., Bogumil, R., Röhring, C., Suhre, K., Adamski, J., 2012. Procedure for tissue sample preparation and metabolite extraction for high-throughput targeted metabolomics. Metabolomics 8 (1), 133–142.

Sacks, J.D., Stanek, L.W., Luben, T.J., Johns, D.O., Buckley, B.J., Brown, J.S., Ross, M., 2011. Particulate matter-induced health effects: who is susceptible? Environ. Health Perspect. 119 (4), 446–454.

Soininen, P., Kangas, A.J., Wurtz, P., Suna, T., Ala-Korpela, M., 2015. Quantitative serum nuclear magnetic resonance metabolomics in cardiovascular epidemiology and genetics. Circ. Cardiovasc. Genet. 8 (1), 192–206. https://doi.org/10.1161/CIRCGENETICS.114.000216.

Suhre, K., Meisinger, C., Doring, A., Altmaier, E., Belcredi, P., et al. 2010. Metabolic footprint of diabetes: a multiplatform metabolomics study in an epidemiological setting. PLoS One 5, 11, e13953. https://doi.org/10.1371/journal.pone.0013953.

Sun, S., Cao, W., Chan, K.-P., Ran, J., Ge, Y., Zhang, Y., Feng, Y., Zeng, Q., Lee, R.-Y., Wong, C.-M., Tian, L., Lei, Y., 2020a. Cigarette smoking increases deaths associated with air pollution in Hong Kong. Atmos. Environ. 223, 117266.

Sun, S., Cao, W., Qiu, H., Ran, J., Lin, H., et al. 2020b. Benefits of physical activity not affected by air pollution: a prospective cohort study. Int. J. Epidemiol., 49, 1, 142–152. https://doi.org/10.1093/ije/dyz184.

Sutter, I., Klingenberg, R., Othman, A., Rohrer, L., Landmesser, U., Heg, D., Rodondi, N., Mach, F., Windecker, S., Matter, C.M., Lüscher, T.F., von Eckardstein, A., Hornemann, T., 2016. Decreased phosphatidylcholine plasmalogens–A putative novel lipid signature in patients with stable coronary artery disease and acute myocardial infarction. Atherosclerosis 246, 130–140.

Tainio, M., Jovanovic Andersen, Z., Nieuwenhuijsen, M.J., Hu, L., de Nazelle, A., An, R., Garcia, L.M.T., Goenka, S., Zapata-Diomedi, B., Bull, F., Sá, T.H.d., 2021. Air pollution, physical activity and health: A mapping review of the evidence. Environ. Int. 147, 105954.

Tang, W.H.W., Wang, Z., Levison, B.S., Koeth, R.A., Britt, E.B., Fu, X., Wu, Y., Hazen, S.L., 2013. Intestinal microbial metabolism of phosphatidylcholine and cardiovascular risk. N. Engl. J. Med. 368 (17), 1575–1584.

van der Veen, J.N., Kennelly, J.P., Wan, S., Vance, J.E., Vance, D.E., Jacobs, R.L., 2017. The critical role of phosphatidylcholine and phosphatidylethanolamine metabolism in health and disease. Biochimica et Biophysica Acta (BBA) - Biomembranes 1859 (9), 1558–1572.

van Veldhoven, K., Kiss, A., Keski-Rahkonen, P., Robinot, N., Scalbert, A., Cullinan, P., Chung, K.F., Collins, P., Sinharay, R., Barratt, B.M., Nieuwenhuijsen, M., Rodoreda, A.A., Carrasco-Turigas, G., Vlaanderen, J., Vermeulen, R., Portengen, L., Kyrtopoulos, S.A., Ponzi, E., Chadeau-Hyam, M., Vineis, P., 2019. Impact of short-term traffic-related air pollution on the metabolome – Results from two metabolome-wide experimental studies. Environ. Int. 123, 124–131.

Vlaanderen, J.J., Janssen, N.A., Hoek, G., Keski-Rahkonen, P., Barupal, D.K., Cassee, F. R., Gosens, I., Strak, M., Steenhof, M., Lan, Q., Brunekreef, B., Scalbert, A., Vermeulen, R.C.H., 2017. The impact of ambient air pollution on the human blood metabolome. Environ. Res. 156, 341–348.

Walker, D.I., Lane, K.J., Liu, K., Uppal, K., Patton, A.P., Durant, J.L., Jones, D.P., Brugge, D., Pennell, K.D., 2019. Metabolomic assessment of exposure to near-highway ultrafine particles. J. Eposure Sci. Environ. Epidemiol. 29 (4), 469–483.

Wang, R., Henderson, S.B., Sbihi, H., Allen, R.W., Brauer, M., 2013. Temporal stability of land use regression models for traffic-related air pollution. Atmos. Environ. 64, 312–319. https://doi.org/10.1016/j.atmosenv.2012.09.056.

Ward-Caviness, C.K., Breitner, S., Wolf, K., Cyrys, J., Kastenmüller, G., Wang-Sattler, R., Schneider, A., Peters, A., 2016. Short-term NO2 exposure is associated with long-chain fatty acids in prospective cohorts from Augsburg, Germany: results from an analysis of 138 metabolites and three exposures. Int. J. Epidemiol. 45 (5), 1528–1538.

Weuve, J., Tchetgen Tchetgen, E.J., Glymour, M.M., Beck, T.L., Aggarwal, N.T., Wilson, R.S., Evans, D.A., Mendes de Leon, C.F., 2012. Accounting for bias due to selective attrition: the example of smoking and cognitive decline. Epidemiology 23 (1), 119–128.

Whiley, L., Sen, A., Heaton, J., Proitsi, P., García-Gómez, D., Leung, R., Smith, N., Thambisetty, M., Kloszewska, I., Mecocci, P., Soininen, H., Tsolaki, M., Vellas, B., Lovestone, S., Legido-Quigley, C., 2014. Evidence of altered phosphatidylcholine metabolism in Alzheimer's disease. Neurobiol. Aging 35 (2), 271–278.

Wieder, C., Frainay, C., Poupin, N., Rodríguez-Mier, P., Vinson, F., et al., 2021. Pathway analysis in metabolomics: Recommendations for the use of over-representation analysis. PLoS Comput. Biol. 17, 9, https://doi.org/e1009105. 10.1371/journal.pcbi.1009105.

Winkler, G., Döring, A., 1998. Validation of a short qualitative food frequency list used in several German large scale surveys. Zeitschrift für Ernährungswissenschaft 37 (3), 234–241. https://doi.org/10.1007/PL00007377.

Wolf, K., Cyrys, J., Harciníková, T., Gu, J., Kusch, T., Hampel, R., Schneider, A., Peters, A., 2017. Land use regression modeling of ultrafine particles, ozone, nitrogen oxides and markers of particulate matter pollution in Augsburg, Germany. Sci. Total Environ. 579, 1531–1540.

Wolf, K., Hoffmann, B., Andersen, Z.J., Atkinson, R.W., Bauwelinck, M., Bellander, T., Brandt, J., Brunekreef, B., Cesaroni, G., Chen, J., de Faire, U., de Hoogh, K., Fecht, D., Forastiere, F., Gulliver, J., Hertel, O., Hvidtfeldt, U.A., Janssen, N.A.H., Jørgensen, J.T., Katsouyanni, K., Ketzel, M., Klompmaker, J.O., Lager, A., Liu, S., MacDonald, C.J., Magnusson, P.K.E., Mehta, A.J., Nagel, G., Oftedal, B., Pedersen, N. L., Pershagen, G., Raaschou-Nielsen, O., Renzi, M., Rizzuto, D., Rodopoulou, S., Samoli, E., van der Schouw, Y.T., Schramm, S., Schwarze, P., Sigsgaard, T., Sørensen, M., Stafoggia, M., Strak, M., Tjønneland, A., Verschuren, W.M.M., Vienneau, D., Weinmayr, G., Hoek, G., Peters, A., Ljungman, P.L.S., 2021. Long-term exposure to low-level ambient air pollution and incidence of stroke and coronary heart disease: a pooled analysis of six European cohorts within the ELAPSE project. The Lancet Planetary Health 5 (9) e620–e632.

Xu, T., Holzapfel, C., Dong, X., Bader, E., Yu, Z., Prehn, C., Perstorfer, K., Jaremek, M., Roemisch-Margl, W., Rathmann, W., Li, Y., Wichmann, H.-E., Wallaschofski, H., Ladwig, K.H., Theis, F., Suhre, K., Adamski, J., Illig, T., Peters, A., Wang-Sattler, R., 2013. Effects of smoking and smoking cessation on human serum metabolite profile: results from the KORA cohort study. BMC Med. 11 (1) https://doi.org/10.1186/1741-7015-11-60.

Yan, Q.i., Liew, Z., Uppal, K., Cui, X., Ling, C., Heck, J.E., von Ehrenstein, O.S., Wu, J., Walker, D.I., Jones, D.P., Ritz, B., 2019. Maternal serum metabolome and traffic-related air pollution exposure in pregnancy. Environ. Int. 130, 104872.

Yazdi, M.D., Wang, Y., Di, Q., Requia, W.J., Wei, Y., Shi, L., Sabath, M.B., Dominici, F., Coull, B., Evans, J.S., Koutrakis, P., Schwartz, J.D., 2021. Long-term effect of exposure to lower concentrations of air pollution on mortality among US Medicare participants and vulnerable subgroups: a doubly-robust approach. The Lancet Planetary Health 5 (10) e689–e697.

Zhang, S., Mwiberi, S., Pickford, R., Breitner, S., Huth, C., Koenig, W., Rathmann, W., Herder, C., Roden, M., Cyrys, J., Peters, A., Wolf, K., Schneider, A., 2021. Longitudinal associations between ambient air pollution and insulin sensitivity: results from the KORA cohort study. The Lancet Planetary Health 5 (1), e39–e49.

Longitudinal associations between metabolites and long-term exposure to ambient air pollution: results from the KORA cohort study

Yueli Yao [a,b], Alexandra Schneider [b], Kathrin Wolf [b], Siqi Zhang [b], Rui Wang-Sattler [b,c], Annette Peters [a,b,c,d]*, Susanne Breitner [a,b]*

[a] Institute for Medical Information Processing, Biometry and Epidemiology – IBE, Ludwig-Maximilians-Universität München, Munich, Germany

[b] Institute of Epidemiology, Helmholtz Zentrum München – German Research Center for Environmental Health, Neuherberg, Germany

[c] German Center for Diabetes Research, DZD, Munich-Neuherberg, Germany

[d] German Centre for Cardiovascular Research, DZHK, Partner Site Munich, Munich, Germany

Supplementary Materials

Contents

Fig. S16. Exposure-response relationships between $PM_{2.5abs}$ and PC ae C34:2, PC ae C34:3, PC ae C36:6 and PC ae C36:4.

Fig. S17. Exposure-response relationships between PM_{coarse} and PC ae C34:2, PC ae C34:3, PC ae C36:6 and PC ae C36:4.

Fig. S18. Exposure-response relationships between $PM_{2.5abs}$ and PC ae C34:2, PC ae C34:3, PC ae C36:6 and PC ae C36:4 after excluding $PM_{2.5abs}$ higher than 99% of $PM_{2.5abs}$ values.

Fig. S19. Comparisons between main data (n =5772) and the outlier removed data (n = 5711) of $PM_{2.5abs}$.

Fig. S20. Exposure-response relationships between PM_{coarse} and PC ae C34:2, PC ae C34:3, PC ae C36:6 and PC ae C36:4 with excluding PM_{coarse} concentration < 5% of total PM_{coarse} and PM_{coarse} concentration > 95% of total PM_{coarse}.

Fig. S21. Comparisons between main data (n =5772) and outliers of PM_{coarse} removed (n = 5203).

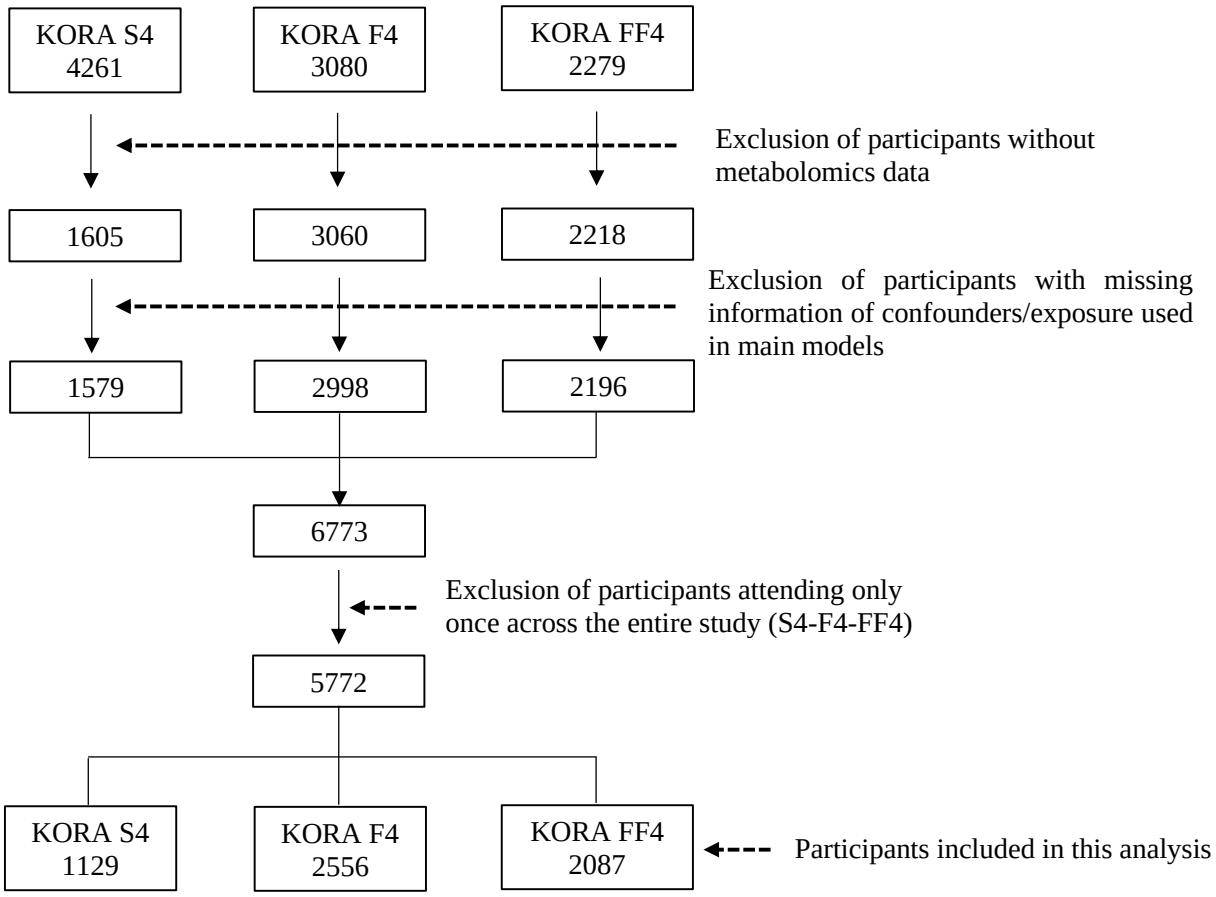


Fig. S1. Flow chart of participant exclusion process in this study.

Table S1 Abbreviations and full biochemical names of the 108 metabolites grouped by their compound classes.

	Abbreviation	**Biochemical name**
Acylcarnitine	C0	Carnitine
	C2	Acetylcarnitine
	C3	Propionylcarnitine
	C4	Butyrylcarnitine
	C10	Decanoylcarnitine
	C12	Dodecanoylcarnitine
	C14:1	Tetradecenoylcarnitine
	C14:2	Tetradecadienylcarnitine
	C16	Hexadecanylcarnitine
	C18	Octadecanoylcarnitine
	C18:1	Octadecenoylcarnitine
	C18:2	Octadecadienylcarnitine
Amino acids	Arg	Arginine
	Gln	Glutamine
	Gly	Glycine
	His	Histidine
	Met	Methionine
	Orn	Ornithine
	Phe	Phenylalanine
	Pro	Proline
	Ser	Serine
	Thr	Threonine
	Trp	Tryptophan
	Tyr	Tryosine
Phosphatidylcholines	PC aa C28:1	Phosphatidylcholine diacyl C28:1
	PC aa C30:0	Phosphatidylcholine diacyl C30:0
	PC aa C32:0	Phosphatidylcholine diacyl C32:0
	PC aa C32:1	Phosphatidylcholine diacyl C32:1
	PC aa C32:2	Phosphatidylcholine diacyl C32:2
	PC aa C32:3	Phosphatidylcholine diacyl C32:3
	PC aa C34:1	Phosphatidylcholine diacyl C34:1
	PC aa C34:2	Phosphatidylcholine diacyl C34:2
	PC aa C34:3	Phosphatidylcholine diacyl C34:3
	PC aa C34:4	Phosphatidylcholine diacyl C34:4
	PC aa C36:1	Phosphatidylcholine diacyl C36:1
	PC aa C36:2	Phosphatidylcholine diacyl C36:2
	PC aa C36:3	Phosphatidylcholine diacyl C36:3
	PC aa C36:4	Phosphatidylcholine diacyl C36:4
	PC aa C36:5	Phosphatidylcholine diacyl C36:5
	PC aa C36:6	Phosphatidylcholine diacyl C36:6
	PC aa C38:0	Phosphatidylcholine diacyl C38:0
	PC aa C38:3	Phosphatidylcholine diacyl C38:3
	PC aa C38:4	Phosphatidylcholine diacyl C38:4
	PC aa C38:5	Phosphatidylcholine diacyl C38:5
	PC aa C38:6	Phosphatidylcholine diacyl C38:6
	PC aa C40:2	Phosphatidylcholine diacyl C40:2
	PC aa C40:3	Phosphatidylcholine diacyl C40:3
	PC aa C40:4	Phosphatidylcholine diacyl C40:4
	PC aa C40:5	Phosphatidylcholine diacyl C40:5

	PC aa C40:6	Phosphatidylcholine diacyl C40:6
	PC aa C42:0	Phosphatidylcholine diacyl C42:0
	PC aa C42:1	Phosphatidylcholine diacyl C42:1
	PC aa C42:2	Phosphatidylcholine diacyl C42:2
	PC aa C42:4	Phosphatidylcholine diacyl C42:4
	PC aa C42:5	Phosphatidylcholine diacyl C42:5
	PC aa C42:6	Phosphatidylcholine diacyl C42:6
	PC ae C32.1	Phosphatidylcholine acyl-alkyl C32:1
	PC ae C32:2	Phosphatidylcholine acyl-alkyl C32:2
	PC ae C34:0	Phosphatidylcholine acyl-alkyl C34:0
	PC ae C34:1	Phosphatidylcholine acyl-alkyl C34:1
	PC ae C34:2	Phosphatidylcholine acyl-alkyl C34:2
	PC ae C34:3	Phosphatidylcholine acyl-alkyl C34:3
	PC ae C36:1	Phosphatidylcholine acyl-alkyl C36:1
	PC ae C36:2	Phosphatidylcholine acyl-alkyl C36:2
	PC ae C36:3	Phosphatidylcholine acyl-alkyl C36:3
	PC ae C36:4	Phosphatidylcholine acyl-alkyl C36:4
	PC ae C36:5	Phosphatidylcholine acyl-alkyl C36:5
	PC ae C38:0	Phosphatidylcholine acyl-alkyl C38:0
	PC ae C38:1	Phosphatidylcholine acyl-alkyl C38:1
	PC ae C38:2	Phosphatidylcholine acyl-alkyl C38:2
	PC ae C38:3	Phosphatidylcholine acyl-alkyl C38:3
	PC ae C38:4	Phosphatidylcholine acyl-alkyl C38:4
	PC ae C38:5	Phosphatidylcholine acyl-alkyl C38:5
	PC ae C38:6	Phosphatidylcholine acyl-alkyl C38:6
	PC ae C40:1	Phosphatidylcholine acyl-alkyl C40:1
	PC ae C40:2	Phosphatidylcholine acyl-alkyl C40:2
	PC ae C40:3	Phosphatidylcholine acyl-alkyl C40:3
	PC ae C40:4	Phosphatidylcholine acyl-alkyl C40:4
	PC ae C40:5	Phosphatidylcholine acyl-alkyl C40:5
	PC ae C40:6	Phosphatidylcholine acyl-alkyl C40:6
	PC ae C42:1	Phosphatidylcholine acyl-alkyl C42:1
	PC ae C42:2	Phosphatidylcholine acyl-alkyl C42:2
	PC ae C42:3	Phosphatidylcholine acyl-alkyl C42:3
	PC ae C42:4	Phosphatidylcholine acyl-alkyl C42:4
	PC ae C42:5	Phosphatidylcholine acyl-alkyl C42:5
	PC ae C44:3	Phosphatidylcholine acyl-alkyl C44:3
	PC ae C44:4	Phosphatidylcholine acyl-alkyl C44:4
	PC ae C44:5	Phosphatidylcholine acyl-alkyl C44:5
	PC ae C44:6	Phosphatidylcholine acyl-alkyl C44:6
Lysophosphatidylcholines	lysoPC a C16:0	lysoPhosphatidylcholine acyl C16:0
	lysoPC a C16:1	lysoPhosphatidylcholine acyl C16:1
	lysoPC a C18:0	lysoPhosphatidylcholine acyl C18:0
	lysoPC a C18:1	lysoPhosphatidylcholine acyl C18:1
	lysoPC a C18:2	lysoPhosphatidylcholine acyl C18:2
	lysoPC a C20:3	lysoPhosphatidylcholine acyl C20:3
	lysoPC a C20:4	lysoPhosphatidylcholine acyl C20:4
Sphingomyelins	SM (OH) C14:1	Hydroxysphingomyeline C14:1
	SM (OH) C16:1	Hydroxysphingomyeline C16:1
	SM (OH) C22:1	Hydroxysphingomyeline C22:1
	SM (OH) C22:2	Hydroxysphingomyeline C22:2
	SM (OH)C24:1	Hydroxysphingomyeline C24:1
	SM C16:0	Sphingomyeline C16:0

	SM C16:1	Sphingomyeline C16:1
	SM C18:0	Sphingomyeline C18:0
	SM C18:1	Sphingomyeline C18:1
	SM C20:2	Sphingomyeline C20:2
	SM C24:1	Sphingomyeline C24:1
Hexose	H1	Hexose

Text S1. Exposure assessment

Back-extrapolated air pollutant concentrations

Briefly, we generated the time series of daily pollutant concentrations covering the study period of KORA S4–FF4. Using data from routine monitoring sites, we calculated the absolute differences in annual average concentrations between the period of each visit (01.01.2000–31.12.2000 for S4, 01.01.2007–31.12.2007 for F4, and 01.07.2013–30.06.2014 for FF4) and the period of ULTRA III measurements (01.03.2014–15.04.2015). To account for the difference in monitoring devices used at routine monitoring sites and in ULTRA III measurements, we calculated the ratio of average concentrations at the monitoring sites for the ULTRA III measurement period to average concentrations at the 20 measurement sites in ULTRA III, and calibrated the absolute difference by multiplying the absolute difference by the ratio. We then calculated for each study participant the back-extrapolated concentration at each visit by adding the calibrated absolute difference to the LUR-model-estimated ULTRA III annual average concentrations. The back-extrapolated air pollution concentrations reflected not only the spatial variation but also the temporal variation in exposure (Zhang et al. 2021).

Short-term air pollutants

The daily average PM_{10} and $PM_{2.5}$ for the years 2004 to 2017 were obtained from an aerosol monitoring station (FH) located 1 km southeast of the city center of Augsburg. This monitoring station was established in 2004 and is considered as a representative site of the urban background in Augsburg. PM_{10} concentrations before 2004 were monitored at an urban background station (Bourges-Platz, BP) located 2 km north of the city center, and measured data were calibrated based on a linear regression model using the overlapping period with the FH site and yielded continuous PM_{10} data for the years 1985 to 2017. $PM_{2.5}$ concentrations before 2004 were calculated as a ratio from PM_{10} time series: $PM_{2.5} = PM_{10} \times 0.68$. PM_{coarse} was calculated by scaling PM_{10} with a factor of 0.32 until the end of 2004. From the beginning of 2005, it was calculated as the difference between PM_{10} and $PM_{2.5}$. Daily NO_2 and NO concentrations for the years 1985 to 2017 were obtained from the BP station, and missing values were imputed by monitoring data from a single urban background monitoring site (LfU, approximately 4 km south of the city center) operated by the Bavarian Environment Agency using linear regressions. Daily NO_x between 2013 and 2017 were obtained by a similar procedure as NO_2, while the years before 2013 was calculated by $NO_x = NO + NO_2$. O_3 concentrations had been measured at the Haunstetten monitoring station (a suburb 7 km south of the Augsburg city center) until 2000. From 2001 on, monitoring has been done at the LfU monitoring station. To account for the difference of monitoring devices at these two stations, we calibrated the measured data at Haunstetten station based on a linear regression model using the overlapping period. A continuous O_3 time-series dataset was derived for the years 1985 to 2017. Further missing rates of air pollutants were all less than 5% and imputed by the median of three-month moving averages, which accounted for the seasonal effects. Since the monitor data of black carbon (equal to $PM_{2.5abs}$) was not available for the KORA S4 period, and there were too many missing values of PNC before 2004, we didn't perform this sensitivity analysis for $PM_{2.5abs}$ and PNC.

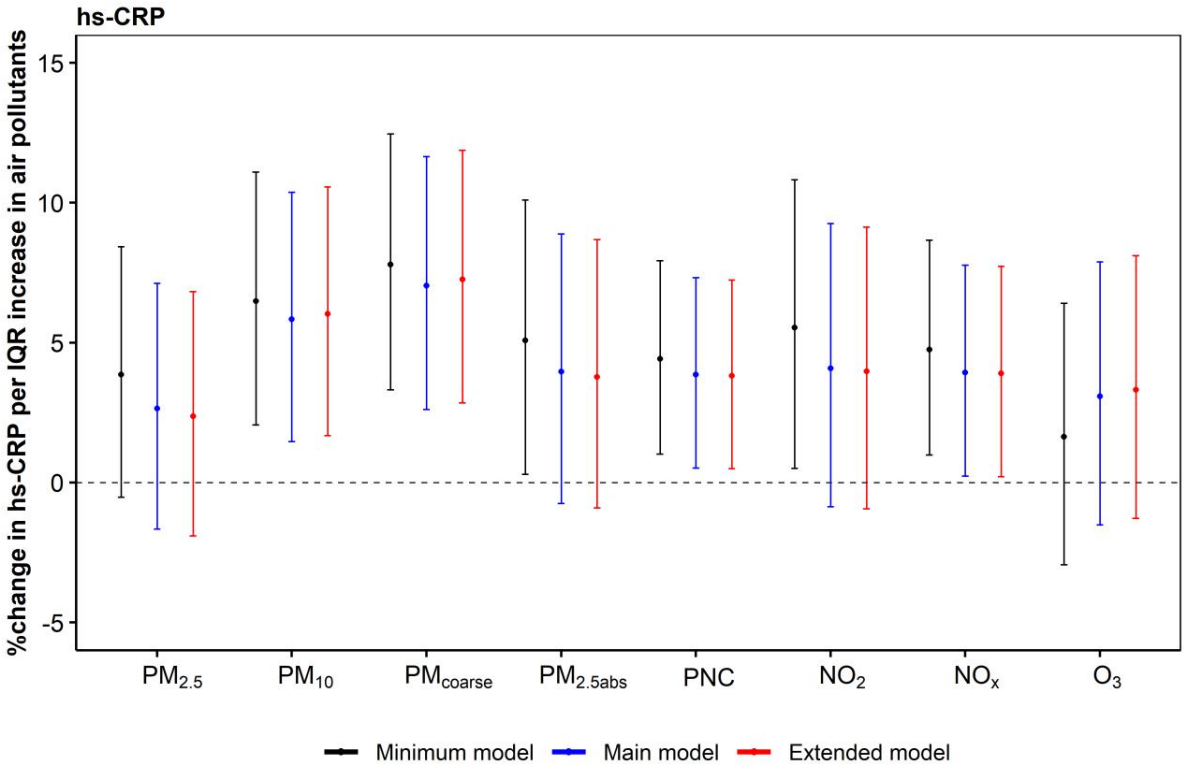


Fig. S2. Comparisons of percent changes (95% CIs) of hs-CRP per IQR increase in air pollutant concentrations between the results from minimum, main and extended models.

Minimum model: minimum models were adjusted for age, sex, BMI, season, an indicator of each visit (KORA S4, KORA F4, or KORA FF4); Main model: further adjusted for educational attainment, smoking status, fasting status, alcohol consumption, physical activity and diet score; Extended model: additionally included diseases status (hypertension, diabetes), medication usages (anti-hypertension, anti-diabetes and lipid lowering medications), HDL, and total cholesterol into the main models. An IQR increase was 1.40 $\mu g/m^3$ for $PM_{2.5}$, 2.06 $\mu g/m^3$ for PM_{10}, 1.36 $\mu g/m^3$ for PM_{coarse}, $1.95\times10^3/cm^3$ for PNC, $0.27\times10^{-5}/m$ for $PM_{2.5abs}$, 6.86 $\mu g/m^3$ for NO_2, 8.69 $\mu g/m^3$ for NO_x, and 3.45 $\mu g/m^3$ for O_3.

$PM_{2.5}$ = particulate matter with an aerodynamic diameter less than or equal to 2.5 μm; PM_{coarse} = particulate matter with an aerodynamic diameter of 2.5-10 μm; PM_{10} = particulate matter with an aerodynamic diameter less than or equal to 10 μm; $PM_{2.5abs}$ = $PM_{2.5}$ absorbance; PNC = particle number concentration; NO_2 = nitrogen dioxide; NO_x = nitrogen oxide; O_3 = ozone.

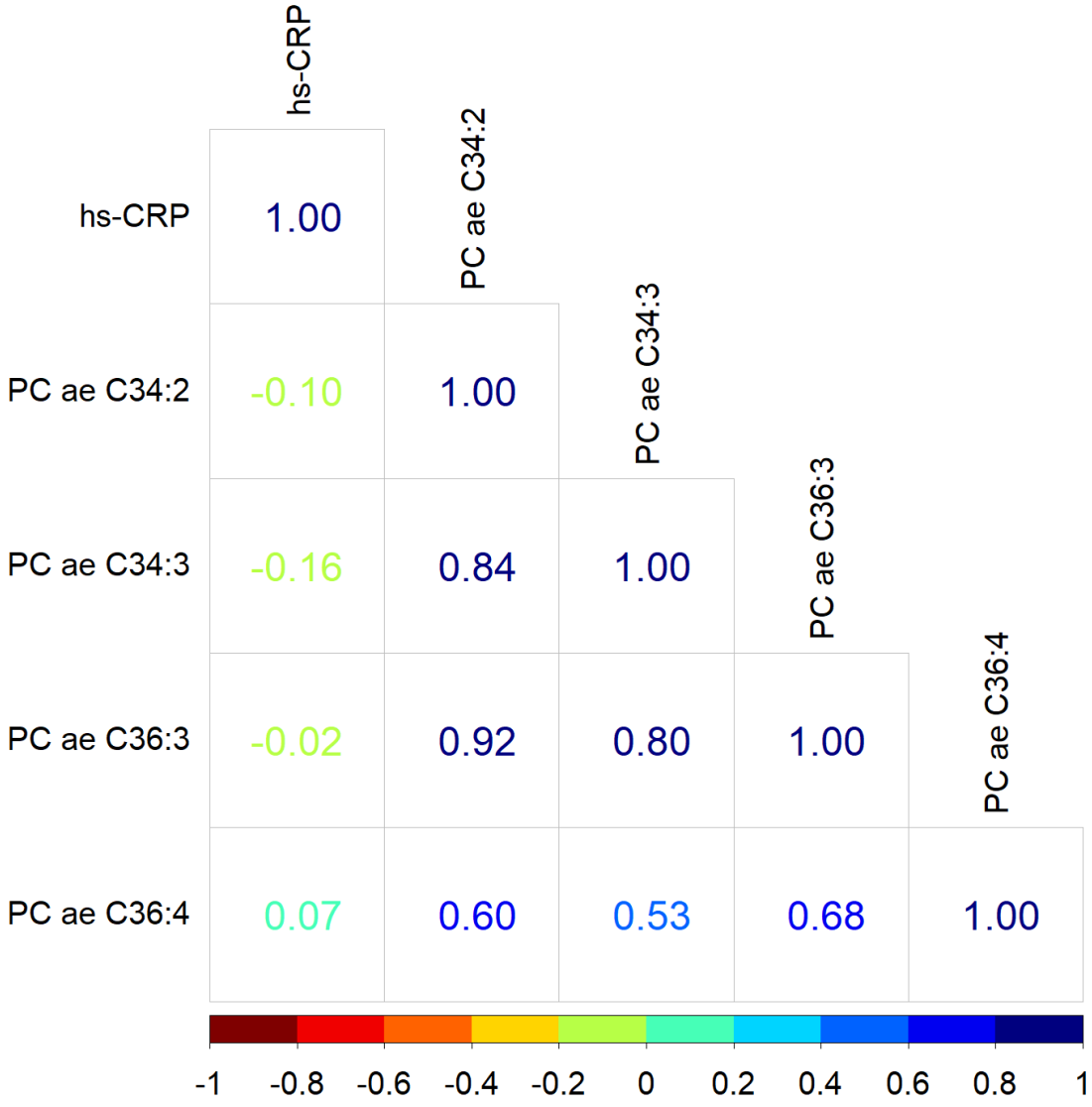


Fig. S3. Spearman correlation coefficients between hs-CRP and PC ae C34:2, PC ae C34:3, PC ae C36:3 and PC ae C36:4.

Table S2 Metabolic pathways identified from pathway analysis that were related to long-term exposure to $PM_{2.5abs}$, PM_{coarse} and NO_2.

	Total	Hits	*p*-value	FDR	Impact factor
Arachidonic acid metabolism	1	1	0.003	0.3	0
Linoleic acid metabolism	1	1	0.008	0.4	0
alpha-Linolenic acid metabolism	1	1	0.02	0.5	0
Glycerophospholipid metabolism	2	1	0.02	0.5	0.1

Total: total number of metabolites in the pathway; Hits: actually matched number from uploaded data; *p*-value: original *p*-value calculated from enrichment analysis; FDR: *p*-value adjusted using False Discovery Rate; Impact factor: pathway impact value calculated from pathway topology analysis.

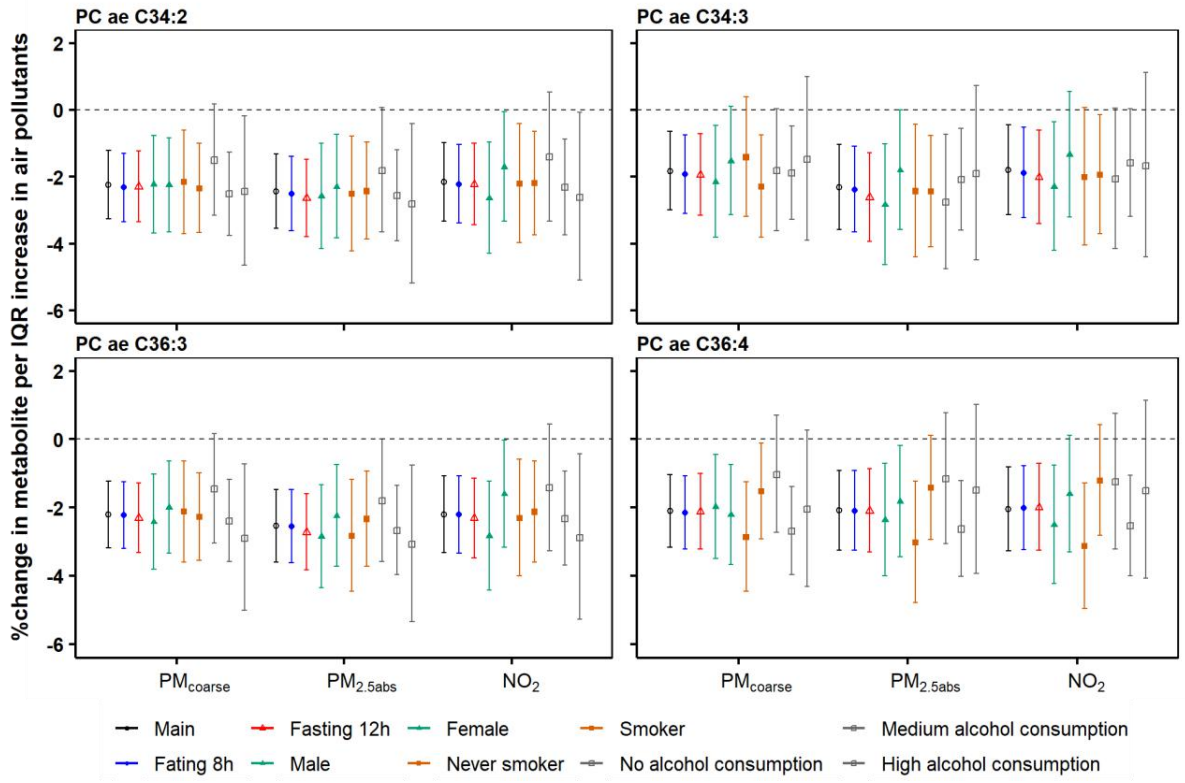


Fig. S4. Percent changes (95% CIs) in metabolite levels per IQR increase in air pollutant concentrations. Effect modification by participants with overnight fasting of 8 or 12 h, sex, smoking, and alcohol consumption.

Results were from our main models adjusted for covariates including age, sex, BMI, season, an indicator of each visit (KORA S4, KORA F4, or KORA FF4), educational attainment, smoking status, fasting status, alcohol consumption, physical activity and diet score, while the continuous variable of alcohol consumption, fasting status, or smoking statues will be replaced by the corresponding effect modifier. An IQR increase was 1.36 μg/m^3 for PM_{coarse}, 0.27×10^{-5}/m for $PM_{2.5abs}$, 6.86 μg/m^3 for NO_2.

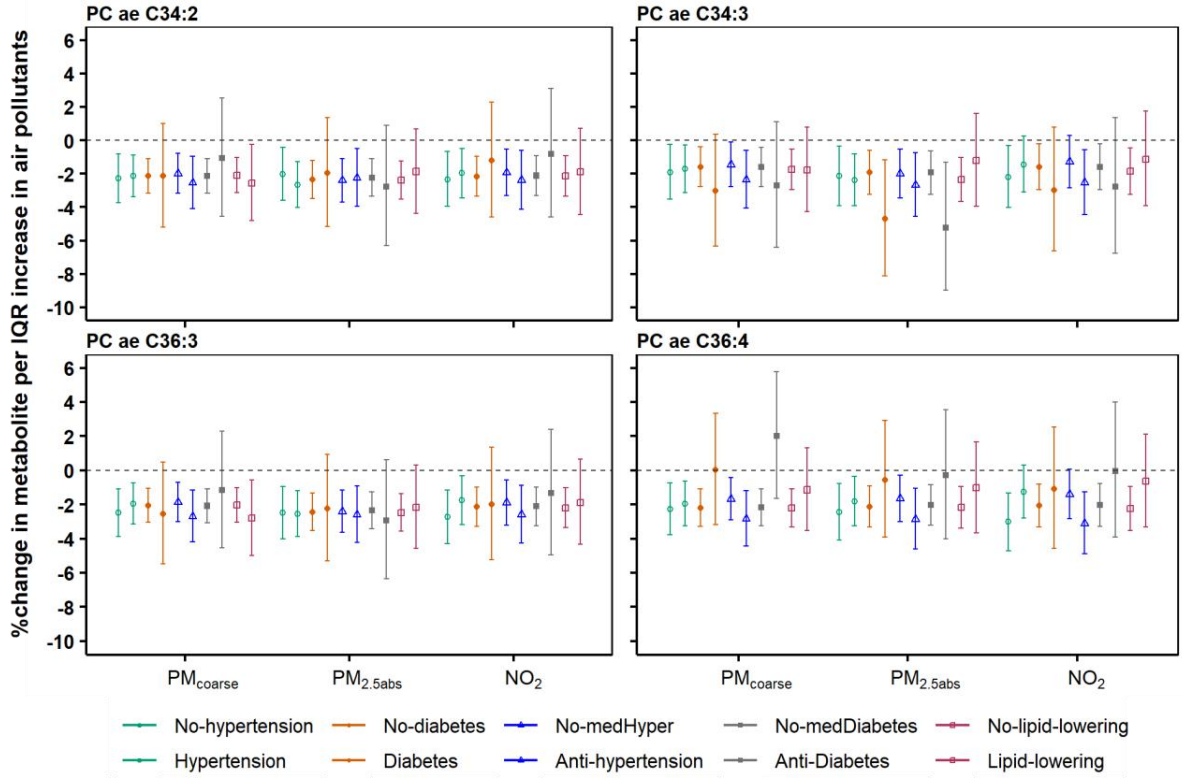


Fig. S5. Percent changes (95% CIs) in metabolite levels per IQR increase in air pollutant concentrations stratified by history of diseases (hypertension, diabetes) and medication intake (anti-hypertensive, anti-diabetic, and lipid-lowering medication).

Results were from our main models adjusted for covariates including age, sex, BMI, season, an indicator of each visit (KORA S4, KORA F4, or KORA FF4), educational attainment, smoking status, fasting status, alcohol consumption, physical activity and diet score. An IQR increase was 1.36 μg/m^3 for PM_{coarse}, 0.27×10^{-5}/m for $PM_{2.5abs}$, 6.86 μg/m^3 for NO_2.

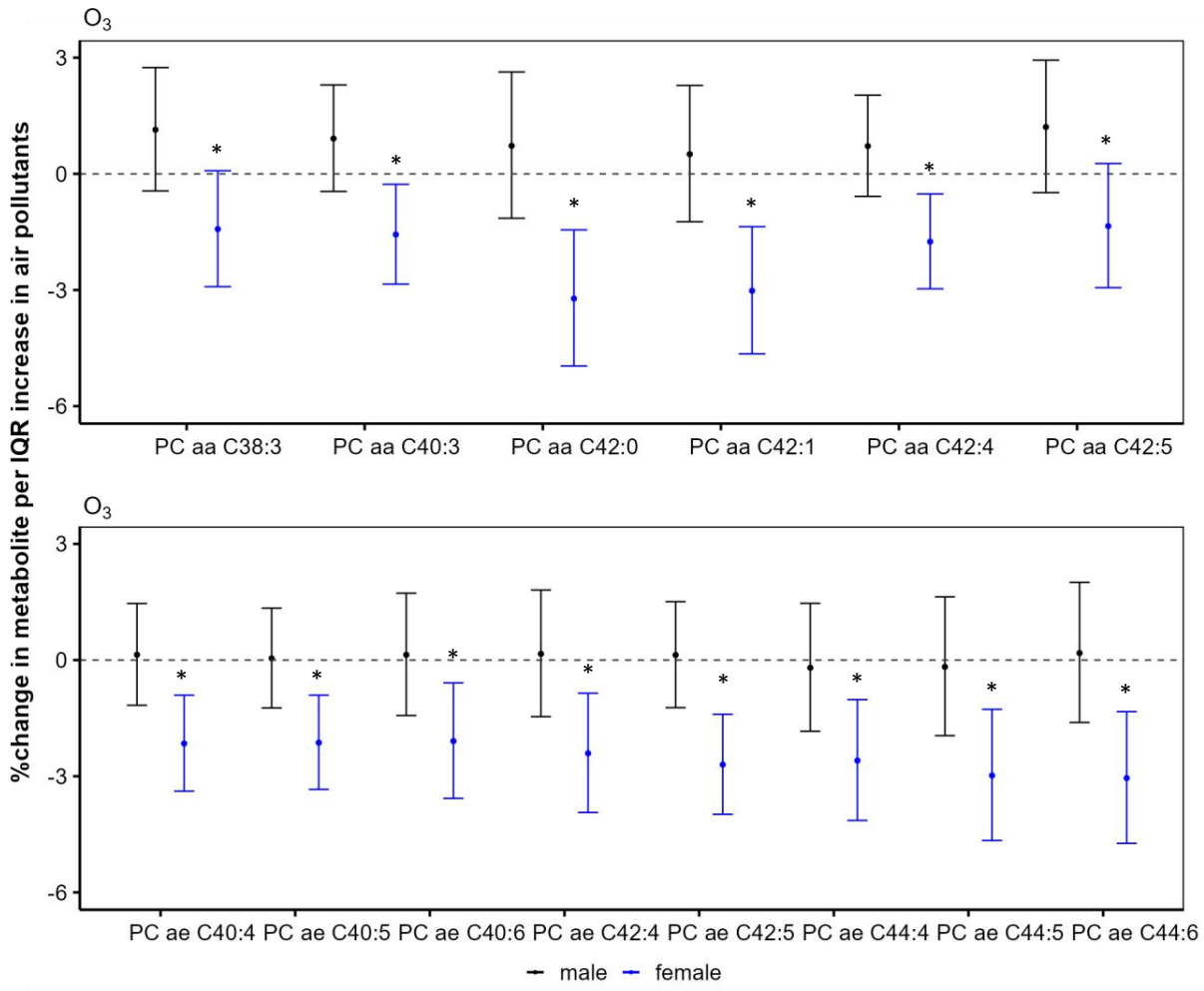


Fig. S6. Percent changes (95% CIs) in metabolite levels per IQR increase in O_3 stratified by sex.

Results were from our main models adjusted for covariates including age, sex, BMI, season, an indicator of each visit (KORA S4, KORA F4, or KORA FF4), educational attainment, smoking status, fasting status, alcohol consumption, physical activity and diet score. An IQR increase was 3.45 μg/m^3 for O_3. PC aa: diacyl phosphatidylcholine; PC ae: acyl-alkyl phosphatidylcholine; O_3 = ozone.

* $p < 0.05$

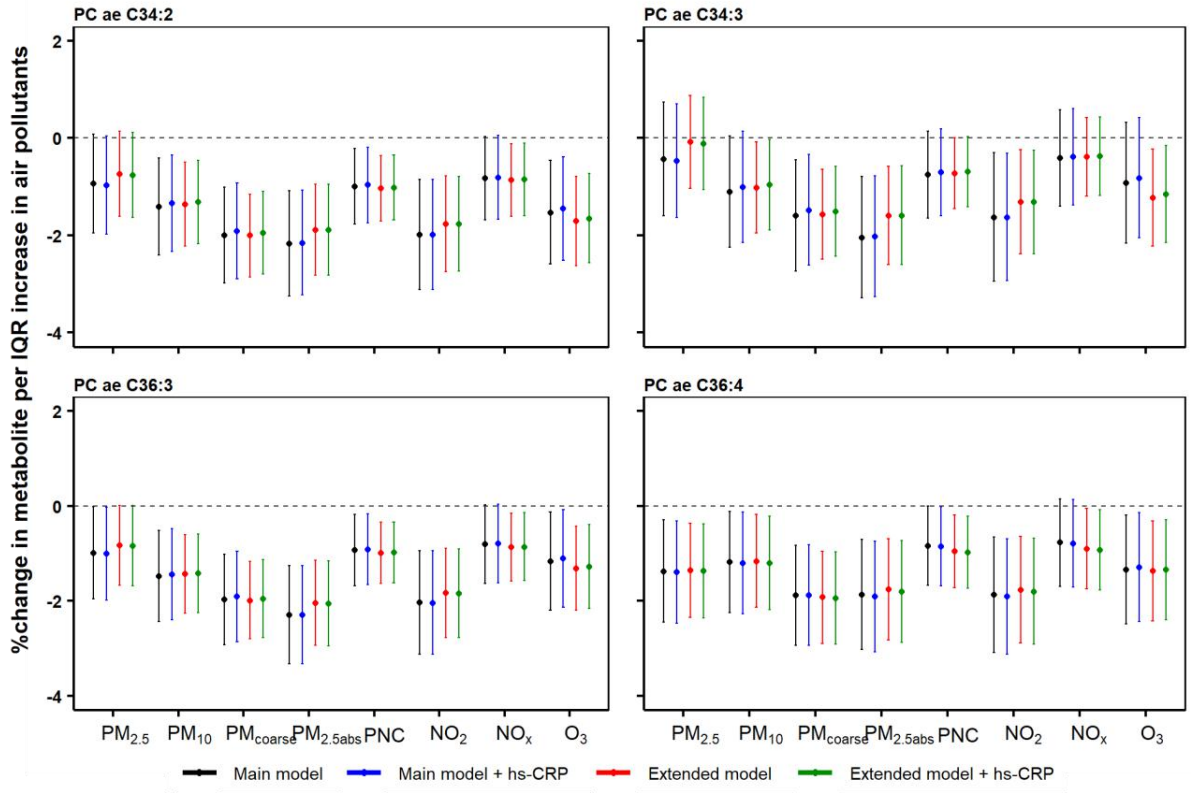


Fig. S7. Comparisons of percent change (95% CI) of metabolite per IQR increase in air pollutant concentrations between the results from main and extended models compared with updated main and extended models of an inclusion of hs-CRP.

Main model: adjusted for age, sex, BMI, season, an indicator of each visit (KORA S4, KORA F4, or KORA FF4), educational attainment, smoking status, fasting status, alcohol consumption, physical activity and diet score; Main model + hs-CRP: hs-CRP and the same covariates in the main model; Extended model: extended model was additionally included diseases status (hypertension, diabetes), and medication usages (anti-hypertension, anti-diabetes and lipid lowering medications), HDL, and total cholesterol into the main models. Extended model + hs-CRP: further adjusted for hs-CRP into the extended models. An IQR increase was 1.40 $\mu g/m^3$ for $PM_{2.5}$, 2.06 $\mu g/m^3$ for PM_{10}, 1.36 $\mu g/m^3$ for PM_{coarse}, 0.27×10^{-5}/m for $PM_{2.5abs}$, $1.95\times10^3/cm^3$ for PNC, 6.86 $\mu g/m^3$ for NO_2, 8.69 $\mu g/m^3$ for NO_x, and 3.45 $\mu g/m^3$ for O_3.

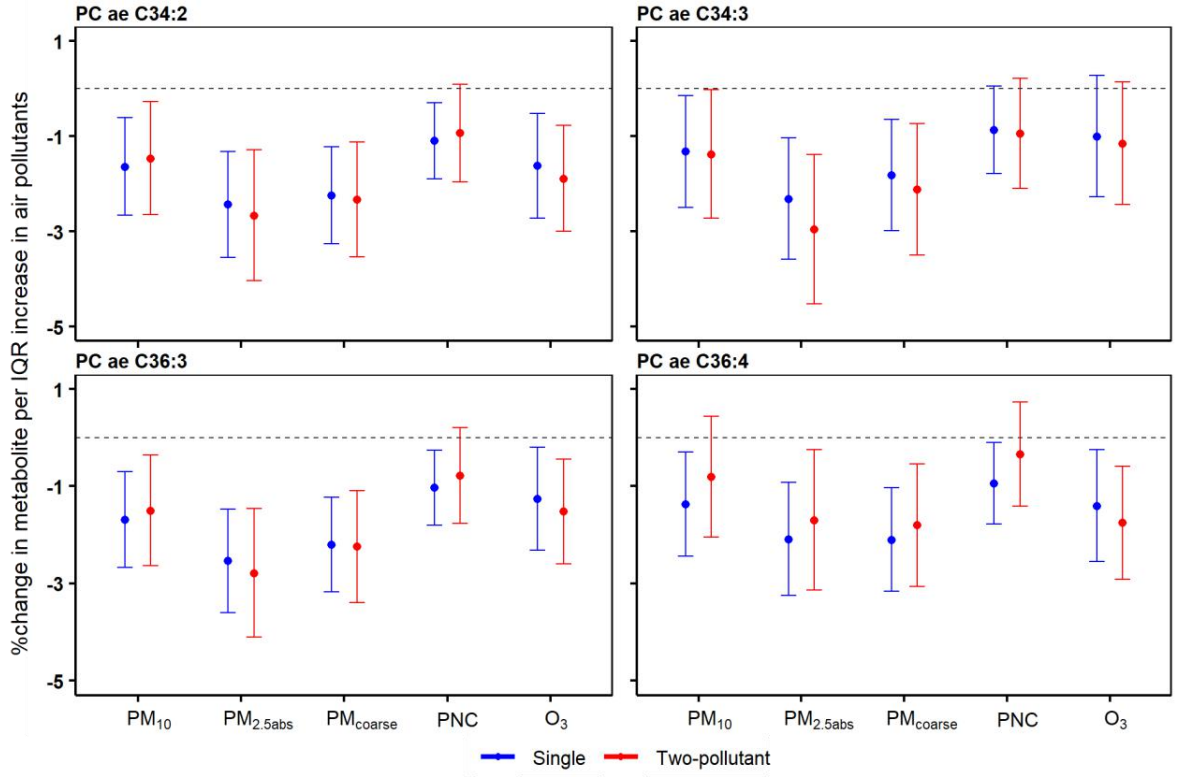


Fig. S8. Comparison between single and two-pollutant models after additional inclusion of $PM_{2.5}$.

Results were from our main models adjusted for covariates including age, sex, BMI, season, an indicator of each visit (KORA S4, KORA F4, or KORA FF4), educational attainment, smoking status, fasting status, alcohol consumption, physical activity and diet score. An IQR increase was 1.40 $\mu g/m^3$ for $PM_{2.5}$, 2.06 $\mu g/m^3$ for PM_{10}, 0.27×10^{-5}/m for $PM_{2.5abs}$, 1.36 $\mu g/m^3$ for PM_{coarse}, $1.95\times10^3/cm^3$ for PNC, and 3.45 $\mu g/m^3$ for O_3.

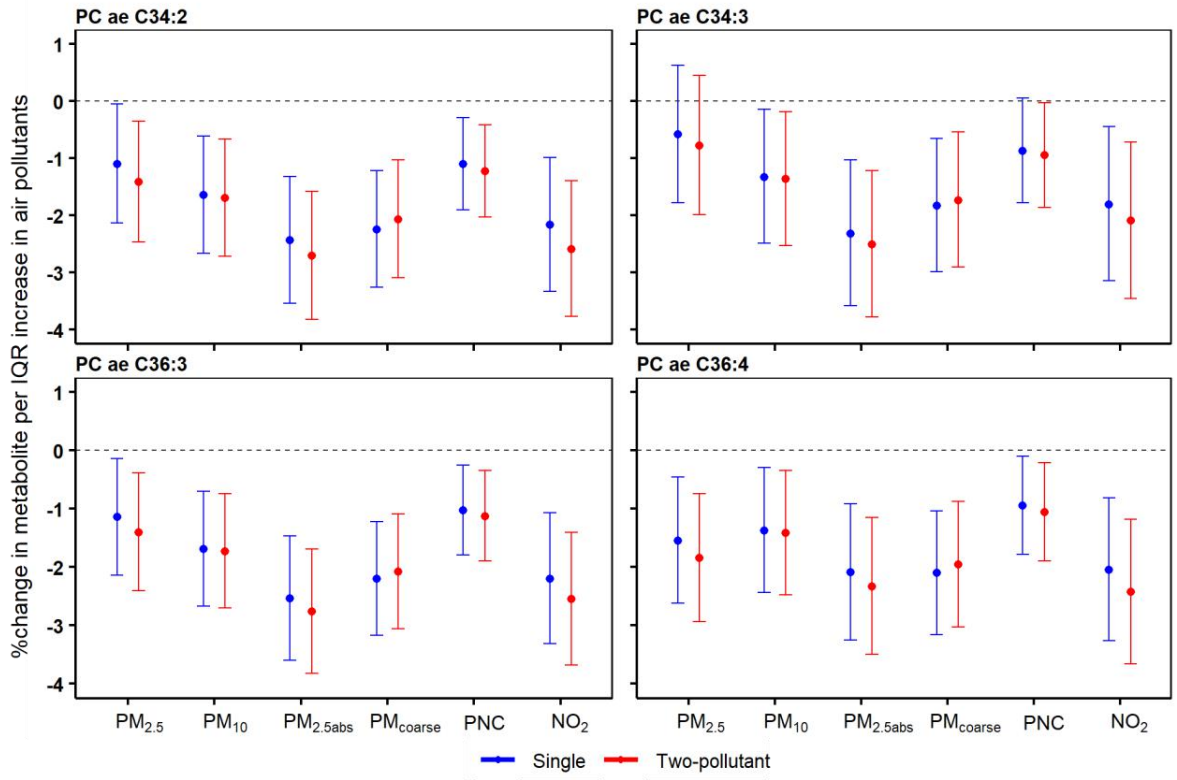


Fig. S9. Comparison between single and two-pollutant models after additional inclusion of O_3.

Models were adjusted for age, sex, BMI, season, an indicator of each visit (KORA S4, KORA F4, or KORA FF4), educational attainment, smoking status, fasting status, alcohol consumption, physical activity and diet score. IQR increases were: 1.40 μg/m^3 for $PM_{2.5}$, 2.06 μg/m^3 for PM_{10}, 1.36 μg/m^3 for PM_{coarse}, 1.95×10^3/cm^3 for PNC, 0.27×10^{-5}/m for $PM_{2.5abs}$, 6.86 μg/m^3 for NO_2.

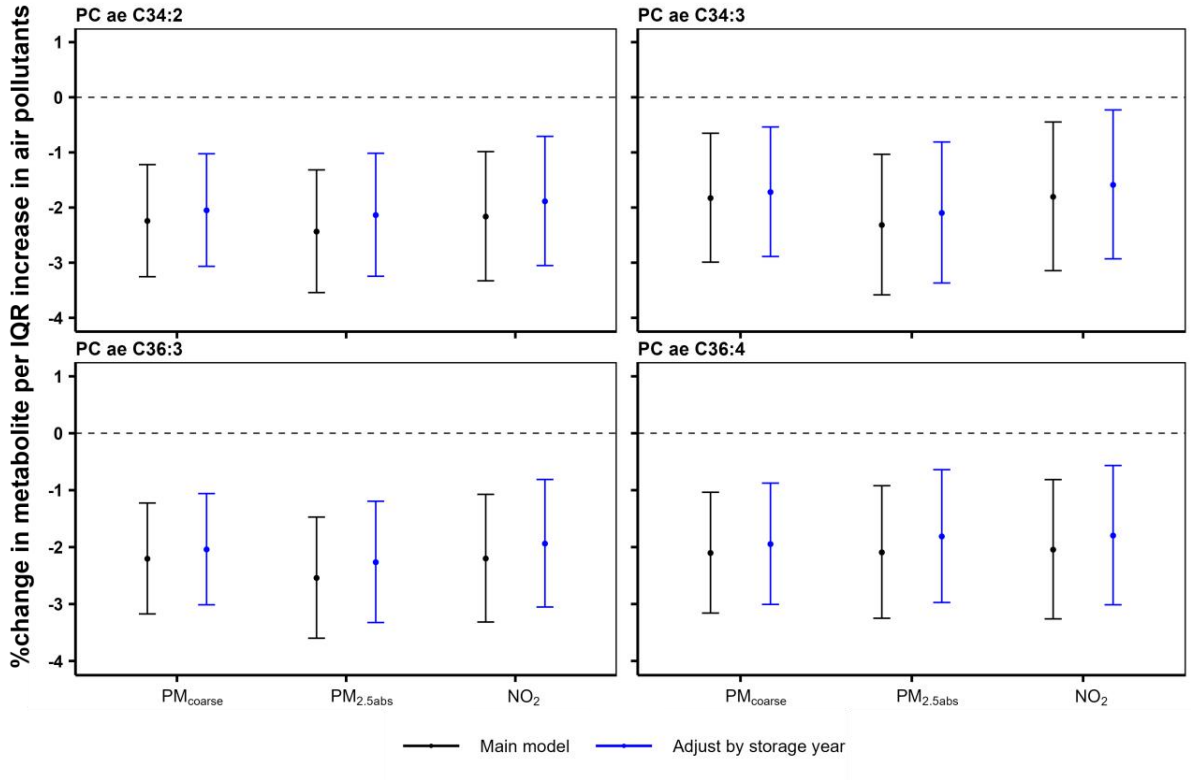


Fig. S10. Comparison between main models and after additional adjusted by storage year of blood samples.

Results were from our main models adjusted for covariates including age, sex, BMI, season, an indicator of each visit (KORA S4, KORA F4, or KORA FF4), educational attainment, smoking status, fasting status, alcohol consumption, physical activity and diet score. An IQR increase was 3.45 μg/m^3 for O_3.

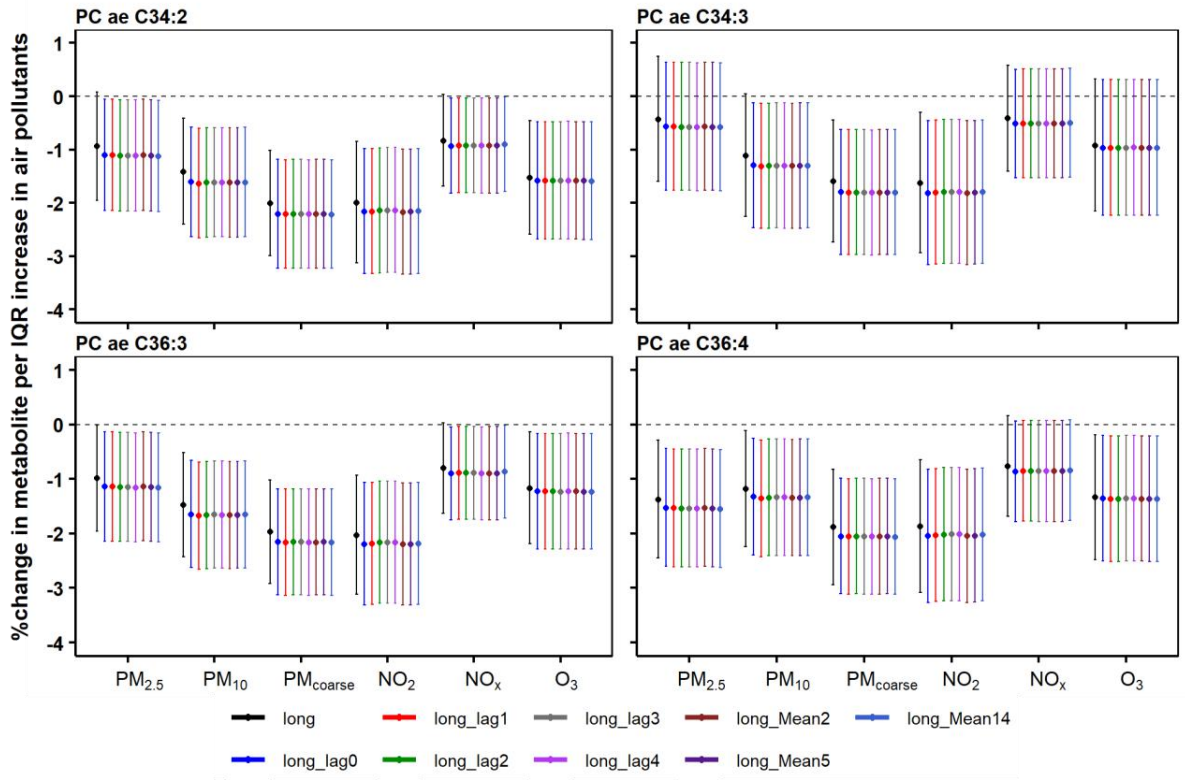


Fig. S11. Inclusion of different exposure window of short-term air pollutants with long-term air pollutants exposure simultaneously.

Long = only long-term exposure; long_lag0 = long-term exposure plus short-term exposure at the day of blood withdrawal; long_lag1 = long-term exposure plus short-term exposure of one day before the blood withdrawal; long_lag2 = long-term exposure plus short-term exposure of two days before the blood withdrawal, long_lag3 = long-term exposure plus short-term exposure of three days before the blood withdrawal, long_lag4 = long-term exposure plus short-term exposure of four days before the blood withdrawal; long_Mean2 = long-term exposure plus short-term exposure of 2-day moving average before the blood withdrawal; long_Mean5 = long-term exposure plus short-term exposure of 5-day moving average before the blood withdrawal, long_Mean14 = long-term exposure plus short-term exposure of 2-week moving average before the blood withdrawal. Results were from our main models adjusted for covariates including age, sex, BMI, season, an indicator of each visit (KORA S4, KORA F4, or KORA FF4), educational attainment, smoking status, fasting status, alcohol consumption, physical activity and diet score. An IQR increase was 1.40 $\mu g/m^3$ for $PM_{2.5}$, 2.06 $\mu g/m^3$ for PM_{10}, 1.36 $\mu g/m^3$ for PM_{coarse}, 6.86 $\mu g/m^3$ for NO_2, 8.69 $\mu g/m^3$ for NO_x, and 3.45 $\mu g/m^3$ for O_3.

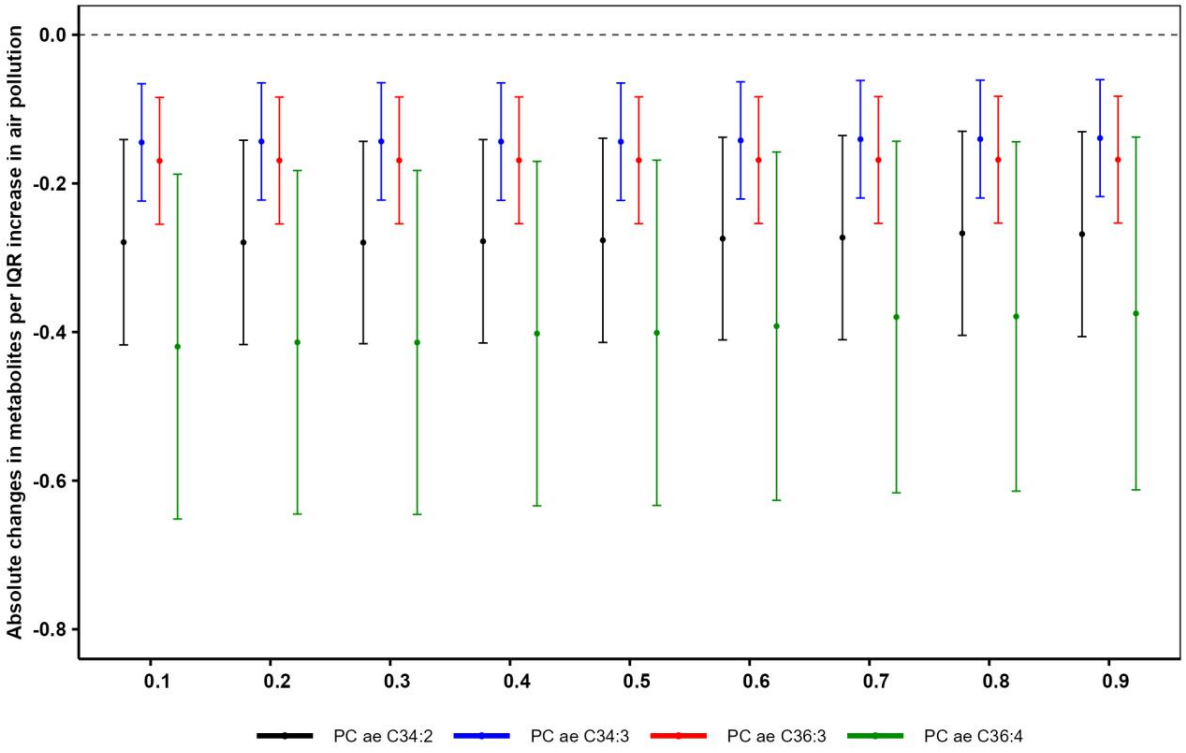


Fig. S12. Absolute changes of metabolites at deciles of the distribution per IQR increase in concentrations of $PM_{2.5abs}$ exposure.

Linear quantile mixed models were used by adjusting for age, sex, BMI, season, an indicator of each visit (KORA S4, KORA F4, or KORA FF4), educational attainment, smoking status, fasting status, alcohol consumption, physical activity and diet score. An IQR increase was 0.27×10^{-5}/m for $PM_{2.5abs}$.

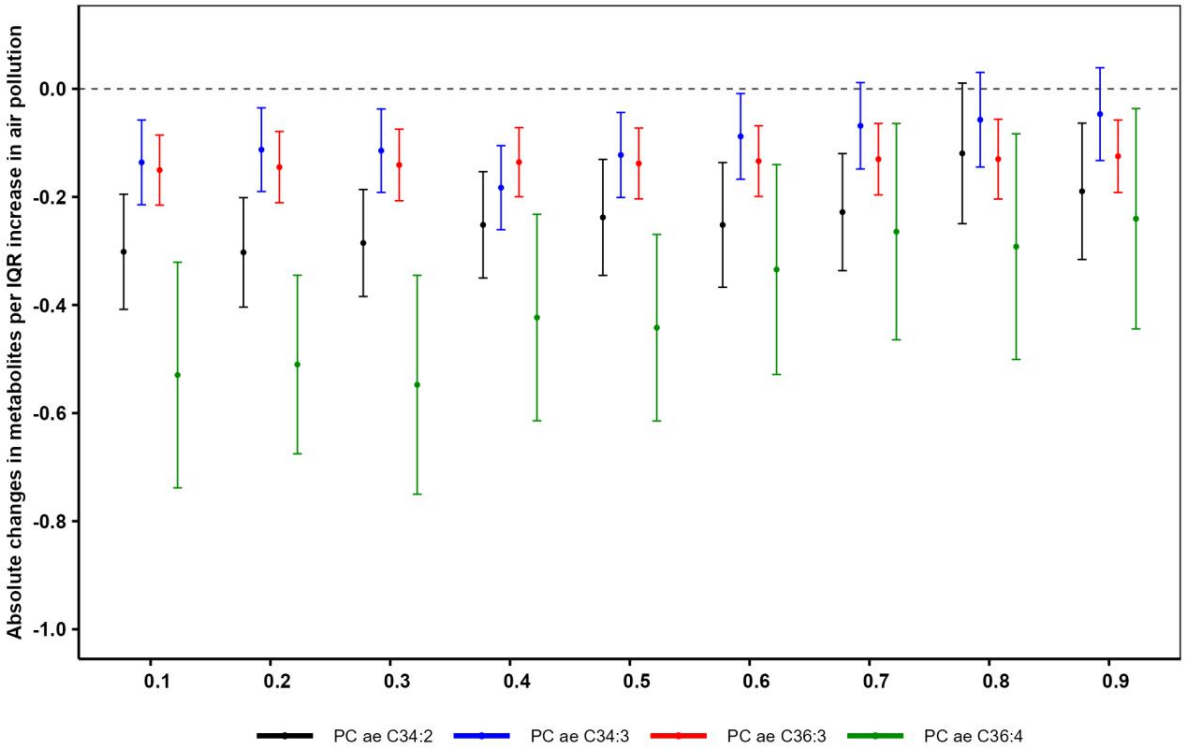


Fig. S13. Absolute changes of metabolites at deciles of the distribution per IQR increase in concentrations of PM_{coarse} exposure.

Linear quantile mixed models were used by adjusting for age, sex, BMI, season, an indicator of each visit (KORA S4, KORA F4, or KORA FF4), educational attainment, smoking status, fasting status, alcohol consumption, physical activity and diet score. An IQR increase was 1.36 $\mu g/m^3$ for PM_{coarse}.

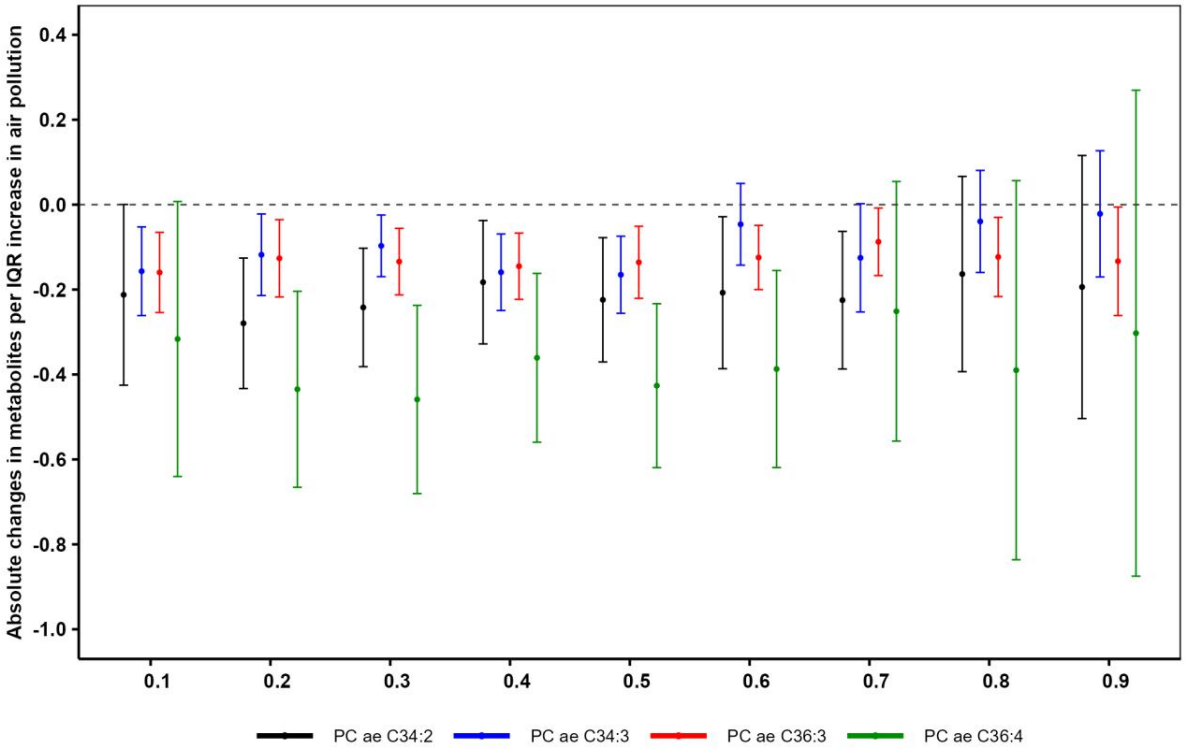


Fig. S14. Absolute changes of metabolites at deciles of the distribution per IQR increase in concentrations of NO_2 exposure.

Linear quantile mixed models were used by adjusting for age, sex, BMI, season, an indicator of each visit (KORA S4, KORA F4, or KORA FF4), educational attainment, smoking status, fasting status, alcohol consumption, physical activity and diet score. An IQR increase was 6.86 $\mu g/m^3$ for NO_2.

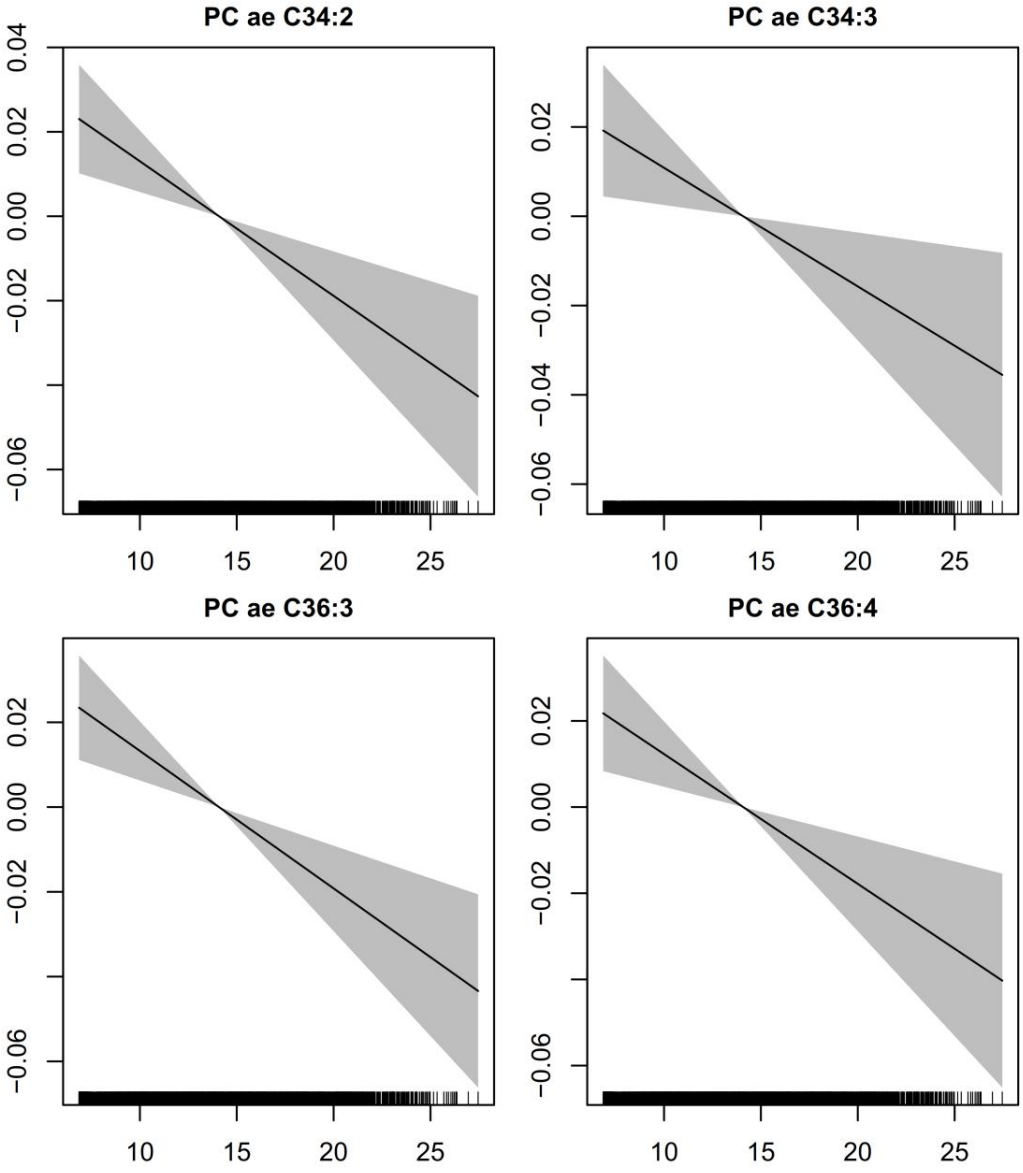


Fig. S15. Exposure-response relationships between NO_2 and PC ae C34:2, PC ae C34:3, PC ae C36:3 and PC ae C36:4.

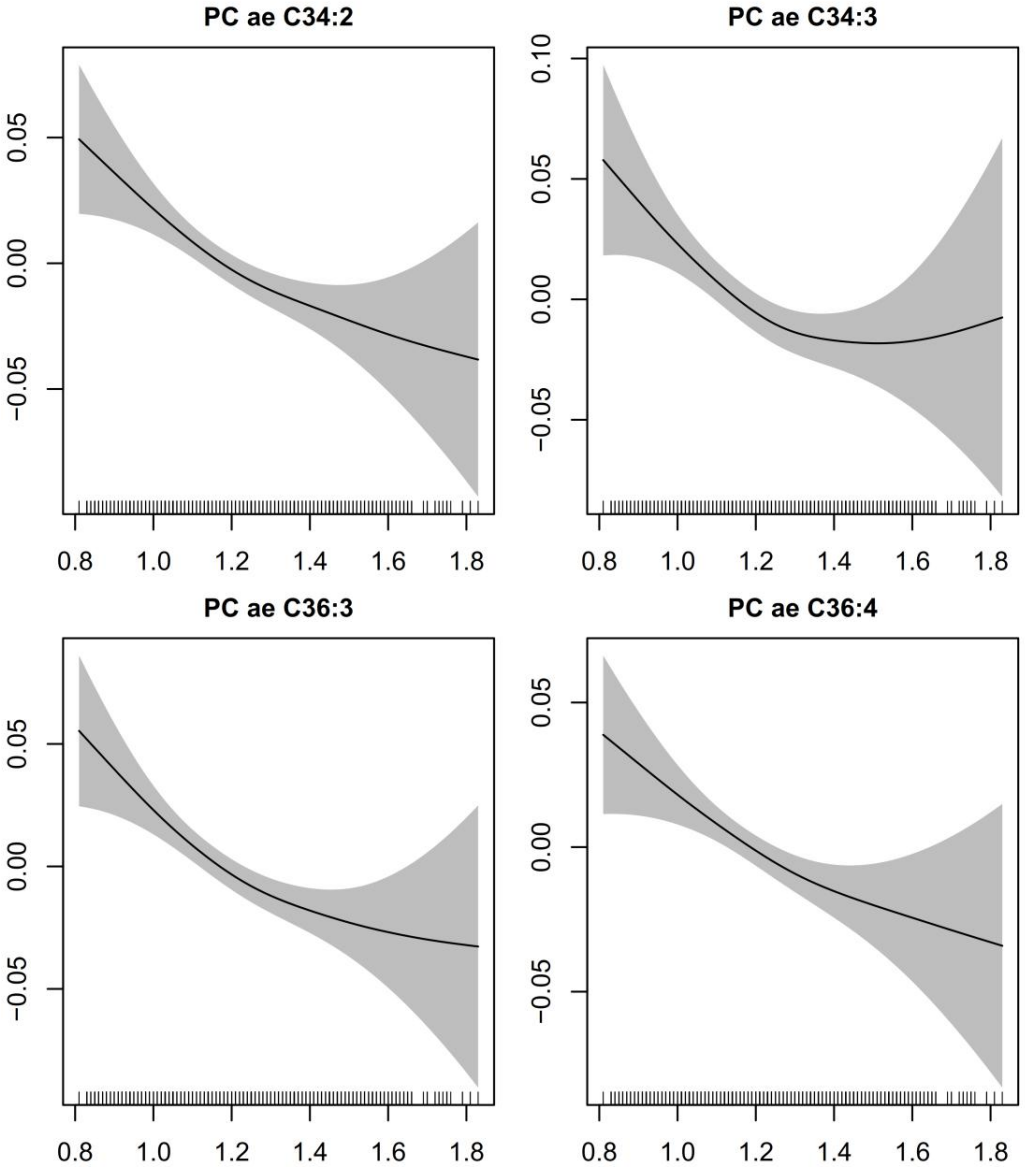


Fig. S16. Exposure-response relationships between $PM_{2.5abs}$ and PC ae C34:2, PC ae C34:3, PC ae C36:3 and PC ae C36:4.

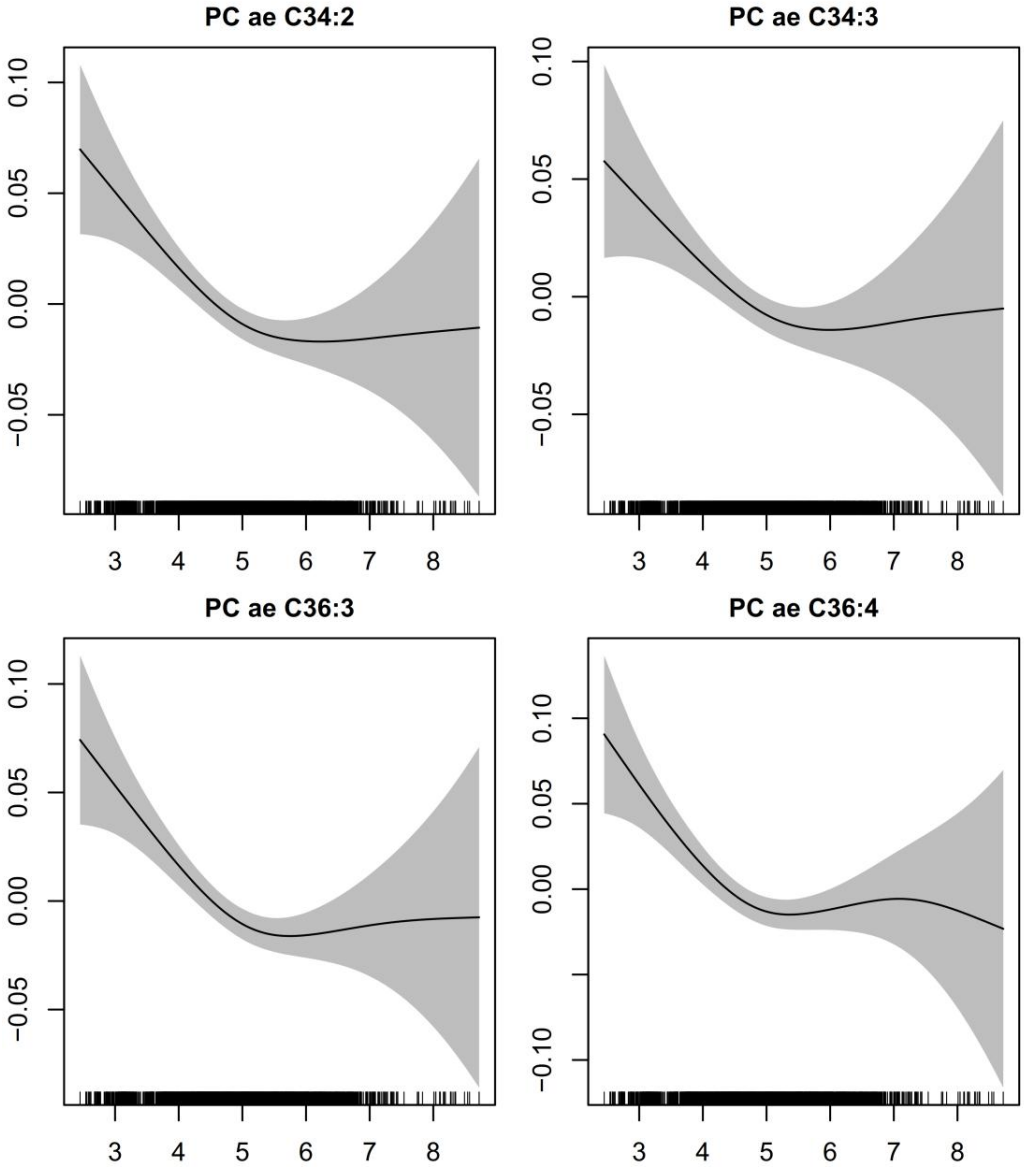


Fig. S17. Exposure-response relationships between PM_{coarse} and PC ae C34:2, PC ae C34:3, PC ae C36:3 and PC ae C36:4.

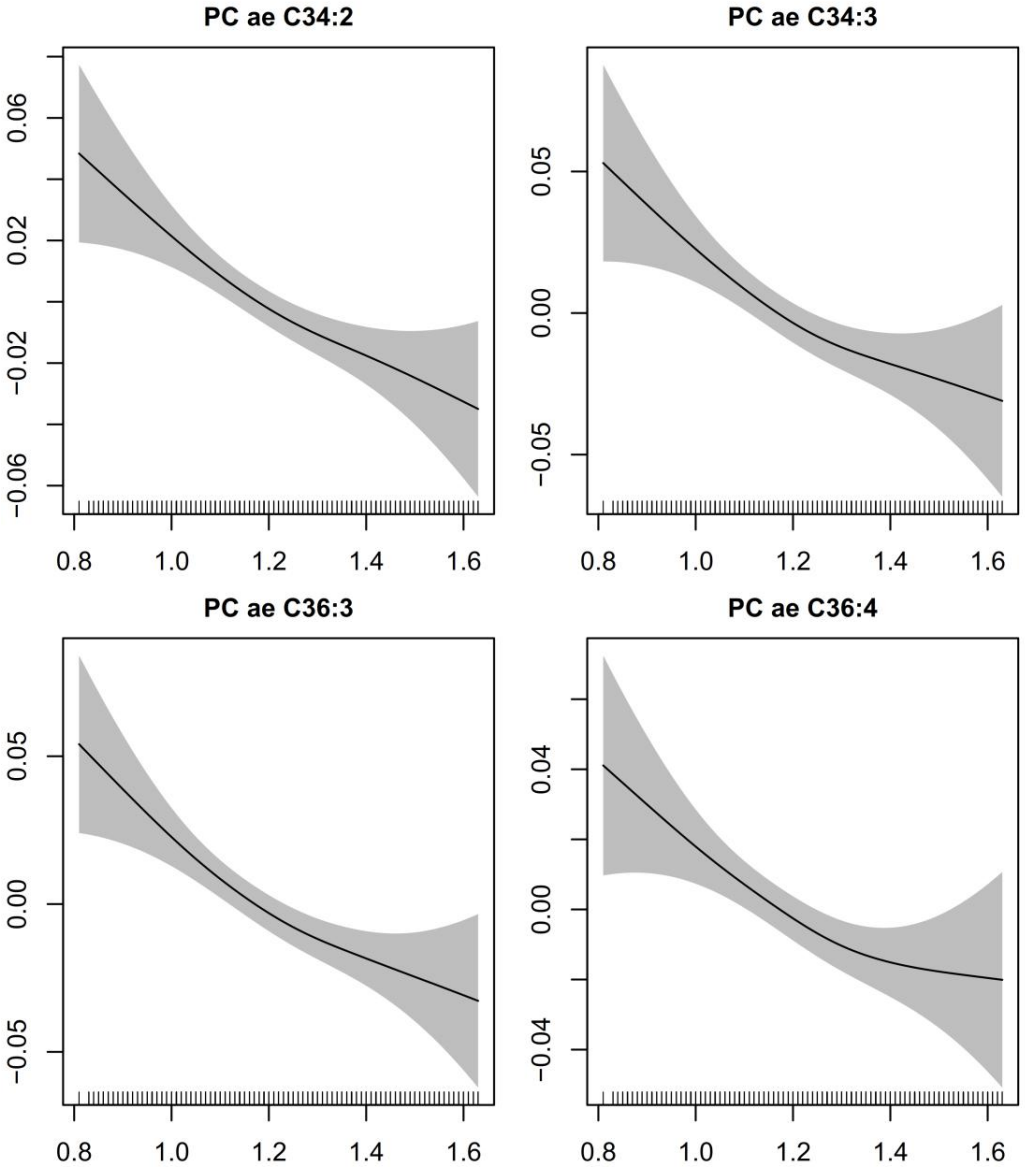


Fig. S18. Exposure-response relationships between $PM_{2.5abs}$ and PC ae C34:2, PC ae C34:3, PC ae C36:3 and PC ae C36:4 after excluding $PM_{2.5abs}$ higher than 99% of $PM_{2.5abs}$ values.

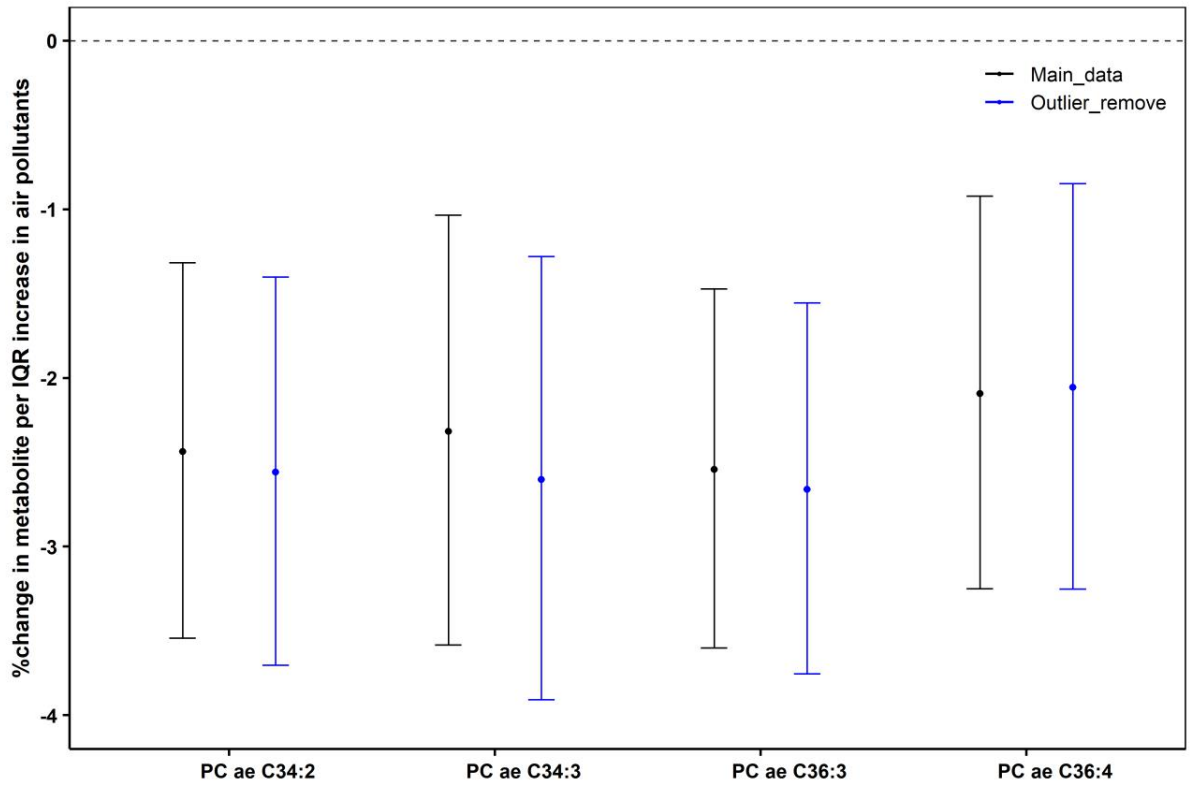


Fig. S19. Comparisons between main data (n =5772) and the outlier removed data (n = 5711) of $PM_{2.5abs}$.

Models were adjusted for age, sex, BMI, season, an indicator of each visit (KORA S4, KORA F4, or KORA FF4), educational attainment, smoking status, fasting status, alcohol consumption, physical activity and diet score. An IQR increase of $PM_{2.5abs}$ didn't change (0.27×10^{-5}/m).

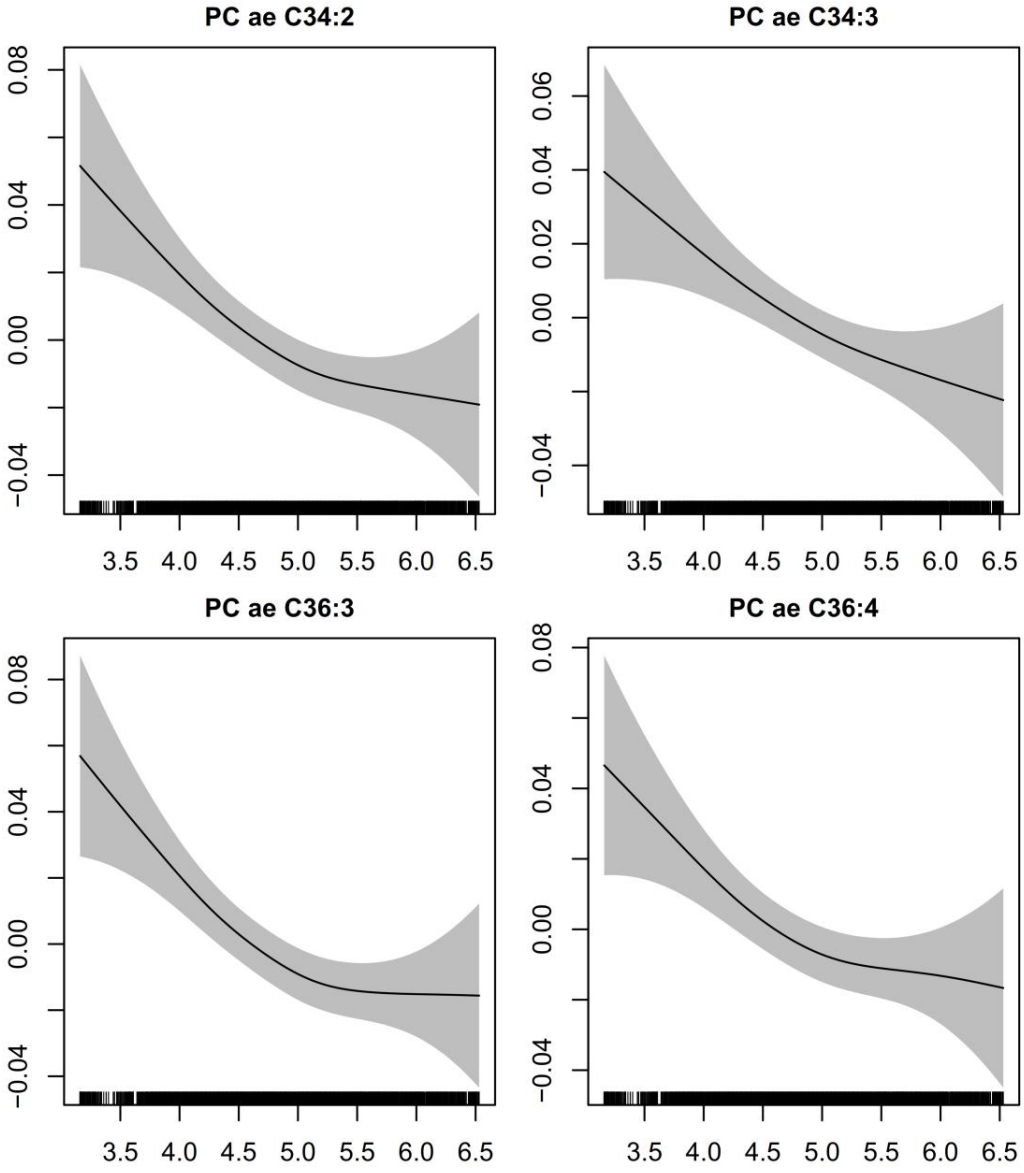


Fig. S20. Exposure-response relationships between PM_{coarse} and PC ae C34:2, PC ae C34:3, PC ae C36:6 and PC ae C36:4 with excluding PM_{coarse} concentration < 5% of total PM_{coarse} and PM_{coarse} concentration > 95% of total PM_{coarse}.

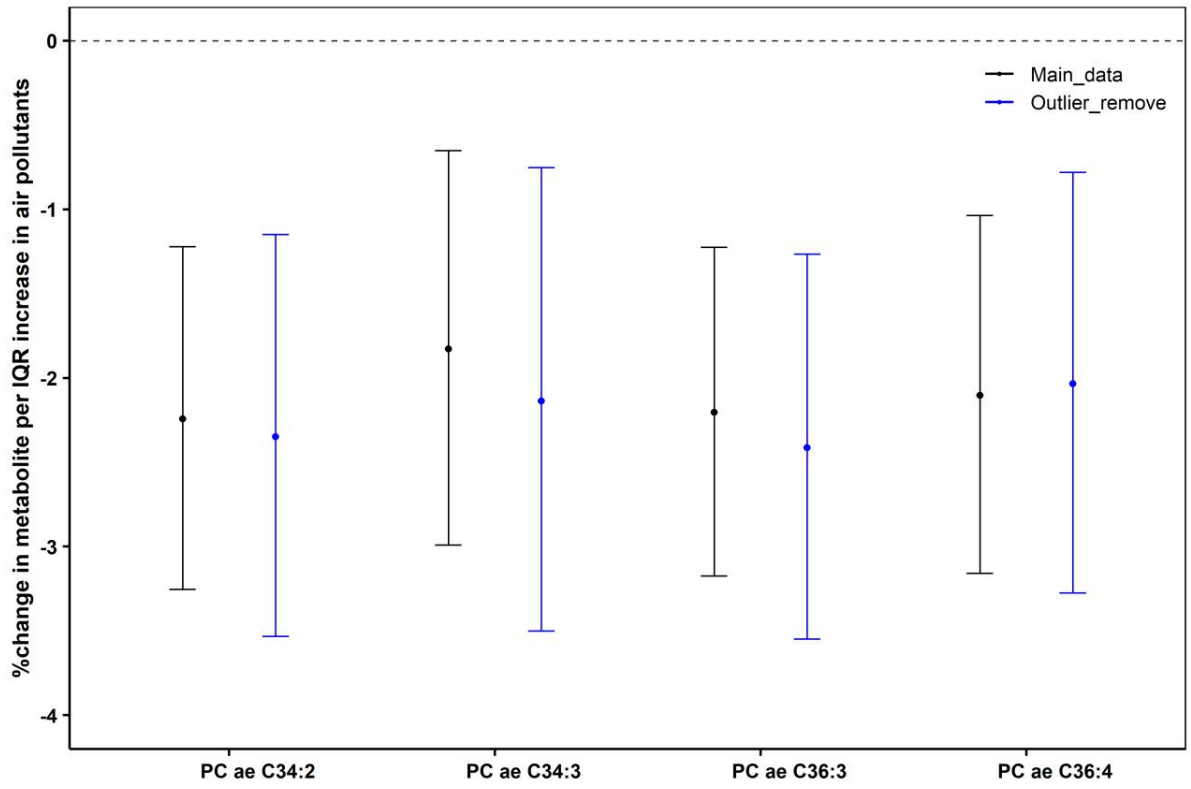


Fig. S21. Comparisons between main data (n =5772) and outliers of PM_{coarse} removed (n = 5203).

Models were adjusted for age, sex, BMI, season, an indicator of each visit (KORA S4, KORA F4, or KORA FF4), educational attainment, smoking status, fasting status, alcohol consumption, physical activity and diet score. The IQR of PM_{coarse} changed from 1.36 $\mu g/m^3$ (main data) to 1.21 $\mu g/m^3$ (outliers removed).

Reference:

Zhang S, Mwiberi S, Pickford R, Breitner S, Huth C, et al. 2021. Longitudinal associations between ambient air pollution and insulin sensitivity: results from the KORA cohort study. The Lancet Planetary Health 5(1): e39-e49. https://doi.org/10.1016/s2542-5196(20)30275-8.

Paper II

Title:	Longitudinal associations between metabolites and immediate, short- and medium-term exposure to ambient air pollution: Results from the KORA cohort study
Authors:	Yueli Yao, Alexandra Schneider, Kathrin Wolf, Siqi Zhang, Rui Wang-Sattler, Cornelia Prehn, Jerzy Adamski, Annette Peters, Susanne Breitner
Status:	Published
Journal:	*Science of the Total Environment*
Year:	2023
Volume:	900
DOI:	https://doi.org/10.1016/j.scitotenv.2023.165780
Supplements:	https://www.sciencedirect.com/science/article/pii/S0048969723044030?via%3Dihub#s0090
Impact factor:	8.2 (Journal Citation Reports®, year 2023)
Rank:	31/358 in Category Environmental Sciences Journals (Journal Citation Reports®, year 2023)

Science of the Total Environment 900 (2023) 165780

Contents lists available at ScienceDirect

Science of the Total Environment

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

Longitudinal associations between metabolites and immediate, short- and medium-term exposure to ambient air pollution: Results from the KORA cohort study

Yueli Yao [a,b,*], Alexandra Schneider [a], Kathrin Wolf [a], Siqi Zhang [a], Rui Wang-Sattler [a,c], Annette Peters [a,b,c,d,1], Susanne Breitner [a,b,1]

[a] Institute of Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany
[b] Institute for Medical Information Processing, Biometry, and Epidemiology (IBE), Faculty of Medicine, Pettenkofer School of Public Health, LMU Munich, Munich, Germany
[c] German Center for Diabetes Research (DZD), Munich-Neuherberg, Germany
[d] German Centre for Cardiovascular Research (DZHK), Partner Site Munich, Munich, Germany

ARTICLE INFO

Editor: Lidia Minguez Alarcon

Keywords:
Air pollution
Nitrogen dioxides
Targeted metabolomics
Phosphatidylcholine
Metabolic pathway

ABSTRACT

Background: Short-term exposure to air pollution has been reported to be associated with cardiopulmonary diseases, but the underlying mechanisms remain unclear. This study aimed to investigate changes in serum metabolites associated with immediate, short- and medium-term exposures to ambient air pollution.
Methods: We used data from the German population-based Cooperative Health Research in the Region of Augsburg (KORA) S4 survey (1999–2001) and two follow-up examinations (F4: 2006–08 and FF4: 2013–14). Mass-spectrometry-based targeted metabolomics was used to quantify metabolites among serum samples. Only participants with repeated metabolites measurements were included in this analysis. We collected daily averages of fine particles ($PM_{2.5}$), coarse particles (PM_{coarse}), nitrogen dioxide (NO_2), and ozone (O_3) at urban background monitors located in Augsburg, Germany. Covariate-adjusted generalized additive mixed-effects models were used to examine the associations between immediate (2-day average of same day and previous day as individual's blood withdrawal), short- (2-week moving average), and medium-term exposures (8-week moving average) to air pollution and metabolites. We further performed pathway analysis for the metabolites significantly associated with air pollutants in each exposure window.
Results: Of 9,620 observations from 4,261 study participants, we included 5,772 (60.0%) observations from 2,583 (60.6%) participants in this analysis. Out of 108 metabolites that passed quality control, multiple significant associations between metabolites and air pollutants with several exposure windows were identified at a Bonferroni corrected *p*-value threshold ($p < 3.9 \times 10^{-5}$). We found the highest number of associations for NO_2, particularly at the medium-term exposure windows. Among the identified metabolic pathways based on the metabolites significantly associated with air pollutants, the glycerophospholipid metabolism was the most robust pathway in different air pollutants exposures.
Conclusions: Our study suggested that short- and medium-term exposure to air pollution might induce alterations of serum metabolites, particularly in metabolites involved in metabolic pathways related to inflammatory response and oxidative stress.

1. Introduction

Short- and medium-term exposures to ambient air pollution could lead to adverse health effects such as reduced lung function, higher blood pressure, and cognitive dysfunction, and further increase the risk for pulmonary, cardiovascular, and neurological diseases (Brook and

* Corresponding author at: Institute of Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, D-85764 Neuherberg, Germany.
E-mail address: yueli.yao@helmholtz-munich.de (Y. Yao).
[1] Authors share last authorship

https://doi.org/10.1016/j.scitotenv.2023.165780
Received 5 May 2023; Received in revised form 21 July 2023; Accepted 23 July 2023
Available online 24 July 2023

Rajagopalan, 2009; Brook et al., 2010; Franchini and Mannucci, 2007; Lee et al., 2017; Pope 3rd et al., 2006; Rice et al., 2013; Wang et al., 2019; Wu et al., 2016; Zhang et al., 2022). Oxidative stress and systemic inflammation have been shown to be among the underlying biological mechanisms mediating the adverse health effects of air pollution (Dauchet et al., 2018; Li et al., 2020).

Ambient air pollutants could induce local inflammatory responses, further leading to indirect systemic inflammation (ambient accumulation-mode particles) or enter into the blood and lead to alterations of blood biomarkers directly (ambient ultrafine particles) (Brook et al., 2010). The blood metabolome has been defined as the collection of biologically active chemicals in human blood derived from endogenous processes and exogenous exposure to food, medicines, and pollutants (Rappaport et al., 2014). Furthermore, metabolomics has become a well-developed tool to identify smaller molecular metabolites in the biological systems and the corresponding cellular responses perturbed by endogenous or exogenous stimuli (Holmes et al., 2008). Therefore, examining serum metabolite alterations and investigating the involved metabolic pathways is a novel approach to better understand the underlying biological mechanisms of air pollution-associated diseases.

Some studies have reported associations between short- and/or medium-term air pollution exposure and the alteration of metabolites in the blood (Breitner et al., 2016; Li et al., 2017; Liang et al., 2018; Nassan et al., 2021a; Nassan et al., 2021b; Surowiec et al., 2016; van Veldhoven et al., 2019; Vlaanderen et al., 2017; Ward-Caviness et al., 2016; Zhu et al., 2021). However, only a few were conducted in a cohort setting (Breitner et al., 2016; Nassan et al., 2021a; Nassan et al., 2021b; Ward-Caviness et al., 2016). Furthermore, air pollution-related changes in metabolites or metabolic pathways observed in these studies were not consistent due to the complexity of metabolites and differences in techniques (targeted metabolomics *vs.* untargeted metabolomics), air pollution exposures (*e.g.*, different foci on single air pollutants or multiple air pollutants and exposure windows), and targeted populations (limited to older men, pregnant women, or teenagers). The inconsistency led to the limited generalizability of the studies' findings to other populations. In addition, most of these studies were based on a cross-sectional design, which could not provide evidence of a potential temporal relationship between air pollution exposures and metabolites.

Given the above gaps, we aimed to determine the associations between immediate, short- and medium-term exposures to air pollution and targeted metabolomics within a longitudinal analysis of the German population-based Cooperative Health Research of Augsburg (KORA) cohort. Moreover, we have previously shown that long-term exposure to air pollution was associated with metabolic alterations in this sample, particularly in PCs with unsaturated long-chain fatty acids (Yao et al., 2022). We hypothesized that immediate, short- and medium-term exposures to air pollution are associated with the perturbation of serum metabolite levels involved in metabolic pathways related to adverse health effects from ambient air pollution, such as inflammatory response and oxidative stress.

2. Materials and methods

2.1. Study design and participants

This longitudinal analysis is based on the same data set as our previous analysis investigating changes in serum metabolites associated with long-term exposure to air pollution (Yao et al., 2022). In brief, the KORA (Cooperative Health Research in the Region Augsburg, Germany) cohort is a regional research platform for population-based surveys and subsequent follow-up studies in the fields of epidemiology, health economics, and health care research (https://www.helmholtz-munich.de/en/epi/cohort/kora). Individuals residing in the city of Augsburg and neighbouring administrative districts were recruited covering urban and rural areas, both men and women and a broad age range to represent the general population. The inclusion criteria for the study were German nationality and first place of residence in the study region. More details can be found elsewhere (Holle et al., 2005; Löwel et al., 2005). The fourth cross-sectional health survey of the KORA cohort (KORA S4) was conducted from October 1999 to April 2001 and involved 4,261 participants aged 25–74 years with German citizenship. Two follow-up examinations were carried out: within the first follow-up (KORA F4), 3,080 participants were examined between October 2006 and May 2008; the second follow-up (KORA FF4) consisted of 2,279 participants with examinations between June 2013 and September 2014.

A computer-assisted personal interview, a self-administered questionnaire, and physical examinations were performed at each visit by trained investigators at the study center. Educational attainment was categorized into primary school, high school, and college. The continuous body-mass index (BMI) was categorized into normal weight ($\leq$30 kg/m^2) and obesity (>30 kg/m^2). Physical activity was categorized based on the time spent on physical exercise into low (no or almost no physical exercise), medium (regular or irregular, approximately 1 h/week), and high (more than 2 h/week) levels. Alcohol consumption was categorized into no (0 g/day), moderate (men 0.1–39.9 g/day and women 0.1–19.9 g/day), and high (men $\geq$40 g/day and women $\geq$20 g/day) consumption. Smoking status was categorized into smoker (regular or irregular smoker), former smoker, and never smoker. The individuals' dietary intake was collected using a food-frequency questionnaire investigating 24 food groups. An index was built to rate the frequency with which each food was consumed by assigning either 0, 1, or 2 points based on the German Nutrition Society (DGE) recommendations. Higher scores reflect better compliance with DGE recommendations. A sum dietary score ranging from 0 to 27 was calculated according to DGE guidelines and subsequently grouped into three categories: adverse ($\leq$13 points), ordinary (14–15 points), and favorable ($\geq$16 points) dietary patterns. This approach was established in earlier KORA studies and was validated against a weighed 7-day dietary protocol (Rabel et al., 2018; Winkler and Döring, 1998).

Only participants who attended at least two visits across the entire study period were included in this longitudinal analysis (Supplementary Fig. S1). Additionally, we excluded participants with missing data on covariates used in our main analysis. Written informed content was obtained from all participants. The KORA study was approved by the ethics committee of the Bavarian Chamber of Physicians (Munich, Germany).

2.2. Metabolite quantification and normalization

Blood samples were drawn into serum gel tubes between 8:00 am and 10:30 am after at least 8 h of overnight fasting. Serum was collected, filled into synthetic straws and stored in liquid nitrogen (−80 °C) until the metabolomics analyses were conducted. The metabolite profiling in serum samples was done with the AbsoluteIDQ™ p180 kit (BIOCRATES Life Sciences AG, Innsbruck, Austria) for KORA S4 (March–April 2011) and FF4 (February–October 2019), allowing for the simultaneous quantification of 188 metabolites. The AbsoluteIDQ™ p150 kit was used to quantify 163 metabolites in KORA F4 samples from August 2008 to March 2009. The assay procedures have been described previously in detail (Römisch-Margl et al., 2012).

Identical quality control procedures were used at each of the three time points. Each metabolite should meet the following three criteria: (1) The average value of the coefficient of variance (CV) in the five/six reference samples or three quality control (QC) samples should be <25%; (2) 50% of all measured sample concentrations for the metabolite should be above the limit of detection (LOD), which was defined as three times the median of zero samples; and (3) The rate of missing value of metabolite should be <5%. The non-detectable values of each metabolite were randomly imputed by values ranging from 75% to 125% of half of the lowest measured value of the corresponding metabolite in each plate. In order to minimize the plate effects in each

visit, plate normalization factors were calculated by dividing the mean of reference sample values (QC samples in KORA F4) in each plate by the mean of all reference sample values in all plates and then used to normalize each metabolite (Han et al., 2022; Huang et al., 2020).

Additionally, to control for the effects of the different kits between KORA F4 and KORA S4/FF4, up to eight participants' samples were randomly selected from each of the 36 kit plates in KORA F4 and re-measured using the same AbsoluteIDQ™ p180 kit used in KORA S4/FF4. The difference of each metabolite between the corresponding participants in KORA F4 and re-measured KORA F4, and a further mean difference of each metabolite were calculated. We calculated the kit normalization factor by dividing the mean of each metabolite in KORA F4 by the mean of each metabolite in KORA F4 minus the mean difference between KORA F4 and re-measured KORA F4. This kit normalization factor was then used to correct KORA F4 metabolite data. Extreme outliers of each metabolite were defined as a value beyond the mean ± 5×standard deviations range and imputed by the K-nearest neighbors algorithm (KNN) (Han et al., 2022; Huang et al., 2020). In total, 135 metabolites in KORA S4, 114 in KORA F4, and 145 in KORA FF4 passed the quality control. Out of these, 108 metabolites overlapped across KORA S4, F4, and FF4 and were used in the subsequent analysis. Metabolites covered the following compound classes: 12 amino acids, 12 acylcarnitines, 72 glycerophospholipids (including 32 phosphatidylcholines (PCs) with acyl-acyl (diacyl) side chains (PC aa), 33 phosphatidylcholines with acyl-alkyl side chains (PC ae), and seven lysophosphatidylcholines (LPCs)), 11 sphingomyelins (SM) and a sum of hexoses (including glucose). The complete list of metabolites is presented in the supplementary material (Supplementary Table S1).

2.3. Air pollution

Hourly averages of air pollutants (particulate matter (PM) with an aerodynamic diameter ≤2.5 μm ($PM_{2.5}$) or between 2.5 and 10 μm (PM_{coarse}), nitrogen dioxide (NO_2), and ozone (O_3)) and meteorological parameters (temperature and relative humidity) were collected at fixed background monitoring sites (Wolf et al., 2015). 24-h average concentrations of each air pollutant and the meteorological parameters on the day of blood withdrawal of each visit were calculated for each monitor when at least 75% of the hourly values were available. For O_3, daily maximum 8-h average levels were calculated. Furthermore, 2-day (same day and previous day as individual's blood withdrawal), 2-week, and 8-week moving averages prior to the examination day were calculated based on the daily average concentrations to explore the effects of immediate, short- and medium-term exposure.

Due to the different operating periods of monitoring sites, the time series of daily air pollutant concentrations were created by combining exposure data from different monitoring sites for specific periods. To obtain one consistent time series for each air pollutant, we selected the station with the longest monitoring period for each air pollutant as the master station. We then imputed missing values by running linear regressions between the master monitor and the other monitors and selecting the one with the highest explained variance (R^2) (Supplementary Table S2). The daily averages of $PM_{2.5}$ for 2004 to 2014 were obtained from an aerosol monitoring station (FH) located approximately 1 km southeast of the city center of Augsburg with a distance of 100 m to the nearest major road. This monitoring station was established in 2004 and is considered a representative site of the urban background in Augsburg (Cyrys et al., 2008). Since $PM_{2.5}$ was not available before 2004, we used PM_{10} (PM with an aerodynamic diameter ≤10 μm) concentrations and downscaled them by a factor of 0.68 to derive $PM_{2.5}$. These PM_{10} concentrations were monitored at an urban background station (Bourgesplatz, BP) located approximately 2 km north of the city center, as well as 20 m to the nearest road with low traffic intensity and 100 m to the nearest road with high traffic intensity (Cyrys et al., 2008). The factor 0.68 to predict $PM_{2.5}$ by PM_{10} was based on the median ratio of $PM_{2.5}/PM_{10}$ during the period of ULTRA III measurements (Environmental Nanoparticles and Health: Exposure, Modeling and Epidemiology of Nanoparticles and their Composition) (Wolf et al., 2017). The measured data were then calibrated based on a linear regression model using the overlapping period with FH to yield continuous $PM_{2.5}$ data from 1999 to 2014. PM_{coarse} was calculated by scaling PM_{10} with a factor of 0.32 until the end of 2004. From the beginning of 2005, it was calculated as the difference between PM_{10} and $PM_{2.5}$ following the method of the ESCAPE project where the KORA cohort was part of Eeftens et al. (2012). Daily NO_2 concentrations from 1999 to 2014 were obtained from the BP station. Missing values were imputed by monitoring data from a single urban background monitoring site operated by the Bavarian Environment Agency (LfU), approximately 4 km south of the city center, using linear regressions. O_3 concentrations were measured at the Haunstetten monitoring station (HAU, a suburb 7 km south of the Augsburg city center) until 2001 (Chen et al., 2020; Wolf et al., 2015). From 2001 on, monitoring was done at the LfU monitoring station. To account for the difference between monitoring devices at these two stations, we calibrated the measured data at the HAU monitoring station using a linear regression model using the overlapping period. Hence, a continuous O_3 time-series dataset was derived from 1999 to 2014. Meteorological parameters were collected separately from the LfU site, and missing days were imputed by the corresponding data measured at the HAU station (Chen et al., 2020; Wolf et al., 2015). Further missing rates of air pollutants were all less than 5% and imputed by the median of three-month moving averages, which account for the seasonal effects.

2.4. Statistical methods

Basic descriptive analyses were performed for participant characteristics, air pollutants, and meteorological parameters. The differences between surveys were examined by a Kruskal-Wallis test (one-way ANOVA) applied for continuous variables and a Pearson's Chi-squared test for categorical variables. Spearman's rank correlation coefficient was used to calculate correlations between air pollutants and meteorological parameters.

We applied generalized additive mixed-effects models with random intercepts for participants to examine the associations between repeatedly measured metabolite concentrations and air pollutants. All metabolites were natural log-transformed to increase the conformity to normal distributions of residuals. A total of three exposure windows were considered for each pollutant, covering immediate and cumulative short-term and medium-term effects: 2-day (same-day and previous-day exposure as the individual's blood withdrawal), 2-week and 8-week moving averages before the examination day. Related exposure windows of air temperature and relative humidity were calculated. Covariates included in the models were selected *a priori* based on literature (Breitner et al., 2016; Ward-Caviness et al., 2016). Minimum models were adjusted for age, sex, BMI, an indicator of each study wave (KORA S4, F4, or FF4), the day of the week, time trend, temperature, and relative humidity. We used regression splines to account for non-linearity in the relationships between time trends, temperature or relative humidity, and metabolites. Three degrees of freedom were used for the temperature and relative humidity splines, and four degrees of freedom per year in each study wave for the time trend spline. Main models additionally included smoking status (never/former/current), alcohol consumption (g/day), physical activity index (low/medium/high), educational attainment (primary school/high school/college), fasting status (overnight fasting of 8 h or not) and dietary score (continuous). In the extended models, we further added hypertension, diabetes, medication intake (anti-hypertensive, anti-diabetic, or lipid-lowering medication), high-density lipoprotein (HDL), and total cholesterol. Effect estimates are presented as percent changes in the geometric mean (together with 95% confidence intervals [95% CI]) of the repeatedly assessed outcomes per interquartile range (IQR) increase in air pollutant concentrations.

Effect modification was investigated by including an interaction term between each air pollutant at each exposure window and the potential effect modifier assessed at each visit. The examined modifiers included age (<65 years *vs* ≥65 years; 65 years is the current official retirement age in Germany), sex (male *vs* female), obesity (BMI ≥30 kg/m^2 *vs* <30 kg/m^2), smoking status (current/former *vs* never smoker), physical activity (low *vs* medium *vs* high), dietary pattern (adverse *vs* ordinary *vs* favorable), hypertension (yes *vs* no), and diabetes (yes *vs* no).

Several sensitivity analyses were performed in this study: 1) We included all participants with at least one visit in KORA S4, F4, or FF4 with complete data on air pollution, phenotypes, and metabolites in the analysis; 2) Only participants with fasting blood samples throughout the entire study period were included; 3) To control for selection bias introduced by selecting participants with more than one measurement, we estimated weights for those included using the inverse probability weighting (IPW) method (Weuve et al., 2012). Briefly, the probability of being included in our main analysis among all study participants in KORA S4 was calculated using logistic regression. We used individual characteristics of our main analysis as possible predictors. Then, we applied the inverse of the predicted probability determined from the logistic regression as the weight in our main model; 4) We applied a crude model on the immediate and short-term analysis, which excluded age, sex, and BMI from our minimum model to investigate the acute effects from air pollutants and meteorological parameters; 5) To investigate the co-effects between short-term and long-term air pollution exposure, we simultaneously included the corresponding long-term exposures estimated by the land use regression models for each air pollutant (Yao et al., 2022); and 6) We performed two-pollutant models by including two air pollutants simultaneously if their correlation was smaller than 0.7.

All statistical analyses were done with R (version 4.1.2). The *p*-value cut-off was set as 3.9×10^{-5} to account for multiple testing introduced by four air pollutants ($PM_{2.5}$, PM_{coarse}, NO_2, and O_3), three exposure windows, and 108 metabolites used in this study (0.05 / (4 * 3 * 108)).

2.5. Pathway analysis

We performed pathway analysis using the air pollutants-associated metabolites *via* the "Pathway Analysis" module in MetaboAnalyst 5.0 (Pang et al., 2021). More details were described in our previous study (Yao et al., 2022). In brief, this module supports pathway analysis by integrating pathway enrichment analysis and pathway topology analysis based on the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. The *p*-value of the pathway enrichment analysis indicated the probability of seeing at least a particular number of metabolites from a collection of pathways of the KEGG pathway database in the metabolomics platform we used (Wieder et al., 2021). The pathway topology analysis uses two well-established node centrality measures to estimate node importance. Furthermore, to take into account the comparison among different pathways, the node importance values calculated from centrality measures are further normalized by the sum of the importance of the pathway. Therefore, the total/maximum importance of each pathway is one. The importance measure of each metabolite node reflects the percentage with regard to the total pathway importance, and the pathway impact value is the cumulative percentage from the matched metabolite nodes. Pathways with a *p*-value ≤ 0.1, or with an impact value > 0.5 while *p*-value ≤ 0.3 were considered the most relevant pathways.

3. Results

3.1. Demographic characteristics

Participant characteristics are summarized in Table 1. Only participants attending at least two visits during the entire study period with no missing information in the main confounders were included in our main analyses. Therefore, among 9,620 observations from 4,261 study participants in the KORA cohort between 1999 and 2014, we included 5,772 (60.0%) observations from 2,583 (60.6%) participants in this analysis. Specifically, 1,977 (76.5%) out of the 2,583 participants attended two examinations, and 606 (23.5%) attended all three examinations (Table 1).

Since metabolite concentrations in KORA S4 were only determined in 1,610 participants between 55 and 74 years as only these were asked to be fasting, 1,605 of the overall 4,261 participants could be included in the analysis after the data quality control. Therefore, on average, KORA S4 participants were older than those of KORA F4 and KORA FF4 (Table 1). Meanwhile, the average educational attainment, the percentages of 8 h overnight fasting before blood withdrawal, current smoking status, unhealthy dietary pattern, and medium and high levels

Table 1
Descriptive statistics of participant characteristics at each study.

	S4 (N = 1,129)	F4 (N = 2,556)	FF4 (N = 2,087)	*p*-Value
	Mean ± SD/ N (%)	Mean ± SD/ N (%)	Mean ± SD/ N (%)	
Age (years)	63.3 ± 5.4	57.5 ± 13.3	60.7 ± 12.3	<0.001
Sex (male)	570 (50.6)	1,240 (48.5)	1,012 (48.5)	0.46
Education				<0.001
Primary school	753 (66.7)	1,357 (53.1)	1,034 (49.5)	
High school	221 (19.6)	621 (24.3)	530 (25.4)	
College	155 (13.7)	578 (22.6)	523 (25.1)	
BMI (kg/m^2)	28.4 ± 4.2	27.7 ± 4.7	27.8 ± 4.9	<0.001
Alcohol consumption (g/day)	16.2 ± 20.9	14.4 ± 19.5	14.9 ± 20.1	0.025
Dietary score	16.2 ± 3.6	15.3 ± 3.6	15.1 ± 3.6	<0.001
Dietary patterns				<0.001
Adverse	271 (24.0)	817 (31.9)	715 (34.3)	
Ordinary	212 (18.8)	541 (21.2)	451 (21.6)	
Favorable	646 (57.2)	1,198 (46.9)	921 (44.1)	
Fasting (8 h) (% yes)	1,016 (90.0)	2,543 (99.5)	2,074 (99.4)	<0.001
Smoking status				0.002
Current smoker	137 (12.1)	384 (15.0)	307 (14.7)	
Former smoker	437 (38.7)	1,066 (41.7)	902 (43.2)	
Never smoker	555 (49.2)	1,106 (43.3)	878 (42.1)	
Physical activity				<0.001
Low	444 (39.3)	818 (32.0)	589 (28.2)	
Medium	479 (42.4)	1,115 (43.6)	952 (45.6)	
High	206 (18.3)	623 (24.4)	546 (26.2)	
Hypertension (% yes)	609 (53.9)	1,016 (39.8)	825 (39.5)	<0.001
Diabetes (% yes)	92 (8.2)	224 (8.8)	215 (10.3)	0.11
Medication usage (% yes)				
Anti-hypertension medication	397 (35.2)	861 (33.7)	782 (37.5)	0.03
Anti-diabetes medication	53 (4.7)	154 (6.0)	174 (8.3)	<0.001
Lipid lowering medication	128 (11.3)	351 (13.7)	342 (16.4)	<0.001
Cholesterol (mg/dL)*	243.6 ± 40.8	216.1 ± 38.8	216.7 ± 39.5	<0.001
HDL (mg/dL)*	58.1 ± 16.5	56.1 ± 14.4	65.9 ± 18.8	<0.001

S4 = fourth cross-sectional health survey of the KORA cohort; F4 = first follow-up examination of KORA S4; FF4 = second follow-up examination of KORA S4; BMI = body mass index; HDL = high density lipoprotein; S4 participants were selected based on whether they did fast or not. Dietary patterns were classified by a dietary score basing on the assessment of individual's daily diet (questionnaire): Adverse = ≤13 points, Ordinary = 14–15 points, Favorable = ≥16 points. Physical activity was defined according to the exercise time per week: Low = almost or no sporting activity, Medium = regular/irregular approx. 1 h/week, High = regularly >2 h in the week.
SD = standard deviation; *p*-value was based on the Kruskal-Wallis test for continuous variables, and Pearson's Chi-squared test for categorical variables.
* Cholesterol was missing for one (0.09%) participant in KORA S4, and one (0.05%) in KORA FF4; HDL was missing for one participant in each KORA S4 (0.09%), F4 (0.04%), and FF4 (0.05%). 1977 participants attended two examinations, and 606 attended three examinations.

of physical activity of KORA S4 were lower (p-value < 0.01). In contrast, mean BMI, alcohol consumption, cholesterol, HDL, and the percentage of hypertension were higher in KORA S4 (p-value < 0.01).

3.2. Characteristics of air pollutants

The daily average concentrations of $PM_{2.5}$, PM_{coarse}, NO_2, and O_3 at different survey times are listed in Table 2. The daily average concentrations of $PM_{2.5}$ in KORA S4 were slightly higher than the WHO Air Quality Guideline (AQG) value of 15 μg/m^3, while the values at the other two survey times were similar or lower. For O_3, the daily average exposures at all three KORA study visits were lower than the WHO AQG levels of 100 μg/m^3. In contrast, the daily average concentrations of NO_2 in KORA S4 (39.2 μg/m^3), KORA F4 (33.0 μg/m^3), and KORA FF4 (29.6 μg/m^3) were all higher than the WHO AQG level (25 μg/m^3). In general, air pollutant levels and relative humidity during the survey times gradually decreased from KORA S4 to KORA FF4, except for O_3 and temperature, which overall showed an increasing trend (Table 2). All pollutants showed strong positive relationships with each other; only O_3 showed weak or negative correlations with other air pollutants but a positive correlation with temperature. Like O_3, relative humidity also showed weak or negative associations with other air pollutants and temperature (Supplementary Fig. S2).

3.3. Associations between metabolites and air pollutants in different exposure windows

There were multiple significant associations between metabolites and air pollutants in different exposure windows at a Bonferroni-corrected p-value after adjusting for covariates ($p < 3.9 \times 10^{-5}$) (Fig. 1 and Supplementary Table S3). In general, medium-term exposures (8-week exposure) to each air pollutant showed the strongest effects. For $PM_{2.5}$, medium-term exposures were significantly associated with arginine (Arg), tryptophan (Trp), and eight phosphatidylcholines. Only Trp and one phosphatidylcholine (PC aa C30:0) were negatively associated with short-term (2-week moving average) $PM_{2.5}$ exposure. Medium-term PM_{coarse} exposure was positively associated with amino acids, including glycine (Gly), methionine (Met), ornithine (Orn), phenylalanine (Phe), serine (Ser) and threonine (Thr), one lysophosphatidylcholine (LPC a C16:0), one phosphatidylcholine (PC ae C40:4), and one sphingomyelin (SM C16:0). In addition, PC aa C40:4 was negatively associated with immediate exposure to PM_{coarse}, and PC ae C44:3 was positively associated with short-term exposure to PM_{coarse}. NO_2 exposure showed the largest number of associations with metabolites, particularly at the medium-term exposure windows, which covered each metabolite group in this study, including 33 PCs, five LPCs, seven SMs, three amino acids (Arg, Trp, and tyrosine (Tyr)), and one acylcarnitine (C0). A few metabolites from PCs, LPCs, or the amino acid group were negatively associated with immediate or short-term exposure to NO_2. Those metabolites remained in the medium-term exposure window of NO_2. Only two metabolites were positively associated with O_3 exposures. Trp was associated with both short- and medium-term O_3 exposures, whereas SM (OH) C24:1 was only associated with immediate O_3 exposures. In addition, we listed the non-overlapping metabolites throughout the three exposure windows as well, which were written in blue. The direction of the effect estimates was generally the same representing the consistency of our results in different exposure windows (Supplementary Table S3).

To summarize, metabolites showing significant associations with all three air pollutants were mainly from the PCs group (Supplementary Table S3). Particularly, we observed that Trp, PC aa C40:4, and PC ae C42:5 showed robust associations with all NO_2 exposure windows (Fig. 2). There were also some overlapping metabolite associations for different air pollutants in the same exposure windows. For example, Trp was associated with short- and medium-term exposures to NO_2, $PM_{2.5}$, and O_3 (Fig. 2 and Supplementary Table S3). Moreover, PC aa C40:4 was significantly associated with immediate exposures to both PM_{coarse} and NO_2, with the effect estimates in the same direction (Supplementary

Table 2
Descriptive statistics of the 2-day, 2- and 8-week moving averages of air pollutant concentrations and meteorological parameters* (N = 5,772).

		2-day moving average		2-week moving average		8-week moving average	
		Mean ± SD	IQR	Mean ± SD	IQR	Mean ± SD	IQR
S4	$PM_{2.5}$ (μg/m^3)	15.6 ± 7.0	8.2	15.7 ± 4.3	5.9	15.8 ± 2.5	4.6
	PM_{coarse} (μg/m^3)	7.3 ± 3.3	3.8	7.4 ± 6.3	2.8	7.4 ± 1.2	2.2
	NO_2 (μg/m^3)	38.3 ± 10.8	14.2	37.0 ± 5.8	7.0	36.4 ± 3.6	4.4
	O_3 (μg/m^3)	41.4 ± 20.4	34.1	41.5 ± 17.6	31.5	42.7 ± 15.9	28.2
	Relative humidity (%)	79.2 ± 11.1	16.0	79.9 ± 8.1	12.2	79.4 ± 7.0	10.7
	Air temperature (°C)	8.5 ± 6.7	8.8	8.4 ± 6.2	10.9	9.0 ± 6.0	10.8
F4	$PM_{2.5}$ (μg/m^3)	14.7 ± 10.5	12.0	14.4 ± 5.6	7.1	15.0 ± 11.2	5.1
	PM_{coarse} (μg/m^3)	4.9 ± 3.4	3.7	4.6 ± 2.3	2.8	5.0 ± 3.9	1.8
	NO_2 (μg/m^3)	32.4 ± 10.6	16.0	31.3 ± 6.0	7.7	33.0 ± 11.8	5.3
	O_3 (μg/m^3)	39.6 ± 21.7	36.7	40.4 ± 17.6	30.3	39.2 ± 22.6	28
	Relative humidity (%)	76.5 ± 9.1	11.7	76.1 ± 6.8	8.9	76.6 ± 9.9	9.5
	Air temperature (°C)	7.9 ± 6.5	9.4	8.1 ± 5.9	9.1	7.9 ± 6.6	9.8
FF4	$PM_{2.5}$ (μg/m^3)	10.5 ± 7.0	7.9	10.5 ± 4.3	5.6	11.0 ± 2.9	4.6
	PM_{coarse} (μg/m^3)	5.5 ± 3.6	3.4	5.0 ± 2.0	2.6	5.2 ± 1.4	2.3
	NO_2 (μg/m^3)	28.8 ± 10.2	13.0	27.2 ± 5.4	7.7	27.5 ± 4.8	8.9
	O_3 (μg/m^3)	44.3 ± 20.8	33.9	43.8 ± 17.9	33.8	43.3 ± 16.3	30.9
	Relative humidity (%)	72.7 ± 11.1	17.8	73.4 ± 8.5	13.0	73.2 ± 6.9	11.8
	Air temperature (°C)	12.4 ± 7.3	12.3	12.0 ± 6.5	11.7	11.9 ± 5.8	9.8
Entire study period	$PM_{2.5}$ (μg/m^3)	13.4 ± 9.0	10.0	13.3 ± 5.3	7.0	13.7 ± 3.7	6.9
	PM_{coarse} (μg/m^3)	5.6 ± 3.6	3.9	5.3 ± 2.4	3.3	5.3 ± 1.8	2.9
	NO_2 (μg/m^3)	32.2 ± 11.0	14.8	30.9 ± 6.7	9.5	31.0 ± 5.3	8.9
	O_3 (μg/m^3)	41.6 ± 21.2	35.0	41.8 ± 17.8	32.4	41.2 ± 16.0	30.1
	Relative humidity (%)	75.7 ± 10.6	15.6	75.8 ± 8.0	11.5	76.0 ± 6.8	11.9
	Air temperature (°C)	9.7 ± 7.1	10.8	9.6 ± 6.4	10.5	9.6 ± 6.0	11.2

* Exposure levels were measured at fixed monitoring sites. S4 = fourth cross-sectional health survey of the KORA cohort; F4 = first follow-up examination of KORA S4; FF4 = second follow-up examination of KORA S4; Entire study period = overall study period covered all three visits. $PM_{2.5}$ = particulate matter with an aerodynamic diameter less than or equal to 2.5 μm; PM_{coarse} = particulate matter with an aerodynamic diameter of 2.5–10 μm; NO_2 = nitrogen dioxide; O_3 = ozone; 2-day moving average = 2-day moving average of exposure levels before the examination day; 2-week moving average = 2-week moving average of exposure levels before the examination day; 8-week moving average = 8-week moving average of exposure levels before the examination day. SD = standard deviation; IQR = interquartile range.

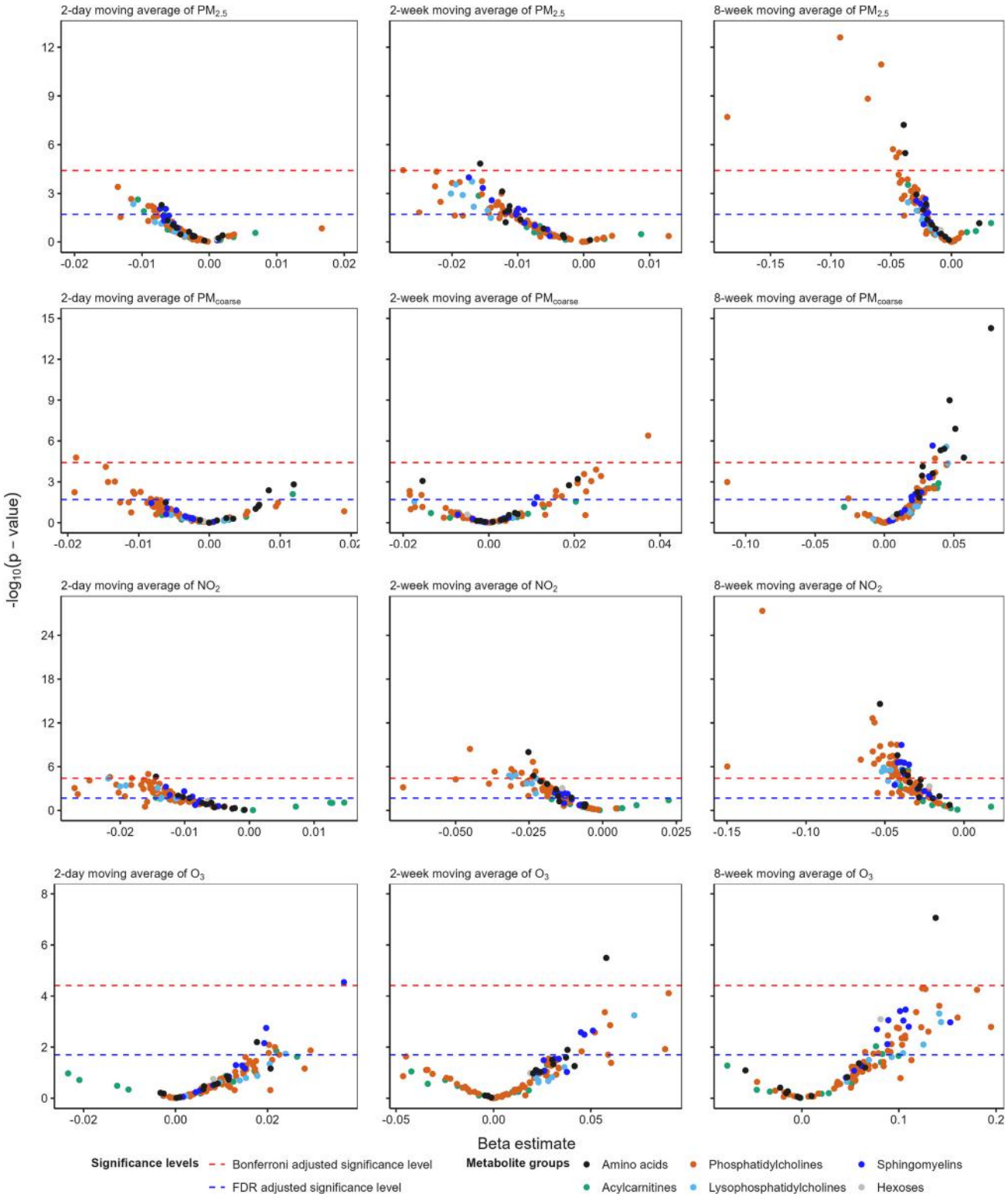


Fig. 1. Volcano plots of the associations between immediate, short- and medium-term exposures to air pollutants and metabolites. The results were derived from the main models adjusted for age, sex, body mass index (BMI), an indicator of each survey time (KORA S4, KORA F4, or KORA FF4), day of the week, time trend, temperature, relative humidity, smoking status, alcohol consumption, physical activity index, educational attainment, fasting status, and dietary score. The Y-axis shows the negative logarithm of the *p*-value (logarithmic base of 10). The X-axis indicates the association between air pollutant exposures and metabolites. The red and blue dashed lines represent the statistical significance levels of the *p*-value adjusted by the Bonferroni and the FDR method, respectively. The six different colors for the points represent the six metabolite groups involved in this study, including amino acids (black), acylcarnitines (green), phosphatidylcholines (orange), lysophosphatidylcholines (light blue), sphingomyelins (blue), and hexoses (grey). $PM_{2.5}$ = particulate matter with an aerodynamic diameter less than or equal to 2.5 μm; PM_{coarse} = particulate matter with an aerodynamic diameter of 2.5–10 μm; NO_2 = nitrogen dioxide; O_3 = ozone.

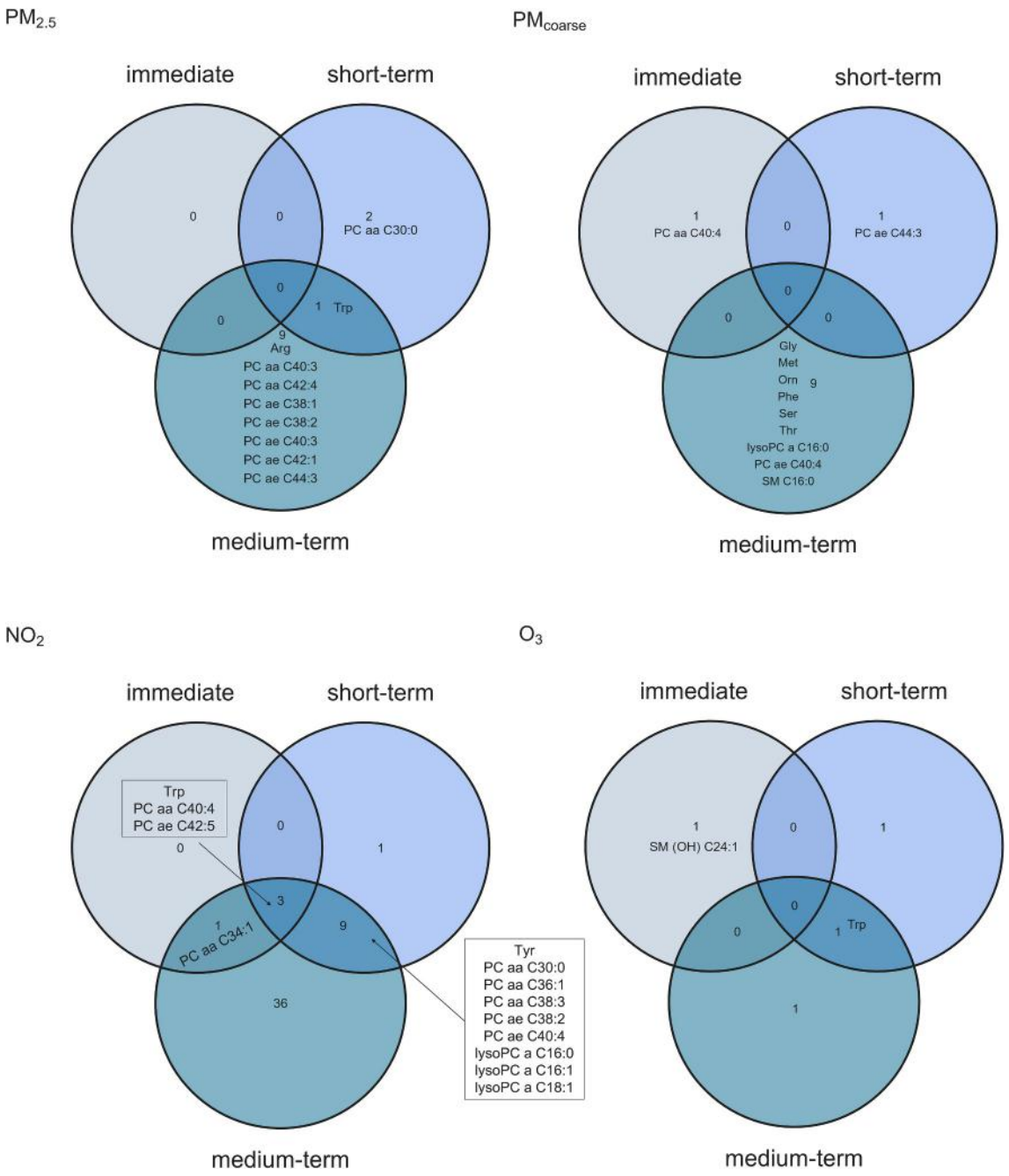


Fig. 2. Venn diagrams of significant associations between immediate, short- and medium-term air pollutant exposures and metabolites. $PM_{2.5}$ = particulate matter with an aerodynamic diameter less than or equal to 2.5 μm; PM_{coarse} = particulate matter with an aerodynamic diameter of 2.5–10 μm; NO_2 = nitrogen dioxide; O_3 = ozone; immediate = 2-day moving average of exposure levels before the examination day; short-term = 2-week moving average of exposure levels before the examination day; medium-term = 8-week moving average of exposure levels before the examination day. Only overlapping metabolites were displayed in the overlapped area instead of showing all the significant results. The positive number in short- and medium-term plots represents the non-overlapping metabolites for each air pollutant; 0 indicates either no overlapping metabolite or no additional significant result for the corresponding air pollutant.

Fig. S3 and Table S3).

3.4. Effect modification

In general, we did not find consistent results from the effect modification analysis between all metabolites and air pollutants in different exposure windows. We restricted the results of effect modification to Trp, PC aa C40:4, and PC ae C42:5 as they showed more robust associations with different air pollutants compared to other metabolites (Supplementary Figs. S4–S7).

3.5. Sensitivity analysis

In the different sensitivity analyses, the associations between serum metabolites and air pollutants with different exposure windows were generally robust, particularly in those metabolites showing consistent associations with air pollutants. The results were similar when including all participants who at least attended once in KORA S4, F4, or FF4, restricting the analysis to fasting individuals, or using predicted IPW to adjust for potential selection bias of participants. The results only slightly changed when we applied the crude models, which only took the effects from short-term exposure to air pollution and meteorological

parameters into account, or when we performed the co-effect models by adding long-term air pollution exposure to the short-term analysis (Supplementary Tables S4–S7). In two-pollutant models with the additional inclusion of PM_{coarse} or O_3, the effect estimates of metabolites attenuated for $PM_{2.5}$ and NO_2, but generally strengthened for PM_{coarse} or O_3 (Supplementary Tables S4–S7).

3.6. Pathway analysis

The results from the pathway analysis showed that short-term exposure to PM_{coarse} was associated with glycerophospholipid metabolism (p-value = 0.04, impact value = 0.09) (Fig. 3 and Supplementary Table S8). Furthermore, medium-term exposure to PM_{coarse} was associated with cysteine and methionine metabolism (p-value = 0.05, impact value = 0.1), glyoxylate and dicarboxylate metabolism (p-value = 0.05, impact value = 0.2), and glycine, serine, and threonine metabolism (p-value = 0.008, impact value = 0.5). The glycerophospholipid metabolism was associated with immediate, short-term, and medium-term cumulative exposures to NO_2 (p-value < 0.1, impact value = 0.1), and the pathway of phenylalanine, tyrosine, and tryptophan biosynthesis was additionally associated with short-term cumulative exposures to NO_2 (p-value = 0.2, impact value = 0.5). However, they were insignificant after using the FDR method to correct the raw p-value (Supplementary Table S8).

4. Discussion

In this study, we investigated the effects of immediate, short- and medium-term exposures to ambient air pollution on serum metabolites in a longitudinal cohort setting. Multiple significant associations were identified between air pollutants within three exposure windows and metabolites, particularly the PCs subgroup. The glycerophospholipid metabolism showed robust associations with NO_2 in all exposure windows, as well as with short-term exposure to PM_{coarse} and medium-term exposure to $PM_{2.5}$. Furthermore, we observed that longer exposure windows (*e.g.*, 8-week moving averages) generally showed more consistent associations with metabolites.

Metabolites are the intermediates or end products of metabolism, can affect cellular physiology through modulation of other "omics" levels and represent changes induced by exposures (Rinschen et al., 2019). So far, only a few studies investigated the effects of acute, short- and medium-term exposures to air pollution on metabolites using population-based cohort data, especially in a longitudinal setting. A cross-sectional analysis based on the KORA F4 and F3 surveys found

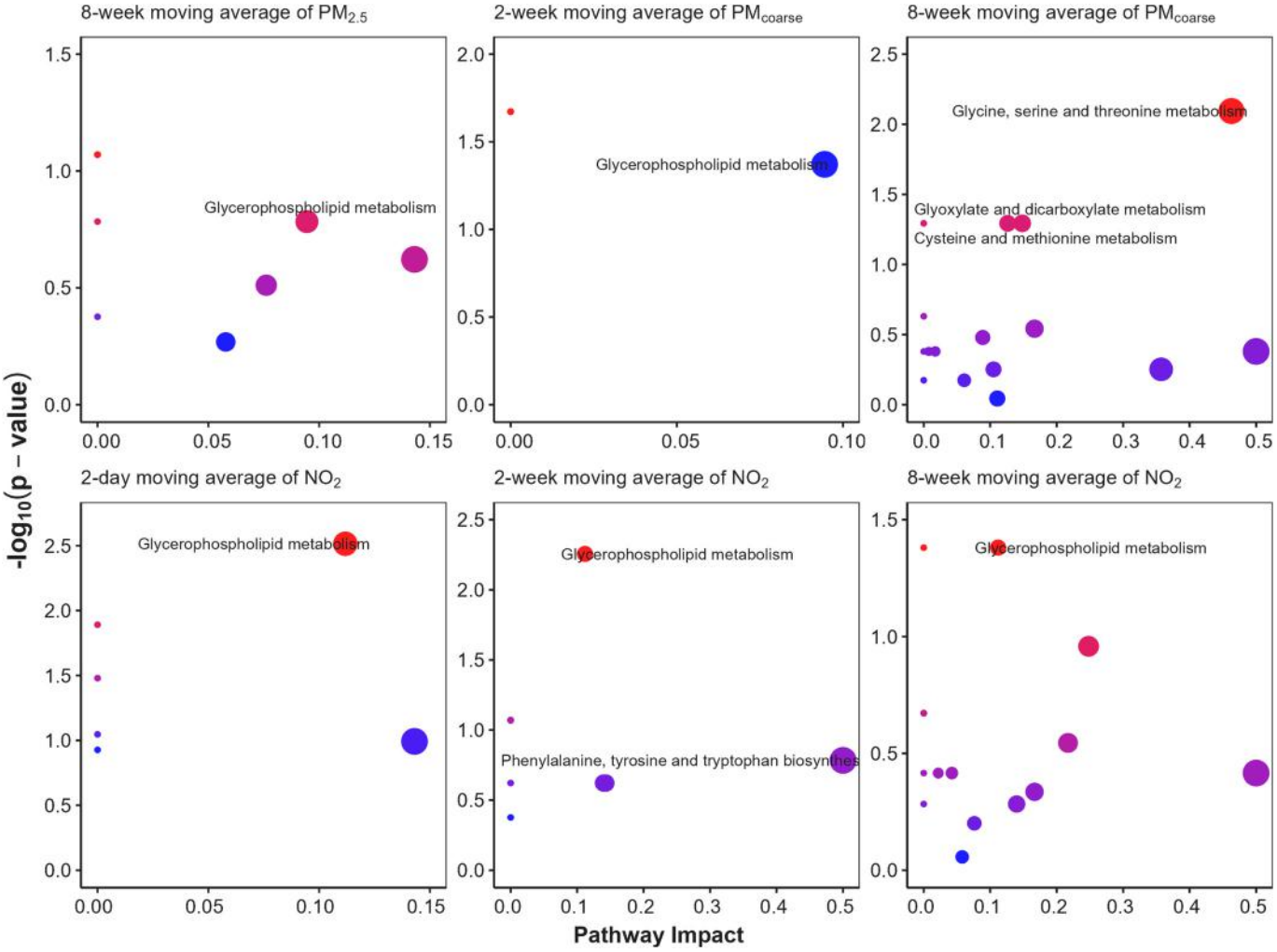


Fig. 3. Metabolic pathways identified for immediate, short-, and medium-term exposure to air pollutants. The pathway analysis is based on both enrichment analysis and pathway topology analysis. The Y-axis is the negative logarithm of the p-value (logarithmic base of 10) from the enrichment test. The X-axis indicates the structural impact of $PM_{2.5}$, PM_{coarse}, or NO_2-related metabolites in the enriched pathways, based on the cumulative importance of all the significant metabolites within the pathway based on the KEGG database. The size of each bubble represents the impact value (corresponding to the X-axis). The color of each bubble represents the significance of the enrichment (corresponding to the Y-axis from blue to red). $PM_{2.5}$ = particulate matter with an aerodynamic diameter less than or equal to 2.5 μm; PM_{coarse} = particulate matter with an aerodynamic diameter of 2.5–10 μm; NO_2 = nitrogen dioxide.

associations between short-term exposures to air pollutants and metabolite levels, mainly from the LPCs and PCs subgroup (Ward-Caviness et al., 2016). NO_2 was the most consistently associated exposure, with five out of ten significant associations. In the current longitudinal analysis based on the KORA S4 survey and two follow-ups, NO_2 still showed the most consistent effects; moreover, PCs were the major metabolite subgroup affected. Additionally, medium-term exposure to $PM_{2.5}$ was also negatively associated with PCs. PCs are one of the most abundant glycerophospholipids and are the most abundant phospholipid in all mammalian cell membranes and subcellular organelles. They can be attacked by reactive oxygen species (ROS), leading to lipid peroxidation (Ayala et al., 2014; Cole et al., 2012; van der Veen et al., 2017). The higher abundance of PCs in human tissues compared to other phospholipid classes has been shown to play an essential role in health and diseases (van der Veen et al., 2017). The altered PC metabolism may also promote the development of Alzheimer's and cardiovascular diseases (Tang et al., 2013; Whiley et al., 2014). In other metabolomics-related studies using the KORA cohort, a decrease in some PCs was associated with smoking and aging (Chak et al., 2019; Xu et al., 2013). We also found negative associations between four PCs and long-term exposure to air pollutants in our previous analysis based on the same subset of KORA participants (Yao et al., 2022). We further observed that adjusting for long-term air pollution exposure did not change the short- and medium-term effects in this study. However, the negative associations between LPCs and medium-term exposure to NO_2 we observed in the current study did not coincide with LPC alterations associated with short-term NO_2 in the previous cross-sectional study of KORA (Ward-Caviness et al., 2016). This might have resulted from the imbalance between PCs and LPCs triggered by the cumulative effects of NO_2 exposure over a longer exposure period. Similar results were observed in an *in vivo* study conducted in rats where both PCs and LPCs decreased after chronic exposure to $PM_{2.5}$, with repeated inflammation as the potential mechanism (Chen et al., 2014). Given that this finding was restricted to medium-term but not immediate, short-, or long-term exposures to air pollutants, the dynamic metabolic changes might have happened from immediate up to long-term effects of air pollutant exposures. Nevertheless, the consistent findings suggest that PCs might be underlying biomarkers involved in the biological mechanisms of air pollution exposure-related adverse outcomes.

We further identified some of these PCs associated with $PM_{2.5}$ (medium-term exposure) and NO_2 (short- and medium-term exposures), which were involved in glycerophospholipid metabolism. Similarly, an analysis using data from the US Normative Aging study identified several metabolic pathways - including the glycerophospholipid metabolism - to be affected by short-term exposures to $PM_{2.5}$ and NO_2 (Nassan et al., 2021a). In another study based on the same cohort, this metabolic pathway was also associated with short- and long-term exposures to $PM_{2.5}$ species (Nassan et al., 2021b). Further associations between glycerophospholipid metabolism and short-term traffic-related air pollution were found in pregnant women (Yan et al., 2019) and college students (Chen et al., 2019). This metabolic pathway is mainly involved in inflammatory and oxidative stress responses, atherosclerosis, and coronary artery disease (CAD) resulting from air pollution (Chen et al., 2021; Li et al., 2022b; Zhang et al., 2020). Therefore, the alterations in the lipid metabolism due to the imbalance between anti- and pro-inflammatory factors and oxidative stress levels might be one of the underlying mechanisms linking air pollution and adverse health outcomes.

In addition, we observed that arginine showed robust negative associations with medium-term exposures to $PM_{2.5}$ and NO_2. Similarly, the perturbation of arginine metabolism was reported by Liang and colleagues, who observed inverse associations between arginine and traffic-related air pollution (Liang et al., 2019; Liang et al., 2018). In a study conducted on cardiac catheterization patients, decreased arginine concentrations were associated with an IQR increase in short-term (particularly 1-day lag) exposure to $PM_{2.5}$ exposure (Breitner et al., 2016), which also supports the findings of this study. As a conditional essential amino acid, arginine has been demonstrated to be the precursor of nitric oxide (NO) synthesis. NO, acting as an endogenous anti-atherogenic molecule, plays an important role in the cardiovascular system. Therefore, the arginine-NO pathway plays a vital role in vascular endothelial function. Its perturbation might lead to endothelial dysfunction and act as a sign of various diseases, *e.g.*, cardiovascular diseases (Li et al., 2022a; Popolo et al., 2014). Furthermore, we found tryptophan, an essential amino acid, to be negatively associated with short-term exposure to NO_2 and medium-term exposures to $PM_{2.5}$ and NO_2. Similarly, a study of 31 healthy volunteers observed the perturbation of tryptophan metabolisms after acute exposure (blood samples were collected 2 h after exposure) to ambient air pollution (Vlaanderen et al., 2017). The decrease of tryptophan concentrations in serum could be induced by the pro-inflammatory cytokine interferon-γ (IFN-γ), mediating chronic immune-related inflammation; meanwhile, an unfavorable tryptophan metabolite profile was associated with atherosclerosis and cardiovascular diseases (Mangge et al., 2014).

Furthermore, a few amino acids, including glycine (Gly), methionine (Met), ornithine (Orn), phenylalanine (Phe), and serine (Ser), were positively associated with PM_{coarse} in our analysis. Some epidemiological studies specifically investigated the associations between amino acids and air pollution (Breitner et al., 2016; Cao et al., 2019; Feng et al., 2021; Gaskins et al., 2021; Hood et al., 2022; Hu et al., 2021; Li et al., 2021; Mu et al., 2019). However, the results were inconsistent among different studies since the influences of air pollution on amino acids could vary due to the differences in air pollutants and exposure periods. For example, Breitner et al. observed decreased Gly concentrations with an increase of 8.1 $\mu g/m^3$ in $PM_{2.5}$ (short-term exposure) but an increased Orn with increased O_3 exposure (Breitner et al., 2016). In another study, which collected the air pollution levels before, during, and after the 2008 Beijing Olympics, the Gly concentrations decreased following a substantial decrease in air pollution (Mu et al., 2019). In two recent studies, Feng et al. did not find significant associations between Gly and air pollutants. However, the effects of air pollutants ($PM_{2.5}$, NO_2, and carbon monoxide (CO)) on Gly fluctuated from 1-day to 7-day exposure windows, with enhanced effects for longer exposure windows (5–7 days) (Feng et al., 2021). In another study, air pollution effects on amino acids varied from acute (12h) to short-term (2 weeks) exposure windows (Hu et al., 2021). For example, Orn increased per 10-ppb increase in both 12-h and 24-h O_3 exposure; however, the magnitude of those associations attenuated with longer exposure windows of O_3 (1-week and 2-week). In this study, Gly and Orn were markedly influenced by longer exposure windows of PM_{coarse}, which is supported by Feng et al.'s study but different from Hu et al. This might have partly resulted from the differences in air pollutants, exposure periods, and sample sizes of studies because we analyzed longer cumulative effects of each air pollutant within a large population-representative cohort. Given the decreased Arg levels associated with medium-term exposure to $PM_{2.5}$ and NO_2 mentioned above, air pollution exposure might cause a perturbation of the urea cycle. In the urea cycle metabolism, arginine could be degraded and synthesized to Orn and urea. The imbalance between the depletion of Arg and the synthesis to Orn has been reported to be related to type 2 diabetes (Cao et al., 2019). Further evidence from the pathway analysis identified that those amino acids were involved in cysteine and methionine metabolism, glyoxylate and dicarboxylate metabolism, and glycine, serine, and threonine metabolism. A study among women undergoing infertility treatment also reported that glycine, serine, and threonine metabolism was only associated with intermediate (2-week) and long-term (3-month) exposures to $PM_{2.5}$. In contrast, cysteine and methionine metabolisms were associated with acute (1–3 days) and long-term exposure windows (Hood et al., 2022), partly coinciding with our results. Similarly, Li et al. observed that cysteine and methionine metabolisms were identified in association with acute traffic air pollution (moving average of lag 1–2 days) (Li et al., 2021). In the study among women undergoing infertility treatment, cysteine, methionine,

and glycine, serine, and threonine metabolisms were associated with exposures to black carbon (soot), NO_2, O_3, and $PM_{2.5}$ exposures, and glyoxylate and dicarboxylate were related to exposures to black carbon and NO_2 (Gaskins et al., 2021). The methionine metabolism could regulate metabolic processes, induce antioxidant defense, and counteract oxidative stress (Martínez et al., 2017). The glyoxylate and dicarboxylate metabolism contributed to insulin resistance and was reported to be associated with atherosclerosis and obesity (Proffitt et al., 2022). Meanwhile, the glycine, serine, and threonine metabolism regulates the tricarboxylic acid cycle and glycolysis and is also involved in antioxidant defense (Amelio et al., 2014).

All those pathways play important roles in antioxidant defense, oxidative stress, or insulin metabolism, which might be the link between air pollution exposure and diseases such as cardiovascular diseases and type 2 diabetes (Wang-Sattler et al., 2012; Ward-Caviness et al., 2017). Although overlapping metabolic pathways were identified across different studies, the results probably varied due to different exposures and exposure periods. Therefore, more studies are needed to validate these findings.

5. Strengths and limitations

The major strength of this study is the use of repeated metabolite measurements to explore the health effects of immediate, short-, and medium-term exposures to ambient air pollution within a large, population-representative cohort study of adults. The KORA cohort is a well-characterized study with standardized and comprehensive methods to collect individual information, enhancing the reliability of our results. Furthermore, the longitudinal study design with repeated measurements of biomarkers strengthened statistical power compared to the cross-sectional design (Guo et al., 2013; Wen et al., 2023) and reduced potential residual confounding from unmeasured factors. The targeted metabolomics approach used in our study has the strength to give correct annotations of each metabolite and quantitate all metabolites compared to untargeted metabolomics analysis in which unknown metabolites are also quantified, which might mislead false annotation for metabolites and is not sufficient to provide quantitative values for certain metabolites.

Some limitations, however, should also be mentioned. Firstly, the targeted metabolomics approach lowered the opportunity for new biomarkers discovery and could not fully represent the whole metabolome. Secondly, the air pollution data was collected from fixed monitoring sites, which might differ from real personal exposure and cause exposure misclassification, particularly for NO_2. However, we expect this misclassification error to be a Berkson-type error, which is likely independent of the real exposure (Armstrong, 1998; Zeger et al., 2000). It is assumed that the Berkson-type error does not lead to bias in the regression coefficients and only impacts the standard error and leads to imprecise estimates with wider confidence interval (Armstrong, 1998). Considering repeated measurements, both individual-specific and autocorrelated Berkson-type errors might exist, and higher Berkson-type error reduces the bias of the effect estimate (Deffner et al., 2018). In a study from the NAS cohort using exposures from both address-predicted models and a central-site monitoring (Nassan et al., 2021a), the authors observed more significant hits with the central-site exposures, suggesting that the central site either captures purely temporal variation or captures regional patterns better than the address-specific predictions. Therefore, the higher temporal variability might be more important than the spatial variability in short-term exposures. Except for O_3, the other pollutants in our study generally have a similar day-to-day changing trend (going up and down simultaneously over the region), which might indicate a more relevant-temporal variability in the Augsburg region in short-term exposures (Supplementary Fig. S8). Thirdly, the monitors we used are considered urban background sites, which might underestimate the real exposure concentrations for some participants. Studies in urban areas have shown that the spatial variability for $PM_{2.5}$ and PM_{10} is generally small whereas the temporal correlation from measurements at different sites is high (Monn, 2001), which was also observed in our study (Supplementary Table S2). NO_2 levels could decline with distance from the roadside (Roorda-Knape et al., 1998). As described before, the NO_2 monitoring station in this study (BP) was located approximately 2 km north of the city center, as well as 20 m to the nearest road with low traffic intensity and 100 m to the nearest road with high traffic intensity. Since most of our participants (5,095 out of 5,772 observations) lived within 1 km of a major road, we believe that the monitored NO_2 exposures can capture the exposure level of the KORA participants. Fourthly, the constant exposure across individuals on the same day might cause difficulty in identifying to what extent individual characteristics might act as potential confounders. We generally did not find consistent effect modifications in different exposure windows though for some metabolites and air pollutants, we observed stronger or weaker associations in different subgroups. Therefore, we included these characteristics as potential confounders to take into account the variations in the exposure-outcome relationship across different levels of individual characteristics. In addition, the results from the sensitivity analysis considering only the effects of short-term exposure to air pollution and meteorological parameters in the immediate and short-term exposure windows were robust compared to our main results suggesting minor effects from individual characteristics. Finally, the results presented here are built on repeated measurements. However, given the uniqueness of the ability to assess metabolite trajectories, we were unable to replicate the findings in an independent dataset.

6. Conclusion

In conclusion, immediate, short- and medium-term exposures to air pollutants were associated with multiple serum metabolites. More associations were seen for NO_2 with longer exposure windows and phosphatidylcholines, particularly PC aa C40:4 and PC ae C42:5. Glycerophospholipid metabolism, involved in inflammation and oxidative stress, was identified as the most robust pathway associated with different air pollutants in different exposure windows. These findings might help to better understand the underlying mechanisms of air pollution-associated adverse health outcomes.

Funding

This work was supported by a scholarship under the State Scholarship Fund by the China Scholarship Council (File No. 201906180003). The KORA study was initiated and financed by the Helmholtz Zentrum München – German Research Center for Environmental Health, which is funded by the German Federal Ministry of Education and Research (BMBF) and by the State of Bavaria. Data collection in the KORA study is done in cooperation with the University Hospital of Augsburg. Furthermore, this study was supported by the Helmholtz Alliance "Aging and Metabolic Programming" (AMPro).

CRediT authorship contribution statement

Yueli Yao: Conceptualization, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft, Writing – review & editing. **Alexandra Schneider:** Conceptualization, Supervision, Methodology, Writing – review & editing. **Kathrin Wolf:** Conceptualization, Supervision, Methodology, Writing – review & editing. **Siqi Zhang:** Conceptualization, Methodology, Writing – review & editing. **Rui Wang-Sattler:** Conceptualization, Writing – review & editing. **Annette Peters:** Conceptualization, Data curation, Funding acquisition, Supervision, Writing – review & editing. **Susanne Breitner:** Conceptualization, Supervision, Methodology, Writing – review & editing.

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.

Acknowledgments

We thank all participants for their long-term commitment to the KORA study, the staff for data collection and research data management and the members of the KORA Study Group (https://www.helmholtz-munich.de/en/epi/cohort/kora) who are responsible for the design and conduct of the study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2023.165780.

References

Amelio, I., Cutruzzolá, F., Antonov, A., Agostini, M., Melino, G., 2014. Serine and glycine metabolism in cancer. Trends Biochem. Sci. 39 (4), 191–198. https://doi.org/10.1016/j.tibs.2014.02.004.

Armstrong, B.G., 1998. Effect of measurement error on epidemiological studies of environmental and occupational exposures. Occup. Environ. Med. 55 (10), 651–656. https://doi.org/10.1136/oem.55.10.651.

Ayala, A., Muñoz, M.F., Argüelles, S., 2014. Lipid peroxidation: production, metabolism, and signaling mechanisms of malondialdehyde and 4-hydroxy-2-nonenal. Oxidative Med. Cell. Longev. 2014, 360438. https://doi.org/10.1155/2014/360438.

Breitner, S., Schneider, A., Devlin, R.B., Ward-Caviness, C.K., Diaz-Sanchez, D., et al., 2016. Associations among plasma metabolite levels and short-term exposure to $PM_{2.5}$ and ozone in a cardiac catheterization cohort. Environ. Int. 97, 76–84. https://doi.org/10.1016/j.envint.2016.10.012.

Brook, R.D., Rajagopalan, S., 2009. Particulate matter, air pollution, and blood pressure. J. Am. Soc. Hyperten. 3 (5), 332–350. https://doi.org/10.1016/j.jash.2009.08.005.

Brook, R.D., Rajagopalan, S., Pope, C.A., Brook, J.R., Bhatnagar, A., et al., 2010. Particulate matter air pollution and cardiovascular disease. Circulation 121 (21), 2331–2378. https://doi.org/10.1161/CIR.0b013e3181dbece1.

Cao, Y.F., Li, J., Zhang, Z., Liu, J., Sun, X.Y., et al., 2019. Plasma levels of amino acids related to urea cycle and risk of type 2 diabetes mellitus in Chinese adults. Front. Endocrinol. 10, 50. https://doi.org/10.3389/fendo.2019.00050.

Chak, C.M., Lacruz, M.E., Adam, J., Brandmaier, S., Covic, M., et al., 2019. Ageing investigation using two-time-point metabolomics data from KORA and CARLA studies. Metabolites 9 (3). https://doi.org/10.3390/metabo9030044.

Chen, C., Li, H., Niu, Y., Liu, C., Lin, Z., et al., 2019. Impact of short-term exposure to fine particulate matter air pollution on urinary metabolome: a randomized, double-blind, crossover trial. Environ. Int. 130, 104878 https://doi.org/10.1016/j.envint.2019.05.072.

Chen, H., Wang, Z., Qin, M., Zhang, B., Lin, L., et al., 2021. Comprehensive metabolomics identified the prominent role of glycerophospholipid metabolism in coronary artery disease progression. Front. Mol. Biosci. 8, 632950. https://doi.org/10.3389/fmolb.2021.632950.

Chen, K., Schneider, A., Cyrys, J., Wolf, K., Meisinger, C., et al., 2020. Hourly exposure to ultrafine particle metrics and the onset of myocardial infarction in Augsburg, Germany. Environ. Health Perspect. 128 (1), 17003. https://doi.org/10.1289/EHP5478.

Chen, W.L., Lin, C.Y., Yan, Y.H., Cheng, K.T., Cheng, T.J., 2014. Alterations in rat pulmonary phosphatidylcholines after chronic exposure to ambient fine particulate matter. Mol. BioSyst. 10 (12), 3163–3169. https://doi.org/10.1039/C4MB00435C.

Cole, L.K., Vance, J.E., Vance, D.E., 2012. Phosphatidylcholine biosynthesis and lipoprotein metabolism. Biochim. Biophys. Acta Mol. Cell Biol. Lip. 1821 (5), 754–761. https://doi.org/10.1016/j.bbalip.2011.09.009.

Cyrys, J., Pitz, M., Heinrich, J., Wichmann, H.E., Peters, A., 2008. Spatial and temporal variation of particle number concentration in Augsburg, Germany. Sci. Total Environ. 401 (1–3), 168–175. https://doi.org/10.1016/j.scitotenv.2008.03.043.

Dauchet, L., Hulo, S., Cherot-Kornobis, N., Matran, R., Amouyel, P., et al., 2018. Short-term exposure to air pollution: associations with lung function and inflammatory markers in non-smoking, healthy adults. Environ. Int. 121, 610–619. https://doi.org/10.1016/j.envint.2018.09.036.

Deffner, V., Küchenhoff, H., Breitner, S., Schneider, A., Cyrys, J., et al., 2018. Mixtures of Berkson and classical covariate measurement error in the linear mixed model: bias analysis and application to a study on ultrafine particles. Biom. J. 60 (3), 480–497. https://doi.org/10.1002/bimj.201600188.

Eeftens, M., Tsai, M.-Y., Ampe, C., Anwander, B., Beelen, R., et al., 2012. Spatial variation of $PM_{2.5}$, PM_{10}, $PM_{2.5}$ absorbance and PM_{coarse} concentrations between and within 20 European study areas and the relationship with NO_2 — results of the ESCAPE project. Atmos. Environ. 62, 303–317. https://doi.org/10.1016/j.atmosenv.2012.08.038.

Feng, B., Liu, C., Yi, T., Song, X., Wang, Y., et al., 2021. Perturbation of amino acid metabolism mediates air pollution associated vascular dysfunction in healthy adults. Environ. Res. 201, 111512 https://doi.org/10.1016/j.envres.2021.111512.

Franchini, M., Mannucci, P.M., 2007. Short-term effects of air pollution on cardiovascular diseases: outcomes and mechanisms. J. Thromb. Haemost. 5 (11), 2169–2174. https://doi.org/10.1111/j.1538-7836.2007.02750.x.

Gaskins, A.J., Tang, Z., Hood, R.B., Ford, J., Schwartz, J.D., et al., 2021. Periconception air pollution, metabolomic biomarkers, and fertility among women undergoing assisted reproduction. Environ. Int. 155, 106666 https://doi.org/10.1016/j.envint.2021.106666.

Guo, Y., Logan, H.L., Glueck, D.H., Muller, K.E., 2013. Selecting a sample size for studies with repeated measures. BMC Med. Res. Methodol. 13 (1), 100. https://doi.org/10.1186/1471-2288-13-100.

Han, S., Huang, J., Foppiano, F., Prehn, C., Adamski, J., et al., 2022. TIGER: technical variation elimination for metabolomics data using ensemble learning architecture. Brief. Bioinform. 23 (2) https://doi.org/10.1093/bib/bbab535.

Holle, R., Happich, M., Löwel, H., Wichmann, H.E., 2005. KORA—a research platform for population based health research. Gesundheitswesen 67 (Suppl. 1), S19–S25. https://doi.org/10.1055/s-2005-858235.

Holmes, E., Wilson, I.D., Nicholson, J.K., 2008. Metabolic phenotyping in health and disease. Cell 134 (5), 714–717. https://doi.org/10.1016/j.cell.2008.08.026.

Hood, R.B., Liang, D., Tang, Z., Kloog, I., Schwartz, J., et al., 2022. Length of $PM_{2.5}$ exposure and alterations in the serum metabolome among women undergoing infertility treatment. Environ. Epidemiol. 6 (1), e191. https://doi.org/10.1097/EE9.0000000000000191.

Hu, X., Yan, M., He, L., Qiu, X., Zhang, J., et al., 2021. Associations between time-weighted personal air pollution exposure and amino acid metabolism in healthy adults. Environ. Int. 156, 106623 https://doi.org/10.1016/j.envint.2021.106623.

Huang, J., Huth, C., Covic, M., Troll, M., Adam, J., et al., 2020. Machine learning approaches reveal metabolic signatures of incident chronic kidney disease in individuals with prediabetes and type 2 diabetes. Diabetes 69 (12), 2756–2765. https://doi.org/10.2337/db20-0586.

Lee, H., Myung, W., Kim, D.K., Kim, S.E., Kim, C.T., et al., 2017. Short-term air pollution exposure aggravates Parkinson's disease in a population-based cohort. Sci. Rep. 7 (1), 44741. https://doi.org/10.1038/srep44741.

Li, H., Cai, J., Chen, R., Zhao, Z., Ying, Z., et al., 2017. Particulate matter exposure and stress hormone levels: a randomized, double-blind, crossover trial of air purification. Circulation 136 (7), 618–627. https://doi.org/10.1161/CIRCULATIONAHA.116.026796.

Li, M., Wu, Y., Ye, L., 2022a. The role of amino acids in endothelial biology and function. Cells 11 (8), 1372. https://doi.org/10.3390/cells11081372.

Li, S., Ma, Y., Ye, S., Guo, R., Su, Y., et al., 2022b. Ambient NO_2 exposure induced cardiotoxicity associated with gut microbiome dysregulation and glycerophospholipid metabolism disruption. Ecotoxicol. Environ. Saf. 238, 113583 https://doi.org/10.1016/j.ecoenv.2022.113583.

Li, Z., Liu, Q., Xu, Z., Guo, X., Wu, S., 2020. Association between short-term exposure to ambient particulate air pollution and biomarkers of oxidative stress: a meta-analysis. Environ. Res. 191, 110105 https://doi.org/10.1016/j.envres.2020.110105.

Li, Z., Liang, D., Ye, D., Chang, H.H., Ziegler, T.R., et al., 2021. Application of high-resolution metabolomics to identify biological pathways perturbed by traffic-related air pollution. Environ. Res. 193, 110506 https://doi.org/10.1016/j.envres.2020.110506.

Liang, D., Moutinho, J.L., Golan, R., Yu, T., Ladva, C.N., et al., 2018. Use of high-resolution metabolomics for the identification of metabolic signals associated with traffic-related air pollution. Environ. Int. 120, 145–154. https://doi.org/10.1016/j.envint.2018.07.044.

Liang, D., Ladva, C.N., Golan, R., Yu, T., Walker, D.I., et al., 2019. Perturbations of the arginine metabolome following exposures to traffic-related air pollution in a panel of commuters with and without asthma. Environ. Int. 127, 503–513. https://doi.org/10.1016/j.envint.2019.04.003.

Löwel, H., Döring, A., Schneider, A., Heier, M., Thorand, B., et al., 2005. The MONICA Augsburg surveys — basis for prospective cohort studies. Gesundheitswesen 67, 13–18. https://doi.org/10.1055/s-2005-858234.

Mangge, H., Stelzer, I., Reininghaus, E.Z., Weghuber, D., Postolache, T.T., et al., 2014. Disturbed tryptophan metabolism in cardiovascular disease. Curr. Med. Chem. 21 (17), 1931–1937. https://doi.org/10.2174/0929867321666140304105526.

Martínez, Y., Li, X., Liu, G., Bin, P., Yan, W., et al., 2017. The role of methionine on metabolism, oxidative stress, and diseases. Amino Acids 49 (12), 2091–2098. https://doi.org/10.1007/s00726-017-2494-2.

Monn, C., 2001. Exposure assessment of air pollutants: a review on spatial heterogeneity and indoor/outdoor/personal exposure to suspended particulate matter, nitrogen dioxide and ozone. Atmos. Environ. 35, 1–32. https://doi.org/10.1016/S1352-2310(00)00330-7.

Mu, L., Niu, Z., Blair, R.H., Yu, H., Browne, R.W., et al., 2019. Metabolomics profiling before, during, and after the Beijing Olympics: a panel study of within-individual differences during periods of high and low air pollution. Environ. Health Perspect. 127 (5), 57010. https://doi.org/10.1289/ehp3705.

Nassan, F.L., Kelly, R.S., Koutrakis, P., Vokonas, P.S., Lasky-Su, J.A., et al., 2021a. Metabolomic signatures of the short-term exposure to air pollution and temperature. Environ. Res. 201, 111553 https://doi.org/10.1016/j.envres.2021.111553.

Nassan, F.L., Wang, C., Kelly, R.S., Lasky-Su, J.A., Vokonas, P.S., et al., 2021b. Ambient $PM_{2.5}$ species and ultrafine particle exposure and their differential metabolomic signatures. Environ. Int. 151, 106447 https://doi.org/10.1016/j.envint.2021.106447.

Pang, Z., Chong, J., Zhou, G., de Lima Morais, D.A., Chang, L., et al., 2021. MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. Nucleic Acids Res. 49 (W1), W388–W396. https://doi.org/10.1093/nar/gkab382.

Pope 3rd, C.A., Muhlestein, J.B., May, H.T., Renlund, D.G., Anderson, J.L., et al., 2006. Ischemic heart disease events triggered by short-term exposure to fine particulate air pollution. Circulation 114 (23), 2443–2448. https://doi.org/10.1161/CIRCULATIONAHA.106.636977.

Popolo, A., Adesso, S., Pinto, A., Autore, G., Marzocco, S., 2014. L-Arginine and its metabolites in kidney and cardiovascular disease. Amino Acids 46 (10), 2271–2286. https://doi.org/10.1007/s00726-014-1825-9.

Proffitt, C., Bidkhori, G., Lee, S., Tebani, A., Mardinoglu, A., et al., 2022. Genome-scale metabolic modelling of the human gut microbiome reveals changes in the glyoxylate and dicarboxylate metabolism in metabolic disorders. iScience 25 (7), 104513. https://doi.org/10.1016/j.isci.2022.104513.

Rabel, M., Laxy, M., Thorand, B., Peters, A., Schwettmann, L., et al., 2018. Clustering of health-related behavior patterns and demographics. Results from the population-based KORA S4/F4 cohort study. Front. Publ. Health 6, 387. https://doi.org/10.3389/fpubh.2018.00387.

Rappaport, S.M., Barupal, D.K., Wishart, D., Vineis, P., Scalbert, A., 2014. The blood exposome and its role in discovering causes of disease. Environ. Health Perspect. 122 (8), 769–774. https://doi.org/10.1289/ehp.1308015.

Rice, M.B., Ljungman, P.L., Wilker, E.H., Gold, D.R., Schwartz, J.D., et al., 2013. Short-term exposure to air pollution and lung function in the Framingham Heart Study. Am. J. Respir. Crit. Care Med. 188 (11), 1351–1357. https://doi.org/10.1164/rccm.201308-1414OC.

Rinschen, M.M., Ivanisevic, J., Giera, M., Siuzdak, G., 2019. Identification of bioactive metabolites using activity metabolomics. Nat. Rev. Mol. Cell Biol. 20 (6), 353–367. https://doi.org/10.1038/s41580-019-0108-4.

Römisch-Margl, W., Prehn, C., Bogumil, R., Röhring, C., Suhre, K., et al., 2012. Procedure for tissue sample preparation and metabolite extraction for high-throughput targeted metabolomics. Metabolomics 8 (1), 133–142. https://doi.org/10.1007/s11306-011-0293-4.

Roorda-Knape, M.C., Janssen, N.A.H., De Hartog, J.J., Van Vliet, P.H.N., Harssema, H., et al., 1998. Air pollution from traffic in city districts near major motorways. Atmos. Environ. 32 (11), 1921–1930. https://doi.org/10.1016/S1352-2310(97)00496-2.

Surowiec, I., Karimpour, M., Gouveia-Figueira, S., Wu, J., Unosson, J., et al., 2016. Multi-platform metabolomics assays for human lung lavage fluids in an air pollution exposure study. Anal. Bioanal. Chem. 408 (17), 4751–4764. https://doi.org/10.1007/s00216-016-9566-0.

Tang, W.H., Wang, Z., Levison, B.S., Koeth, R.A., Britt, E.B., et al., 2013. Intestinal microbial metabolism of phosphatidylcholine and cardiovascular risk. N. Engl. J. Med. 368 (17), 1575–1584. https://doi.org/10.1056/NEJMoa1109400.

van der Veen, J.N., Kennelly, J.P., Wan, S., Vance, J.E., Vance, D.E., et al., 2017. The critical role of phosphatidylcholine and phosphatidylethanolamine metabolism in health and disease. Biochim. Biophys. Acta Biomembr. 1859 (9, Part B), 1558–1572. https://doi.org/10.1016/j.bbamem.2017.04.006.

van Veldhoven, K., Kiss, A., Keski-Rahkonen, P., Robinot, N., Scalbert, A., et al., 2019. Impact of short-term traffic-related air pollution on the metabolome — results from two metabolome-wide experimental studies. Environ. Int. 123, 124–131. https://doi.org/10.1016/j.envint.2018.11.034.

Vlaanderen, J.J., Janssen, N.A., Hoek, G., Keski-Rahkonen, P., Barupal, D.K., et al., 2017. The impact of ambient air pollution on the human blood metabolome. Environ. Res. 156, 341–348. https://doi.org/10.1016/j.envres.2017.03.042.

Wang, N., Mengersen, K., Tong, S., Kimlin, M., Zhou, M., et al., 2019. Short-term association between ambient air pollution and lung cancer mortality. Environ. Res. 179, 108748 https://doi.org/10.1016/j.envres.2019.108748.

Wang-Sattler, R., Yu, Z., Herder, C., Messias, A.C., Floegel, A., et al., 2012. Novel biomarkers for pre-diabetes identified by metabolomics. Mol. Syst. Biol. 8, 615. https://doi.org/10.1038/msb.2012.43.

Ward-Caviness, C.K., Breitner, S., Wolf, K., Cyrys, J., Kastenmuller, G., et al., 2016. Short-term NO_2 exposure is associated with long-chain fatty acids in prospective cohorts from Augsburg, Germany: results from an analysis of 138 metabolites and three exposures. Int. J. Epidemiol. 45 (5), 1528–1538. https://doi.org/10.1093/ije/dyw247.

Ward-Caviness, C.K., Xu, T., Aspelund, T., Thorand, B., Montrone, C., et al., 2017. Improvement of myocardial infarction risk prediction via inflammation-associated metabolite biomarkers. Heart 103 (16), 1278–1285. https://doi.org/10.1136/heartjnl-2016-310789.

Wen, T., Liao, D., Wellenius, G.A., Whitsel, E.A., Margolis, H.G., et al., 2023. Short-term air pollution levels and blood pressure in older women. Epidemiology 34 (2), 271–281. https://doi.org/10.1097/ede.0000000000001577.

Weuve, J., Tchetgen Tchetgen, E.J., Glymour, M.M., Beck, T.L., Aggarwal, N.T., et al., 2012. Accounting for bias due to selective attrition: the example of smoking and cognitive decline. Epidemiology 23 (1), 119–128. https://doi.org/10.1097/EDE.0b013e318230e861.

Whiley, L., Sen, A., Heaton, J., Proitsi, P., García-Gómez, D., et al., 2014. Evidence of altered phosphatidylcholine metabolism in Alzheimer's disease. Neurobiol. Aging 35 (2), 271–278. https://doi.org/10.1016/j.neurobiolaging.2013.08.001.

Wieder, C., Frainay, C., Poupin, N., Rodríguez-Mier, P., Vinson, F., et al., 2021. Pathway analysis in metabolomics: recommendations for the use of over-representation analysis. PLoS Comput. Biol. 17 (9), e1009105 https://doi.org/10.1371/journal.pcbi.1009105.

Winkler, G., Döring, A., 1998. Validation of a short qualitative food frequency list used in several German large scale surveys. Z. Ernahrungswiss. 37 (3), 234–241. https://doi.org/10.1007/PL00007377.

Wolf, K., Schneider, A., Breitner, S., Meisinger, C., Heier, M., et al., 2015. Associations between short-term exposure to particulate matter and ultrafine particles and myocardial infarction in Augsburg, Germany. Int. J. Hyg. Environ. Health 218 (6), 535–542. https://doi.org/10.1016/j.ijheh.2015.05.002.

Wolf, K., Cyrys, J., Harciníková, T., Gu, J., Kusch, T., et al., 2017. Land use regression modeling of ultrafine particles, ozone, nitrogen oxides and markers of particulate matter pollution in Augsburg, Germany. Sci. Total Environ. 579, 1531–1540. https://doi.org/10.1016/j.scitotenv.2016.11.160.

Wu, S., Ni, Y., Li, H., Pan, L., Yang, D., et al., 2016. Short-term exposure to high ambient air pollution increases airway inflammation and respiratory symptoms in chronic obstructive pulmonary disease patients in Beijing, China. Environ. Int. 94, 76–82. https://doi.org/10.1016/j.envint.2016.05.004.

Xu, T., Holzapfel, C., Dong, X., Bader, E., Yu, Z., et al., 2013. Effects of smoking and smoking cessation on human serum metabolite profile: results from the KORA cohort study. BMC Med. 11 (1), 60. https://doi.org/10.1186/1741-7015-11-60.

Yan, Q., Liew, Z., Uppal, K., Cui, X., Ling, C., et al., 2019. Maternal serum metabolome and traffic-related air pollution exposure in pregnancy. Environ. Int. 130, 104872 https://doi.org/10.1016/j.envint.2019.05.066.

Yao, Y., Schneider, A., Wolf, K., Zhang, S., Wang-Sattler, R., et al., 2022. Longitudinal associations between metabolites and long-term exposure to ambient air pollution: results from the KORA cohort study. Environ. Int. 170, 107632 https://doi.org/10.1016/j.envint.2022.107632.

Zeger, S.L., Thomas, D., Dominici, F., Samet, J.M., Schwartz, J., et al., 2000. Exposure measurement error in time-series studies of air pollution: concepts and consequences. Environ. Health Perspect. 108 (5), 419–426. https://doi.org/10.1289/ehp.00108419.

Zhang, J., Liang, S., Ning, R., Jiang, J., Zhang, J., et al., 2020. $PM_{2.5}$-induced inflammation and lipidome alteration associated with the development of atherosclerosis based on a targeted lipidomic analysis. Environ. Int. 136, 105444 https://doi.org/10.1016/j.envint.2019.105444.

Zhang, Y., Ke, L., Fu, Y., Di, Q., Ma, X., 2022. Physical activity attenuates negative effects of short-term exposure to ambient air pollution on cognitive function. Environ. Int. 160, 107070 https://doi.org/10.1016/j.envint.2021.107070.

Zhu, K., Browne, R.W., Blair, R.H., Bonner, M.R., Tian, M., et al., 2021. Changes in arachidonic acid (AA)- and linoleic acid (LA)-derived hydroxy metabolites and their interplay with inflammatory biomarkers in response to drastic changes in air pollution exposure. Environ. Res. 200, 111401 https://doi.org/10.1016/j.envres.2021.111401.

Science of the Total Environment 905 (2023) 167050

Contents lists available at ScienceDirect

Science of the Total Environment

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

Corrigendum

Corrigendum to "Longitudinal associations between metabolites and immediate, short- and medium-term exposure to ambient air pollution: Results from the KORA cohort study" [Sci. Total Environ. 900 (2023) 165780]

Yueli Yao [a,b,*], Alexandra Schneider [a], Kathrin Wolf [a], Siqi Zhang [a], Rui Wang-Sattler [a,c], Cornelia Prehn [d], Jerzy Adamski [e,f,g], Annette Peters [a,b,c,h,1], Susanne Breitner [a,b,1]

[a] Institute of Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany
[b] Institute for Medical Information Processing, Biometry, and Epidemiology (IBE), Faculty of Medicine, Pettenkofer School of Public Health, LMU Munich, Munich, Germany
[c] German Center for Diabetes Research (DZD), Munich-Neuherberg, Germany
[d] Metabolomics and Proteomics Core, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany
[e] Institute of Experimental Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany
[f] Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore, 8 Medical Drive, Singapore 117597, Singapore
[g] Institute of Biochemistry, Faculty of Medicine, University of Ljubljana, Vrazov trg 2, 1000 Ljubljana, Slovenia
[h] German Centre for Cardiovascular Research (DZHK), Partner Site Munich, Munich, Germany

The authors regret that the printed version of the above article contained a number of errors. The corrigendum is only given to the additional inclusion of another two co-authors on the author list in this paper. All data, results, and discussion of this study are not affected by the corrigendum in comparison to the original version of our manuscript. The correct and final version follows. The authors would like to apologise for any inconvenience caused.

Longitudinal associations between metabolites and immediate, short- and medium-term exposure to ambient air pollution: Results from the KORA cohort study

Yueli Yao [a,b], Alexandra Schneider [a], Kathrin Wolf [a], Siqi Zhang [a], Rui Wang-Sattler [a,c], Cornelia Prehn [d], Jerzy Adamski [e,f,g], Annette Peters [a,b,c,h,1], Susanne Breitner [a,b,1]

[a] Institute of Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany

[b] Institute for Medical Information Processing, Biometry, and Epidemiology (IBE), Faculty of Medicine, Pettenkofer School of Public Health, LMU Munich, Munich, Germany

[c] German Center for Diabetes Research (DZD), Munich-Neuherberg, Germany

[d] Metabolomics and Proteomics Core, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany

[e] Institute of Experimental Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany

[f] Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore, 8 Medical Drive, Singapore 117597, Singapore

[g] Institute of Biochemistry, Faculty of Medicine, University of Ljubljana, Vrazov trg 2, 1000 Ljubljana, Slovenia

[h] German Centre for Cardiovascular Research (DZHK), Partner Site Munich, Munich, Germany

[1] Authors share last authorship.

CRediT authorship contribution statement

Yueli Yao: Conceptualization, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft, Writing – review & editing. **Alexandra Schneider:** Conceptualization, Supervision, Methodology, Writing – review & editing. **Kathrin Wolf:** Conceptualization, Supervision, Methodology, Writing – review & editing. **Siqi Zhang:** Conceptualization, Methodology, Writing – review & editing. **Rui Wang-Sattler:** Conceptualization, Writing – review & editing. **Cornelia Prehn:** Resources, Data curation. **Jerzy Adamski:** Resources, Data curation. **Annette Peters:** Conceptualization, Data curation, Funding acquisition, Supervision, Writing – review & editing. **Susanne Breitner:** Conceptualization, Supervision, Methodology, Writing – review & editing.

DOI of original article: https://doi.org/10.1016/j.scitotenv.2023.165780.

* Corresponding author at: Institute of Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, D-85764 Neuherberg, Germany.

E-mail address: yueli.yao@helmholtz-munich.de (Y. Yao).

[1] Authors share last authorship.

https://doi.org/10.1016/j.scitotenv.2023.167050

Available online 19 September 2023

Longitudinal associations between metabolites and immediate, short- and medium-term exposure to ambient air pollution: results from the KORA cohort study

Yueli Yao [a,b], Alexandra Schneider [a], Kathrin Wolf [a], Siqi Zhang [a], Rui Wang-Sattler [a,c], Annette Peters [a,b,c,d]*, Susanne Breitner [a,b]*

[a] Institute of Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany

[b] Institute for Medical Information Processing, Biometry, and Epidemiology (IBE), Faculty of Medicine, LMU Munich, Pettenkofer School of Public Health, Munich, Germany

[c] German Center for Diabetes Research (DZD), Munich-Neuherberg, Germany

[d] German Centre for Cardiovascular Research (DZHK), Partner Site Munich, Munich, Germany

*Authors share last authorship

Correspondence to:

Yueli Yao, Institute of Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, D-85764 Neuherberg, Germany; Email: yueli.yao@helmholtz-munich.de.

Supplementary Materials

Contents

Table S1 Biochemical names and categories of the 108 investigated metabolites.

	Abbreviation	Biochemical name
Acylcarnitine	C0	Carnitine
	C2	Acetylcarnitine
	C3	Propionylcarnitine
	C4	Butyrylcarnitine
	C10	Decanoylcarnitine
	C12	Dodecanoylcarnitine
	C14:1	Tetradecenoylcarnitine
	C14:2	Tetradecadienylcarnitine
	C16	Hexadecanylcarnitine
	C18	Octadecanoylcarnitine
	C18:1	Octadecenoylcarnitine
	C18:2	Octadecadienylcarnitine
Amino acids	Arg	Arginine
	Gln	Glutamine
	Gly	Glycine
	His	Histidine
	Met	Methionine
	Orn	Ornithine
	Phe	Phenylalanine
	Pro	Proline
	Ser	Serine
	Thr	Threonine
	Trp	Tryptophan
	Tyr	Tryosine
Phosphatidylcholines	PC aa C28:1	Phosphatidylcholine diacyl C28:1
	PC aa C30:0	Phosphatidylcholine diacyl C30:0
	PC aa C32:0	Phosphatidylcholine diacyl C32:0
	PC aa C32:1	Phosphatidylcholine diacyl C32:1
	PC aa C32:2	Phosphatidylcholine diacyl C32:2
	PC aa C32:3	Phosphatidylcholine diacyl C32:3
	PC aa C34:1	Phosphatidylcholine diacyl C34:1
	PC aa C34:2	Phosphatidylcholine diacyl C34:2
	PC aa C34:3	Phosphatidylcholine diacyl C34:3
	PC aa C34:4	Phosphatidylcholine diacyl C34:4
	PC aa C36:1	Phosphatidylcholine diacyl C36:1
	PC aa C36:2	Phosphatidylcholine diacyl C36:2
	PC aa C36:3	Phosphatidylcholine diacyl C36:3
	PC aa C36:4	Phosphatidylcholine diacyl C36:4
	PC aa C36:5	Phosphatidylcholine diacyl C36:5
	PC aa C36:6	Phosphatidylcholine diacyl C36:6
	PC aa C38:0	Phosphatidylcholine diacyl C38:0
	PC aa C38:3	Phosphatidylcholine diacyl C38:3
	PC aa C38:4	Phosphatidylcholine diacyl C38:4
	PC aa C38:5	Phosphatidylcholine diacyl C38:5
	PC aa C38:6	Phosphatidylcholine diacyl C38:6
	PC aa C40:2	Phosphatidylcholine diacyl C40:2
	PC aa C40:3	Phosphatidylcholine diacyl C40:3
	PC aa C40:4	Phosphatidylcholine diacyl C40:4

	PC aa C40:5	Phosphatidylcholine diacyl C40:5
	PC aa C40:6	Phosphatidylcholine diacyl C40:6
	PC aa C42:0	Phosphatidylcholine diacyl C42:0
	PC aa C42:1	Phosphatidylcholine diacyl C42:1
	PC aa C42:2	Phosphatidylcholine diacyl C42:2
	PC aa C42:4	Phosphatidylcholine diacyl C42:4
	PC aa C42:5	Phosphatidylcholine diacyl C42:5
	PC aa C42:6	Phosphatidylcholine diacyl C42:6
	PC ae C32:1	Phosphatidylcholine acyl-alkyl C32:1
	PC ae C32:2	Phosphatidylcholine acyl-alkyl C32:2
	PC ae C34:0	Phosphatidylcholine acyl-alkyl C34:0
	PC ae C34:1	Phosphatidylcholine acyl-alkyl C34:1
	PC ae C34:2	Phosphatidylcholine acyl-alkyl C34:2
	PC ae C34:3	Phosphatidylcholine acyl-alkyl C34:3
	PC ae C36:1	Phosphatidylcholine acyl-alkyl C36:1
	PC ae C36:2	Phosphatidylcholine acyl-alkyl C36:2
	PC ae C36:3	Phosphatidylcholine acyl-alkyl C36:3
	PC ae C36:4	Phosphatidylcholine acyl-alkyl C36:4
	PC ae C36:5	Phosphatidylcholine acyl-alkyl C36:5
	PC ae C38:0	Phosphatidylcholine acyl-alkyl C38:0
	PC ae C38:1	Phosphatidylcholine acyl-alkyl C38:1
	PC ae C38:2	Phosphatidylcholine acyl-alkyl C38:2
	PC ae C38:3	Phosphatidylcholine acyl-alkyl C38:3
	PC ae C38:4	Phosphatidylcholine acyl-alkyl C38:4
	PC ae C38:5	Phosphatidylcholine acyl-alkyl C38:5
	PC ae C38:6	Phosphatidylcholine acyl-alkyl C38:6
	PC ae C40:1	Phosphatidylcholine acyl-alkyl C40:1
	PC ae C40:2	Phosphatidylcholine acyl-alkyl C40:2
	PC ae C40:3	Phosphatidylcholine acyl-alkyl C40:3
	PC ae C40:4	Phosphatidylcholine acyl-alkyl C40:4
	PC ae C40:5	Phosphatidylcholine acyl-alkyl C40:5
	PC ae C40:6	Phosphatidylcholine acyl-alkyl C40:6
	PC ae C42:1	Phosphatidylcholine acyl-alkyl C42:1
	PC ae C42:2	Phosphatidylcholine acyl-alkyl C42:2
	PC ae C42:3	Phosphatidylcholine acyl-alkyl C42:3
	PC ae C42:4	Phosphatidylcholine acyl-alkyl C42:4
	PC ae C42:5	Phosphatidylcholine acyl-alkyl C42:5
	PC ae C44:3	Phosphatidylcholine acyl-alkyl C44:3
	PC ae C44:4	Phosphatidylcholine acyl-alkyl C44:4
	PC ae C44:5	Phosphatidylcholine acyl-alkyl C44:5
	PC ae C44:6	Phosphatidylcholine acyl-alkyl C44:6
Lysophosphatidylcholines	lysoPC a C16:0	lysoPhosphatidylcholine acyl C16:0
	lysoPC a C16:1	lysoPhosphatidylcholine acyl C16:1
	lysoPC a C18:0	lysoPhosphatidylcholine acyl C18:0
	lysoPC a C18:1	lysoPhosphatidylcholine acyl C18:1
	lysoPC a C18:2	lysoPhosphatidylcholine acyl C18:2
	lysoPC a C20:3	lysoPhosphatidylcholine acyl C20:3
	lysoPC a C20:4	lysoPhosphatidylcholine acyl C20:4
Sphingomyelins	SM (OH) C14:1	Hydroxysphingomyeline C14:1
	SM (OH) C16:1	Hydroxysphingomyeline C16:1
	SM (OH) C22:1	Hydroxysphingomyeline C22:1

	SM (OH) C22:2	Hydroxysphingomyeline C22:2
	SM (OH)C24:1	Hydroxysphingomyeline C24:1
	SM C16:0	Sphingomyeline C16:0
	SM C16:1	Sphingomyeline C16:1
	SM C18:0	Sphingomyeline C18:0
	SM C18:1	Sphingomyeline C18:1
	SM C20:2	Sphingomyeline C20:2
	SM C24:1	Sphingomyeline C24:1
Hexose	H1	Hexose

Table S2 The explained variance (R^2) between main monitor and the other monitors of PM_{10} and NO_2.

year		R^2 (with monitor LfU)	R^2 (with monitor BP)		R^2 (with monitor LfU)
2001		---	---		0.70
2002		---	---		0.60
2003		---	---		0.70
2004		0.97	0.97		0.78
2005		0.94	0.91		0.76
2006		0.98	0.98		0.69
2007	PM_{10}	0.96	0.97	NO_2	0.78
2008	(monitor FH)	0.96	0.96	(monitor BP)	0.78
2009		0.97	0.94		0.85
2010		0.90	0.84		0.57
2011		0.94	0.78		0.71
2012		0.80	0.74		0.92
2013		0.76	---		0.92
2014		0.95	0.92		0.76

FH: an aerosol monitoring station located approximately 1 km southeast of the city center of Augsburg with a distance of 100 meters to the nearest major road;

LfU: a single urban background monitoring site operated by the Bavarian Environment Agency, approximately 4 km south of the city center;

BP: an urban background station (Bourgesplatz, BP) located approximately 2 km north of the city center, as well as 20 meters to the nearest road with low traffic intensity and 100 meters to the nearest road with high traffic intensity.

Table S3 Percent changes (95% CIs) in repeated measurements of metabolites per interquartile (IQR) increase in air pollutant concentrations at Bonferroni adjusted *p*-value levels.

air pollutants	metabolites	Exposure windows		
		2-day moving average	2-week moving window	8-week moving average
$PM_{2.5}$	Arg	-0.2 (-0.8, 0.3)	-1.1 (-1.9, -0.3)	-3.8 (-5.3, -2.2)
	Trp	-0.7 (-1.2, -0.2)	-1.6 (-2.3, -0.9)	-3.9 (-5.2, -2.5)
	PC aa C30:0	-1 (-1.8, -0.1)	-2.7 (-4.0, -1.4)	-4.1 (-5.8, -2.3)
	PC aa C40:3	-0.3 (-1, 0.3)	-0.4 (-1.3, 0.5)	-4.5 (-6.3, -2.6)
	PC aa C42:4	-0.3 (-1, 0.3)	-0.7 (-1.6, 0.1)	-4.2 (-5.9, -2.5)
	PC ae C38:1	1.6 (-0.6, 4)	1.5 (-1.6, 4.8)	-17 (-22.2, -11.4)
	PC ae C38:2	-0.2 (-1.1, 0.6)	-2 (-3.1, -0.9)	-8.8 (-11, -6.5)
	PC ae C40:3	-0.2 (-0.7, 0.4)	-0.8 (-1.5, 0)	-5.6 (-7.2, -4.1)
	PC ae C42:1	0 (-0.7, 0.7)	0.3 (-0.6, 1.2)	-4.7 (-6.6, -2.8)
	PC ae C44:3	0.3 (-0.5, 1.1)	0.5 (-0.6, 1.6)	-6.7 (-8.8, -4.6)
PM_{coarse}	Gly	0.7 (0, 1.5)	1.8 (0.6, 3)	5.2 (3.3, 7.2)
	Met	0 (-0.7, 0.7)	0.2 (-0.8, 1.3)	4.1 (2.3, 5.9)
	Orn	1.2 (0.5, 2)	2.2 (1, 3.4)	8 (5.9, 10.1)
	Phe	0.8 (0.3, 1.4)	0.6 (-0.3, 1.6)	4.8 (3.2, 6.4)
	Ser	0.7 (0, 1.4)	0.7 (-0.4, 1.8)	4.4 (2.5, 6.3)
	Thr	0.4 (-0.6, 1.4)	-0.4 (-2, 1.2)	5.9 (3.2, 8.7)
	lysoPC a C16:0	0.2 (-0.5, 1)	1 (-0.1, 2.1)	4.6 (2.6, 6.5)
	PC aa C40:4	-1.9 (-2.7, -1.0)	-1.3 (-2.6, -0.1)	0.6 (-1.4, 2.7)
	PC ae C40:4	-0.6 (-1.2, 0.1)	0.4 (-0.5, 1.4)	3.7 (2, 5.5)
	PC ae C44:3	1 (0.1, 1.9)	3.8 (2.3, 5.3)	0.7 (-1.5, 3)
	SM C16:0	-0.1 (-0.7, 0.4)	1.5 (0.7, 2.2)	3.5 (2.1, 5.0)
NO_2	C0	-0.9 (-1.6, -0.1)	-1.5 (-2.5, -0.6)	-3.5 (-4.9, -2)
	Arg	-0.4 (-1.1, 0.4)	-1.5 (-2.5, -0.6)	-4.1 (-5.5, -2.7)
	Trp	-1.4 (-2.1, -0.9)	-2.5 (-3.3, -1.6)	-5.2 (-6.4, -3.9)
	Tyr	-1.1 (-1.9, -0.3)	-2.3 (-3.4, -1.3)	-3.8 (-5.3, -2.2)
	lysoPC a C16:0	-1.4 (-2.2, -0.6)	-2.2 (-3.3, -1.2)	-3.9 (-5.4, -2.3)
	lysoPC a C16:1	-2 (-3.1, -0.9)	-3.1 (-4.5, -1.7)	-5.1 (-7.1, -2.9)
	lysoPC a C18:0	-1.1 (-2.1, -0.2)	-1.8 (-2.9, -0.6)	-4.5 (-6.3, -2.7)
	lysoPC a C18:1	-2.1 (-3.2, -1.1)	-2.9 (-4.2, -1.6)	-4.9 (-6.9, -3)
	lysoPC a C20:3	-1.3 (-2.4, -0.2)	-3.1 (-4.4, -1.6)	-4.1 (-6.1, -2)
	lysoPC a C20:4	-1.9 (-2.9, -0.9)	-2.6 (-3.9, -1.3)	-4.8 (-6.8, -2.8)
	PC aa C28:1	-1 (-1.8, -0.2)	-1.3 (-2.2, -0.3)	-3.6 (-5.2, -1.9)
	PC aa C30:0	-2.5 (-3.6, -1.3)	-3.6 (-5.1, -2.1)	-6.3 (-8.5, -4)
	PC aa C32:0	-1.6 (-2.3, -0.8)	-1.4 (-2.3, -0.6)	-5.5 (-7, -4)
	PC aa C34:1	-1.8 (-2.7, -1.0)	-2 (-3.0, -1.0)	-4.3 (-5.9, -2.7)
	PC aa C36:1	-1.8 (-2.8, -0.9)	-3.1 (-4.3, -1.8)	-5.6 (-7.4, -3.7)
	PC aa C36:2	-0.6 (-1.4, 0.1)	-1.1 (-1.9, -0.2)	-3.5 (-5, -2)
	PC aa C36:3	-1.1 (-1.9, -0.3)	-1.6 (-2.5, -0.7)	-3.8 (-5.4, -2.2)
	PC aa C38:3	-1.2 (-2.1, -0.3)	-2.7 (-3.9, -1.5)	-4.6 (-6.4, -2.9)
	PC aa C40:2	-1 (-2.1, 0)	-0.2 (-1.4, 1.1)	-4.4 (-6.4, -2.4)
	PC aa C40:3	-1.1 (-2, -0.2)	-0.5 (-1.6, 0.6)	-5 (-6.7, -3.3)
	PC aa C40:4	-2.1 (-3.1, -1.1)	-3 (-4.2, -1.7)	-5.3 (-7.1, -3.4)
	PC aa C40:5	-1.5 (-2.5, -0.5)	-1.8 (-2.9, -0.6)	-4.5 (-6.4, -2.6)
	PC aa C42:4	-1.3 (-2.1, -0.4)	-1.6 (-2.6, -0.6)	-4.6 (-6.1, -3)
	PC aa C42:5	-2.1 (-3.1, -1)	-2.4 (-3.7, -1)	-5.5 (-7.6, -3.4)
	PC ae C32:2	-1.2 (-2, -0.4)	-1.7 (-2.6, -0.8)	-3.7 (-5.3, -2.1)

	PC ae C34:0	-1.6 (-2.6, -0.7)	-1.6 (-2.7, -0.4)	-4.3 (-6.1, -2.4)
	PC ae C34:1	-1.5 (-2.3, -0.7)	-1.8 (-2.8, -0.9)	-4.4 (-5.9, -2.9)
	PC ae C34:2	-0.9 (-1.9, 0)	-2 (-3.1, -0.9)	-4.3 (-6.1, -2.5)
	PC ae C36:1	-1.2 (-2.1, -0.4)	-1.3 (-2.3, -0.3)	-3.8 (-5.5, -2.2)
	PC ae C36:2	-0.8 (-1.7, 0)	-1.1 (-2.1, -0.1)	-3.7 (-5.3, -2)
	PC ae C36:3	-1.2 (-2.1, -0.3)	-2 (-3.1, -1)	-4.4 (-6.1, -2.6)
	PC ae C36:4	-1.5 (-2.5, -0.6)	-1.3 (-2.4, -0.1)	-4.1 (-5.9, -2.2)
	PC ae C38:1	-1.6 (-4.5, 1.4)	-6.4 (-10, -2.7)	-13.9 (-18.9, -8.6)
	PC ae C38:2	-1.1 (-2.2, 0.1)	-4.4 (-5.8, -3)	-12 (-14, -10)
	PC ae C38:4	-1.3 (-2.1, -0.5)	-1.5 (-2.5, -0.6)	-4 (-5.5, -2.5)
	PC ae C38:5	-1.6 (-2.4, -0.9)	-1.8 (-2.7, -0.8)	-4.2 (-5.7, -2.6)
	PC ae C40:3	-0.9 (-1.7, -0.1)	-1.7 (-2.6, -0.8)	-5.6 (-7.1, -4.1)
	PC ae C40:4	-1.5 (-2.2, -0.7)	-2.3 (-3.2, -1.3)	-4.5 (-5.9, -3.1)
	PC ae C40:5	-1.4 (-2.1, -0.7)	-1.4 (-2.3, -0.5)	-3.7 (-5.1, -2.3)
	PC ae C42:2	-1.2 (-2, -0.3)	-0.7 (-1.8, 0.3)	-4.1 (-5.7, -2.3)
	PC ae C42:3	-1.4 (-2.3, -0.6)	-1.6 (-2.6, -0.5)	-5.2 (-6.8, -3.5)
	PC ae C42:4	-1.5 (-2.4, -0.6)	-2.0 (-3.1, -0.9)	-4.3 (-5.9, -2.6)
	PC ae C42:5	-1.6 (-2.2, -0.9)	-2.3 (-3.2, -1.5)	-4.2 (-5.4, -2.8)
	SM (OH) C14:1	-0.7 (-1.5, 0.1)	-1.1 (-2.1, -0.2)	-3.8 (-5.3, -2.2)
	SM (OH) C16:1	-0.8 (-1.6, 0.1)	-1.2 (-2.2, -0.2)	-3.7 (-5.3, -2.1)
	SM C16:0	-1 (-1.6, -0.4)	-1 (-1.7, -0.2)	-3.9 (-5.1, -2.7)
	SM C16:1	-0.9 (-1.6, -0.2)	-1.1 (-1.8, -0.3)	-3.4 (-4.7, -2.1)
	SM C18:0	-0.6 (-1.4, 0.1)	-1 (-1.9, -0.1)	-3.9 (-5.3, -2.4)
	SM C18:1	-0.6 (-1.4, 0.2)	-1.3 (-2.3, -0.4)	-4.1 (-5.6, -2.6)
	SM C24:1	-1.3 (-2, -0.6)	-1 (-1.8, -0.1)	-3.7 (-5.1, -2.3)
O_3	Trp	1.8 (0.5, 3.1)	6 (3.4, 8.6)	14.8 (9.1, 20.7)
	SM (OH) C24:1	3.7 (2, 5.5)	4.3 (1.3, 7.5)	12.4 (5.6, 19.5)

The results were derived from the main models adjusted for age, sex, body mass index (BMI), an indicator of each study wave (KORA S4, KORA F4, or KORA FF4), day of the week, time trend, temperature, relative humidity, smoking status, alcohol consumption, physical activity index, educational attainment, fasting status, and dietary score. Three degrees of freedom were used for the temperature and relative humidity splines, and four degrees of freedom per year in each study wave were used for the time trend spline. The metabolites concentrations were natural log-transformed, and the effect estimates were represented as the percent changes in the geometric mean of metabolites per IQR increase in air pollutants concentrations.

$PM_{2.5}$ = particulate matter with an aerodynamic diameter less than or equal to 2.5 μm; PM_{coarse} = particulate matter with an aerodynamic diameter of 2.5-10 μm; NO_2 = nitrogen dioxide; O_3 = ozone.

The Bonferroni *p*-value cut-off is: 3.9×10^{-5}. The effect estimates listed in blue colour were not significantly associated air pollutants at the Bonferroni *p*-value cut-off.

Table S4 Results comparison between main analysis and sensitivity analysis ($PM_{2.5}$ exposure)

	2-day moving average								2-week moving average								8-week moving average					
	M	S1	S2	S3	S4	S5	S6	S7	M	S1	S2	S3	S4	S5	S6	S7	M	S1	S2	S3	S6	S7
C16																						
Arg																						
Phe																						
Ser																						
Trp																						
PC aa C30:0																						
PC aa C36:1																						
PC aa C38:3																						
PC aa C40:3																						
PC aa C40:4																						
PC aa C42:4																						
PC ae C38:1																						
PC ae C38:2																						
PC ae C38:3																						
PC ae C40:3																						
PC ae C40:4																						
PC ae C42:1																						
PC ae C42:2																						
PC ae C42:3																						
PC ae C42:5																						
PC ae C44:3																						
SM (OH) C14:1																						
SM (OH) C22:2																						

(1) **M:** results from main analysis adjusted for age, sex, body mass index (BMI), an indicator of each study wave (KORA S4, KORA F4, or KORA FF4), day of the week, time trend, temperature, relative humidity, smoking status, alcohol consumption, physical activity index, educational attainment, fasting status, and dietary score. Three degrees of freedom were used for the temperature and relative humidity splines, and four degrees of freedom per year in each study wave were used for the time trend spline.
S1: restricted to all participants attended KORA S4, KORAF4 or KORAFF4 with metabolomics data.
S2: restricted to fasting participants throughout the entire study period.
S3: IPW analysis.
S4: immediate (2-day moving average) and short-term (2-week moving average) analysis results from crude model adjusted for an indicator of each study wave (KORA S4, KORA F4, or KORA FF4), day of the week, time trend, temperature, relative humidity. Three degrees of freedom were used for the temperature and relative humidity splines, and four degrees of freedom per year in each study wave were used for the time trend spline.
S5: co-effect between immediate (2-day moving average) and short-term (2-week moving average) exposure and corresponding long-term exposure (annual average) to air pollutant.
S6: two-pollutant models additionally adjusted for PM_{coarse}.
S7: two-pollutant models additionally adjusted for O_3.
(2) Blue means negative associations between metabolites and air pollutants, while red means positive association between metabolites and air pollutants.

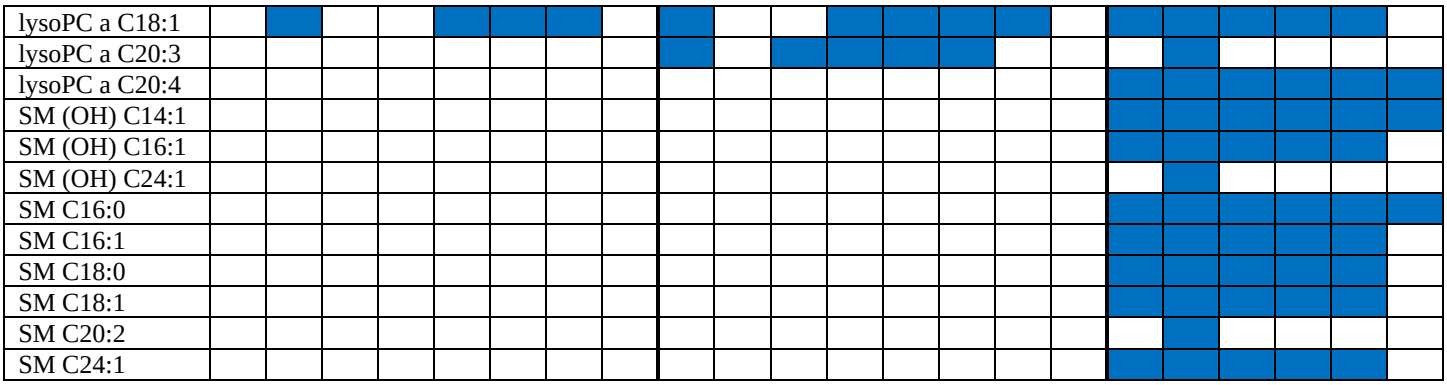
Table S5 Results comparison between main analysis and sensitivity analysis (PM_{coarse} exposure)

	2-day moving average							2-week moving average							8-week moving average				
	M	S1	S2	S3	S4	S5	S6	M	S1	S2	S3	S4	S5	S6	M	S1	S2	S3	S6
C16																			
C18:2																			
Gly																			
His																			
Met																			
Orn																			
Phe																			
Ser																			
Thr																			
Tyr																			
PC aa C28:1																			
PC aa C40:3																			
PC aa C40:4																			
PC aa C42:0																			
PC aa C42:4																			
PC ae C40:1																			
PC ae C40:2																			
PC ae C40:3																			
PC ae C40:4																			
PC ae C40:5																			
PC ae C42:5																			
PC ae C44:3																			
lysoPC a C16:0																			
lysoPC a C18:0																			
SM (OH) C14:1																			
SM (OH) C22:2																			
SM C16:0																			

(1) **M:** results from main analysis adjusted for age, sex, body mass index (BMI), an indicator of each study wave (KORA S4, KORA F4, or KORA FF4), day of the week, time trend, temperature, relative humidity, smoking status, alcohol consumption, physical activity index, educational attainment, fasting status, and dietary score. Three degrees of freedom were used for the temperature and relative humidity splines, and four degrees of freedom per year in each study wave were used for the time trend spline.
S1: restricted to all participants attended KORA S4, KORAF4 or KORAFF4 with metabolomics data.
S2: restricted to fasting participants throughout the entire study period.
S3: IPW analysis.
S4: short-term (lag0 and 2-week moving average, respectively) analysis results from crude model adjusted for an indicator of each study wave (KORA S4, KORA F4, or KORA FF4), day of the week, time trend, temperature, relative humidity. Three degrees of freedom were used for the temperature and relative humidity splines, and four degrees of freedom per year in each study wave were used for the time trend spline.
S5: co-effect between immediate (2-day moving average) and short-term (2-week moving average) exposure and corresponding long-term exposure (annual average) to air pollutant.
S6: two-pollutant models additionally adjusted for O_3.
(2) Blue means negative associations between metabolites and air pollutants, while red means positive association between metabolites and air pollutants.

Table S6 Results comparison between main analysis and sensitivity analysis (NO_2 exposure)

	2-day moving average								2-week moving average								8-week moving average					
	M	S1	S2	S3	S4	S5	S6	S7	M	S1	S2	S3	S4	S5	S6	S7	M	S1	S2	S3	S6	S7
C0																	■	■	■	■	■	
Arg															■		■	■	■	■	■	■
Ser															■		■	■	■	■	■	■
Trp	■	■		■	■	■			■	■	■	■	■	■	■		■	■	■	■	■	■
Tyr		■							■	■		■	■	■	■		■	■	■	■	■	■
Phe		■																				■
PC aa C28:1																	■	■	■	■	■	
PC aa C30:0		■							■	■	■	■	■	■			■	■	■	■	■	
PC aa C32:0																	■	■	■	■	■	■
PC aa C34:1	■				■	■			■	■		■	■	■			■	■	■	■	■	
PC aa C36:1									■	■	■	■	■	■			■	■	■	■	■	■
PC aa C36:2																	■	■	■	■	■	
PC aa C36:3																	■	■	■	■	■	
PC aa C36:4																		■				
PC aa C38:3									■	■	■	■	■	■		■	■	■	■	■	■	■
PC aa C38:4																	■	■		■	■	
PC aa C40:2																	■		■		■	■
PC aa C40:3																	■	■	■	■	■	■
PC aa C40:4	■	■		■	■	■			■	■		■	■	■			■	■	■	■	■	■
PC aa C40:5										■						■	■	■	■	■	■	
PC aa C42:4																	■	■	■	■	■	■
PC aa C42:5																	■	■	■	■	■	■
PC ae C32:1																		■				
PC ae C32:2																	■	■	■	■	■	■
PC ae C34:0																	■	■		■	■	■
PC ae C34:1																	■	■	■	■	■	
PC ae C34:2																	■	■	■	■	■	
PC ae C34:3																		■	■		■	
PC ae C36:1																	■	■	■	■	■	■
PC ae C36:2																	■	■	■	■	■	■
PC ae C36:3																	■	■	■	■	■	
PC ae C36:4																	■	■		■	■	
PC ae C38:1																	■	■	■	■	■	■
PC ae C38:2									■	■	■	■	■	■	■	■	■	■	■	■	■	■
PC ae C38:4										■							■	■	■	■	■	■
PC ae C38:5		■								■							■	■	■	■	■	■
PC ae C40:2																		■				■
PC ae C40:3										■					■		■	■	■	■	■	■
PC ae C40:4		■		■					■	■	■	■	■	■			■	■	■	■	■	■
PC ae C40:5		■								■							■	■	■	■	■	■
PC ae C42:2																	■	■	■	■	■	■
PC ae C42:3															■		■	■	■	■	■	■
PC ae C42:4																	■	■	■	■	■	■
PC ae C42:5	■	■	■	■	■	■			■	■	■	■	■	■	■		■	■	■	■	■	■
lysoPC a C16:0							■		■			■	■	■	■		■	■		■	■	■
lysoPC a C16:1									■			■	■	■	■		■			■	■	
lysoPC a C18:0															■		■	■	■	■	■	■



(1) **M:** results from main analysis adjusted for age, sex, body mass index (BMI), an indicator of each study wave (KORA S4, KORA F4, or KORA FF4), day of the week, time trend, temperature, relative humidity, smoking status, alcohol consumption, physical activity index, educational attainment, fasting status, and dietary score. Three degrees of freedom were used for the temperature and relative humidity splines, and four degrees of freedom per year in each study wave were used for the time trend spline.
S1: restricted to all participants attended KORA S4, KORAF4 or KORAFF4 with metabolomics data.
S2: restricted to fasting participants throughout the entire study period.
S3: IPW analysis.
S4: short-term (lag0 and 2-week moving average, respectively) analysis results from crude model adjusted for an indicator of each study wave (KORA S4, KORA F4, or KORA FF4), day of the week, time trend, temperature, relative humidity. Three degrees of freedom were used for the temperature and relative humidity splines, and four degrees of freedom per year in each study wave were used for the time trend spline.
S5: co-effect between immediate (2-day moving average) and short-term (2-week moving average) exposure and corresponding long-term exposure (annual average) to air pollutant.
S6: two-pollutant models additionally adjusted for PM_{coarse}.
S7: two-pollutant models additionally adjusted for O_3.
(2) Blue means negative associations between metabolites and air pollutants, while red means positive association between metabolites and air pollutants.

Table S7 Results comparison between main analysis and sensitivity analysis (O_3 exposure)

	2-day moving average							2-week moving average							8-week moving average				
	M	S1	S2	S3	S4	S5	S6	M	S1	S2	S3	S4	S5	S6	M	S1	S2	S3	S6
Trp																			
PC aa C30:0																			
PC aa C32:0																			
PC ae C34:1																			
PC ae C38:2																			
SM (OH) C24:1																			
SM C16:0																			
SM C16:1																			
SM C18:0																			
SM C20:2																			
SM C24:1																			

(1) **M:** results from main analysis adjusted for age, sex, body mass index (BMI), an indicator of each study wave (KORA S4, KORA F4, or KORA FF4), day of the week, time trend, temperature, relative humidity, smoking status, alcohol consumption, physical activity index, educational attainment, fasting status, and dietary score. Three degrees of freedom were used for the temperature and relative humidity splines, and four degrees of freedom per year in each study wave were used for the time trend spline.
S1: restricted to all participants attended KORA S4, KORAF4 or KORAFF4 with metabolomics data.
S2: restricted to fasting participants throughout the entire study period.
S3: IPW analysis.
S4: short-term (lag0 and 2-week moving average, respectively) analysis results from crude model adjusted for an indicator of each study wave (KORA S4, KORA F4, or KORA FF4), day of the week, time trend, temperature, relative humidity. Three degrees of freedom were used for the temperature and relative humidity splines, and four degrees of freedom per year in each study wave were used for the time trend spline.
S5: co-effect between immediate (2-day moving average) and short-term (2-week moving average) exposure and corresponding long-term exposure (annual average) to air pollutant.
S6: two-pollutant models additionally adjusted for PM_{coarse}.
(2) Blue means negative associations between metabolites and air pollutants, while red means positive association between metabolites and air pollutants.

Table S8 Metabolic pathways related to air pollutants in different exposure windows.

Air pollutants	Exposure windows	Exposure window/Pathways	Total	Hits	*p*-value	FDR	Impact factor
$PM_{2.5}$	8-week moving average	Arachidonic acid metabolism	1	1	0.1	1	0
		Linoleic acid metabolism	1	1	0.1	1	0
		alpha-Linolenic acid metabolism	1	1	0.1	1	0
		D-Arginine and D-ornithine metabolism	2	1	0.1	1	0
		Glycerophospholipid metabolism	2	1	0.1	0.9	0.09
		Tryptophan metabolism	3	1	0.2	1	0.1
		Arginine biosynthesis	4	1	0.3	1	0.08
		Aminoacyl-tRNA biosynthesis	16	2	0.4	1	0
		Arginine and proline metabolism	8	1	0.5	1	0.06
PM_{coarse}	2-week moving average	Arachidonic acid metabolism	1	1	0.02	0.6	0
		Linoleic acid metabolism	1	1	0.02	0.6	0
		alpha-Linolenic acid metabolism	1	1	0.02	0.6	0
		Glycerophospholipid metabolism	2	1	0.04	0.9	0.09
	8-week moving average	Glycine, serine and threonine metabolism	5	4	0.008	0.3	0.5
		Sphingolipid metabolism	2	2	0.05	0.5	0
		Cysteine and methionine metabolism	2	2	0.05	0.5	0.1
		Glyoxylate and dicarboxylate metabolism	2	2	0.05	0.5	0.2
		Porphyrin and chlorophyll metabolism	1	1	0.2	1	0
		Aminoacyl-tRNA biosynthesis	16	5	0.3	1	0.2
		Glutathione metabolism	5	2	0.3	1	0.09
		D-Arginine and D-ornithine metabolism	2	1	0.4	1	0
		Primary bile acid biosynthesis	2	1	0.4	1	0.008
		Glycerophospholipid metabolism	2	1	0.4	1	0.02
		Phenylalanine, tyrosine and tryptophan biosynthesis	2	1	0.4	1	0.5
		Tryptophan metabolism	3	1	0.6	1	0.1
		Phenylalanine metabolism	3	1	0.6	1	0.4
		Valine, leucine and isoleucine biosynthesis	4	1	0.7	1	0
		Arginine biosynthesis	4	1	0.7	1	0.06
		Arginine and proline metabolism	8	1	0.9	1	0.1
NO_2	2-day moving average	Glycerophospholipid metabolism	2	2	0.001	0.3	0.1
		Tryptophan metabolism	1	1	0.1	1	0.1
		Aminoacyl-tRNA biosynthesis	1	1	0.1	1	0
		Arachidonic acid metabolism	1	1	0.04	0.9	0
		Linoleic acid metabolism	1	1	0.04	0.9	0
		alpha-Linolenic acid metabolism	1	1	0.04	0.9	0
	2-week moving average	Glycerophospholipid metabolism	2	2	0.006	0.2	0.1

		Ubiquinone and other terpenoid-quinone biosynthesis	1	1	0.09	0.6	0
		Arachidonic acid metabolism	1	1	0.09	0.6	0
		Linoleic acid metabolism	1	1	0.09	0.6	0
		alpha-Linolenic acid metabolism	1	1	0.09	0.6	0
		Phenylalanine, tyrosine and tryptophan biosynthesis	2	1	0.2	1	0.5
		Phenylalanine metabolism	3	1	0.2	1	0
		Tyrosine metabolism	3	1	0.2	1	0.1
		Tryptophan metabolism	3	1	0.2	1	0.1
		Aminoacyl-tRNA biosynthesis	16	2	0.4	1	0
	8-week moving average	Sphingolipid metabolism	2	2	0.04	0.7	0
		Glycerophospholipid metabolism	2	2	0.04	0.7	0.1
		Tryptophan metabolism	3	2	0.1	1	0.2
		Ubiquinone and other terpenoid-quinone biosynthesis	1	1	0.2	1	0
		Arachidonic acid metabolism	1	1	0.2	1	0
		Linoleic acid metabolism	1	1	0.2	1	0
		alpha-Linolenic acid metabolism	1	1	0.2	1	0
		Glycine, serine and threonine metabolism	5	2	0.3	1	0.2
		D-Arginine and D-ornithine metabolism	2	1	0.4	1	0
		Cysteine and methionine metabolism	2	1	0.4	1	0.02
		Glyoxylate and dicarboxylate metabolism	2	1	0.4	1	0.04
		Phenylalanine, tyrosine and tryptophan biosynthesis	2	1	0.4	1	0.5
		Aminoacyl-tRNA biosynthesis	16	4	0.5	1	0.2
		Phenylalanine metabolism	3	1	0.5	1	0
		Tyrosine metabolism	3	1	0.5	1	0.1
		Arginine biosynthesis	4	1	0.6	1	0.08
		Arginine and proline metabolism	8	1	0.9	1	0.06

Total: total number of metabolites in the pathway; Hits: the actually matched number from uploaded data; *p*-value: original *p*-value calculated from the enrichment analysis; FDR: *p*-value adjusted by False Discovery Rate; Impact factor: pathway impact value calculated from pathway topology analysis.
$PM_{2.5}$ = particulate matter with an aerodynamic diameter less than or equal to 2.5 μm; PM_{coarse} = particulate matter with an aerodynamic diameter of 2.5-10 μm; NO_2 = nitrogen dioxide.

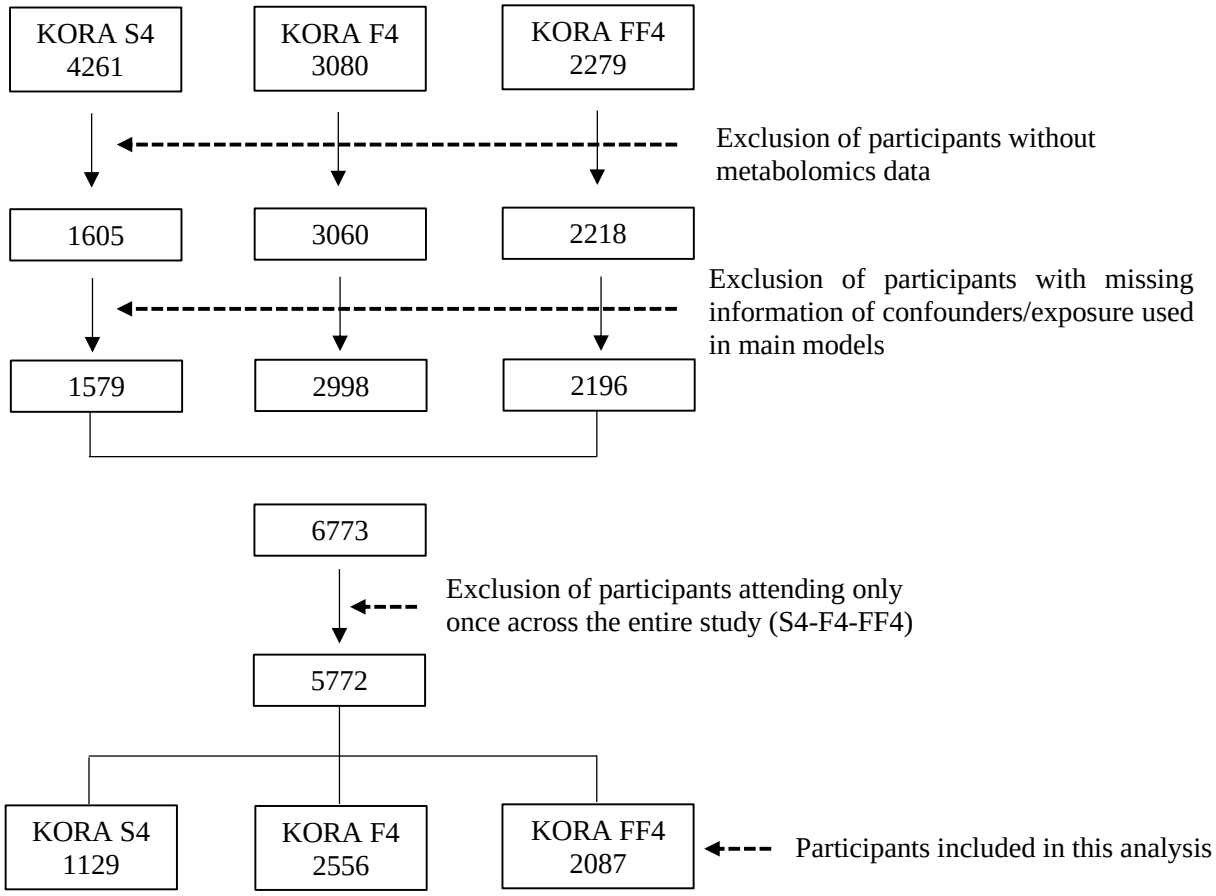


Figure S1. Flow chart of participant exclusion process in this study.

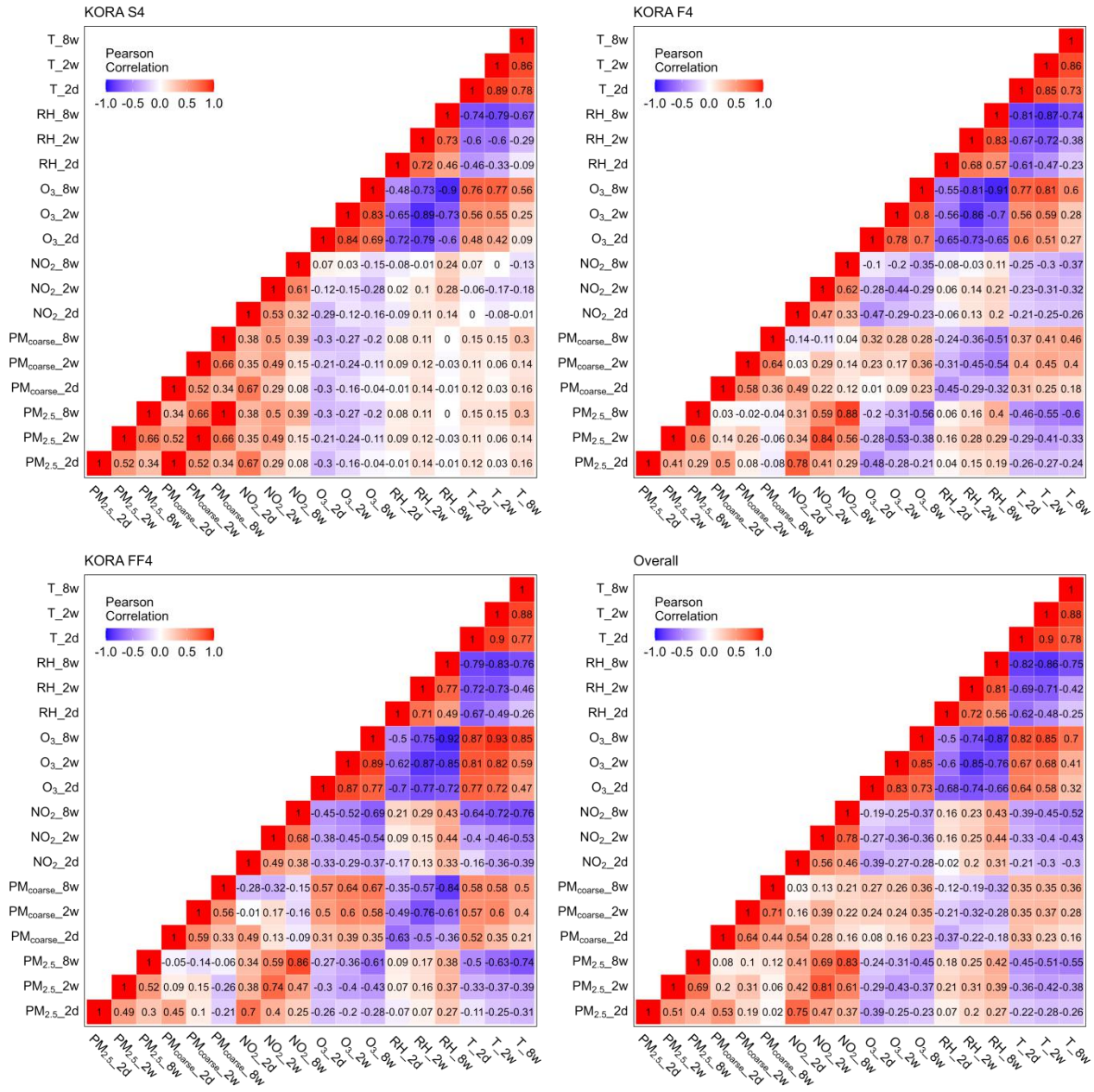


Figure S2. Spearman correlations air pollutants and meteorological parameters in each study survey (S4, F4, and FF4) and throughout the entire study (overall) per exposure window and cross-exposure windows.

$PM_{2.5}$ = particulate matter with an aerodynamic diameter less than or equal to 2.5 μm; **PM_{coarse}** = particulate matter with an aerodynamic diameter of 2.5-10 μm; **NO_2** = nitrogen dioxide; **O_3** = ozone; **RH** = relative humidity; **T** = air temperature; **2d:** 2-day moving average of exposure levels before the examination day; **2w:** 2-week moving average of exposure levels before the examination day; **8w:** 8-week moving average of exposure levels before the examination day.

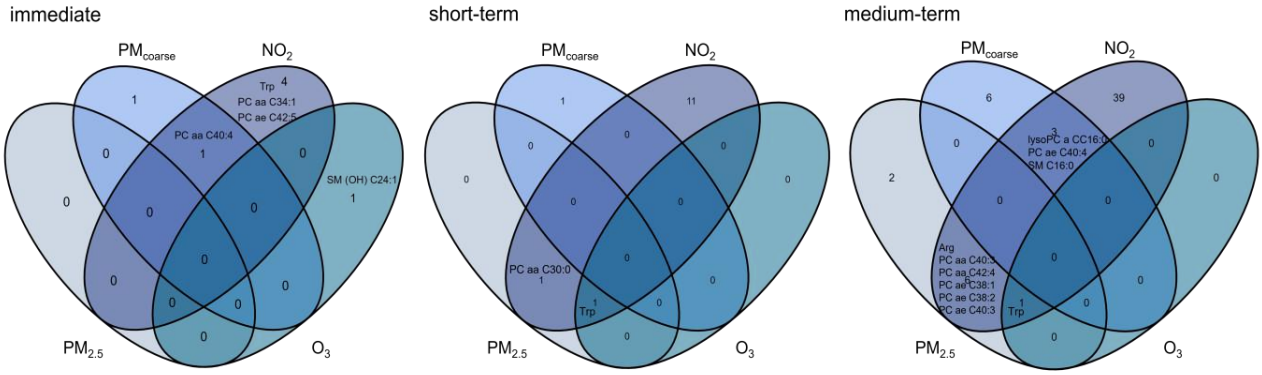


Figure S3. Venn diagrams of significant associations between immediate, short- and medium-term air pollutant exposures and metabolites.

$PM_{2.5}$ = particulate matter with an aerodynamic diameter less than or equal to 2.5 μm; PM_{coarse} = particulate matter with an aerodynamic diameter of 2.5-10 μm; NO_2 = nitrogen dioxide; O_3 = ozone; immediate = 2-day moving average of exposure levels before the examination day; short-term = 2-week moving average of exposure levels before the examination day; medium-term= 8-week moving average of exposure levels before the examination day. Only overlapping metabolites were displayed in the overlapped area instead of showing all the significant results. The positive number in short- and medium-term plots represents the non-overlapping metabolites for each air pollutant; 0 indicates either no overlapping metabolite or no additional significant result for the corresponding air pollutant.

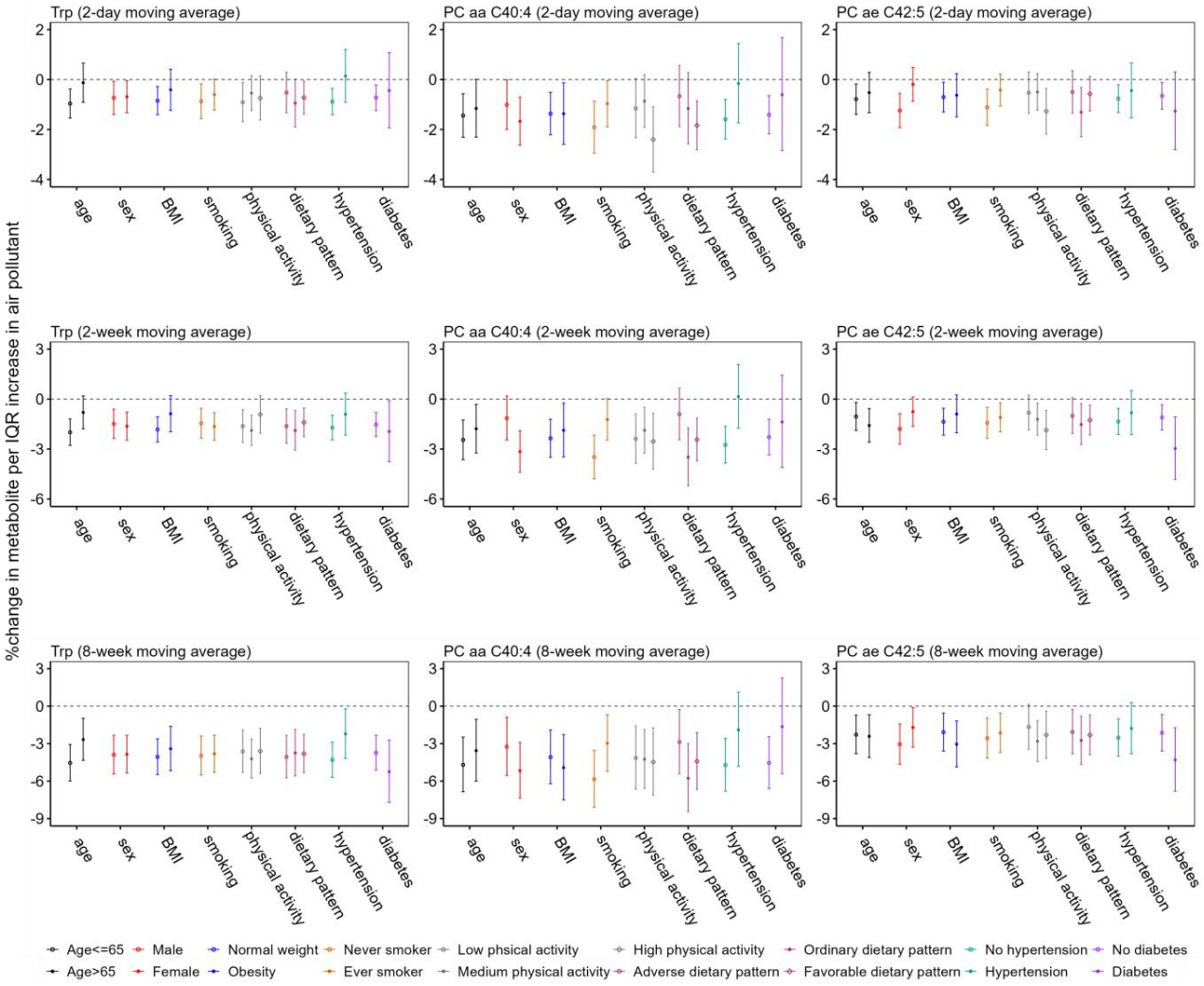


Figure S4. Percent changes (95% CIs) of metabolites (Trp, PC aa C40:4, and PC ae C42:5) per IQR increase in $PM_{2.5}$ exposure stratified by age, sex, BMI, smoking, physical activity level, dietary pattern, hypertension, and diabetes. Results were from our main models adjusted for age, sex, body-mass index (BMI), an indicator of each study wave (KORA S4, KORA F4, or KORA FF4), day of the week, time trend, temperature, relative humidity, smoking status, alcohol consumption, physical activity index, educational attainment, fasting status, and dietary score. Three degrees of freedom were used for the temperature and relative humidity splines, and four degrees of freedom per year in each study wave were used for the time trend spline. The continuous variable will be replaced by each corresponding effect modifier. 2-day, 2-week, and 8-week moving average were the exposure windows, and the IQR increase was 10.0 $\mu g/m^3$, 7.0 $\mu g/m^3$, and 6.9 $\mu g/m^3$ for 2-day, 2-week, and 8-week moving average of $PM_{2.5}$ throughout the entire study period, respectively.

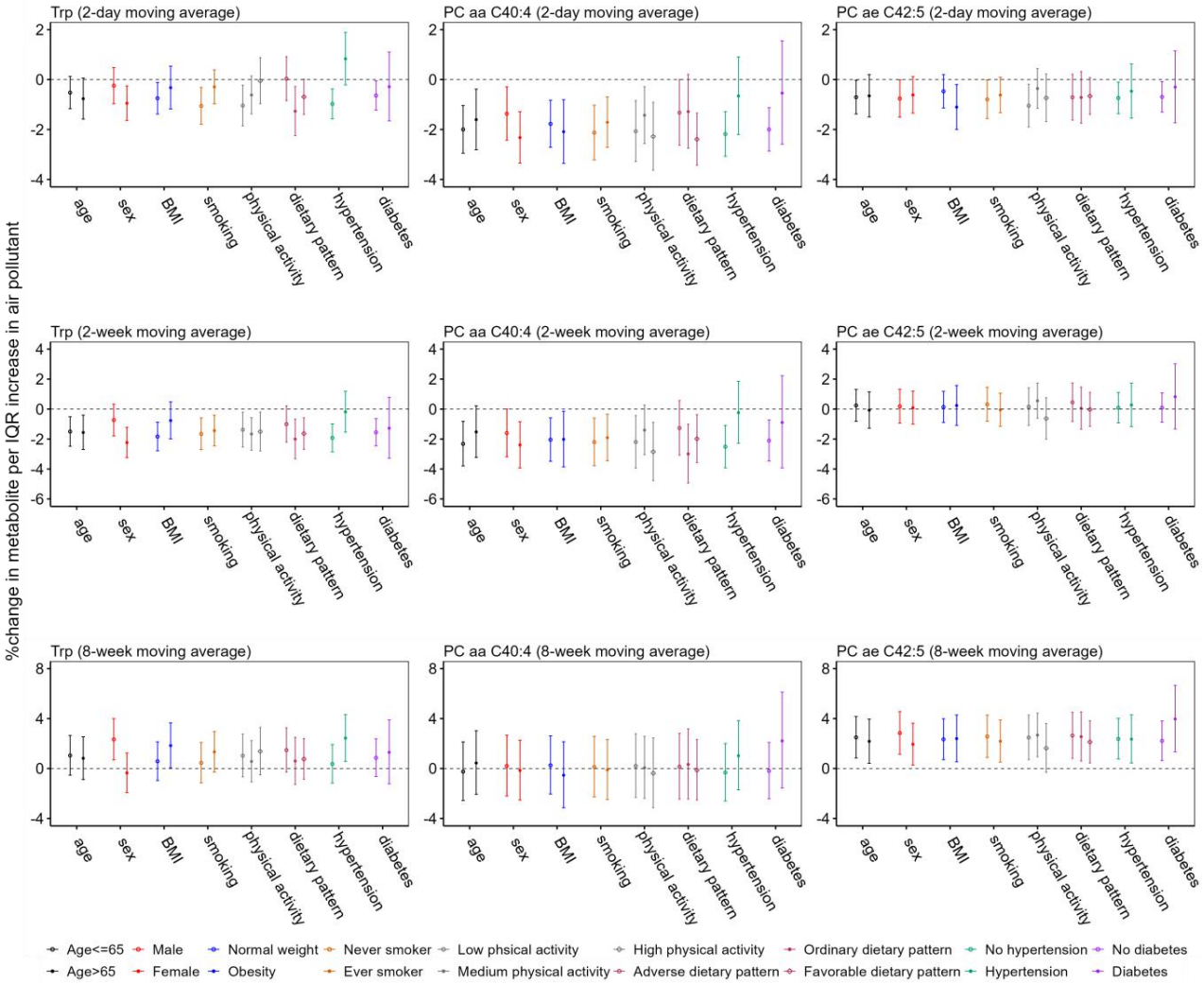


Figure S5. Percent changes (95% CIs) of metabolites (Trp, PC aa C40:4, and PC ae C42:5) per IQR increase in PM_{coarse} exposure stratified by age, sex, BMI, smoking, physical activity level, dietary pattern, hypertension, and diabetes. Results were from our main models adjusted for age, sex, body-mass index (BMI), an indicator of each study wave (KORA S4, KORA F4, or KORA FF4), day of the week, time trend, temperature, relative humidity, smoking status, alcohol consumption, physical activity index, educational attainment, fasting status, and dietary score. Three degrees of freedom were used for the temperature and relative humidity splines, and four degrees of freedom per year in each study wave were used for the time trend spline. The continuous variable will be replaced by each corresponding effect modifier. 2-day, 2-week, and 8-week moving average were the exposure windows, and the IQR increase was 3.9 $\mu g/m^3$, 3.3 $\mu g/m^3$, and 2.9 $\mu g/m^3$ for 2-day, 2-week, and 8-week moving average of PM_{coarse} throughout the entire study period, respectively.

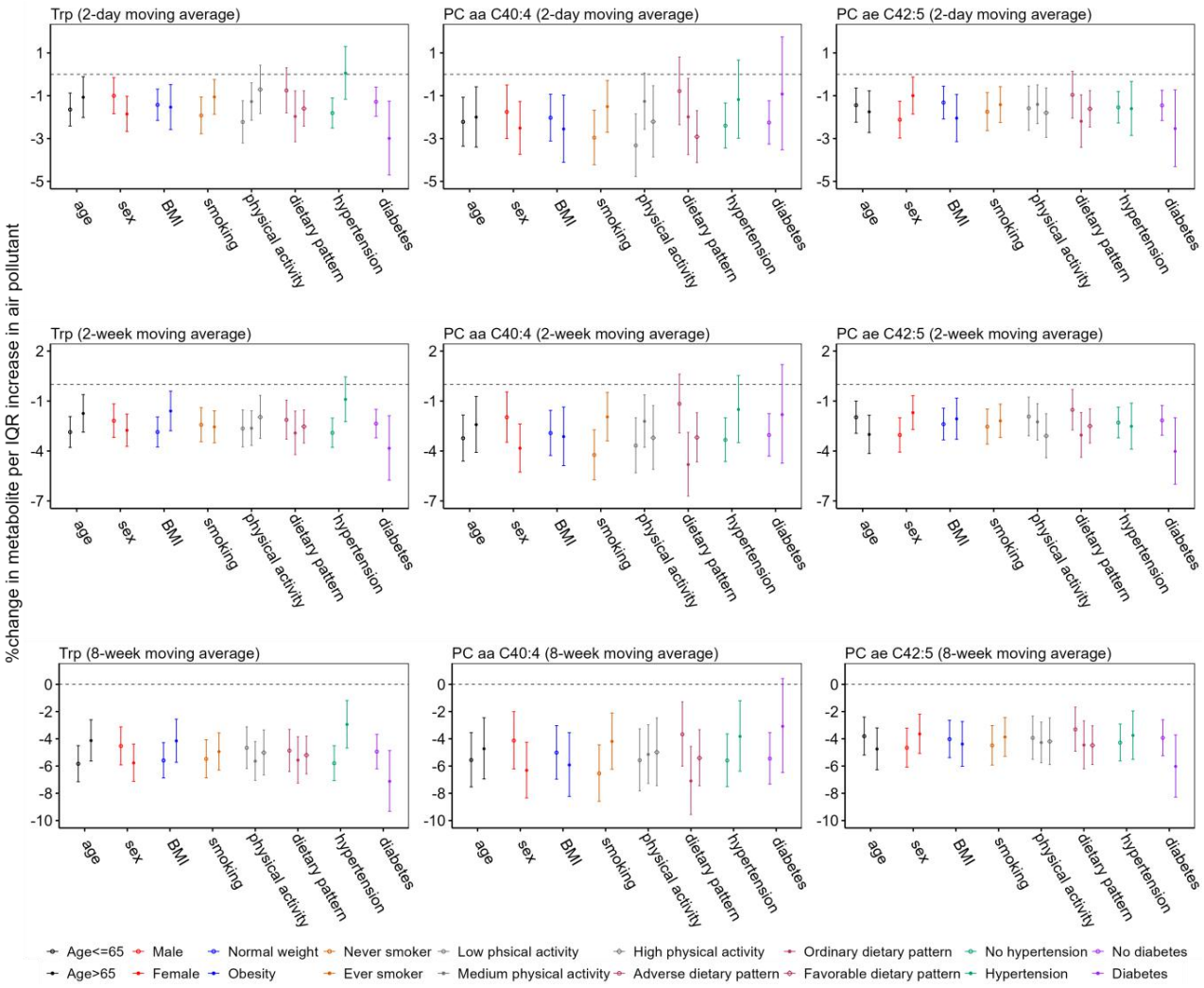


Figure S6. Percent changes (95% CIs) of metabolites (Trp, PC aa C40:4, and PC ae C42:5) per IQR increase in NO_2 exposure stratified by age, sex, BMI, smoking, physical activity level, dietary pattern, hypertension, and diabetes. Results were from our main models adjusted for age, sex, body-mass index (BMI), an indicator of each study wave (KORA S4, KORA F4, or KORA FF4), day of the week, time trend, temperature, relative humidity, smoking status, alcohol consumption, physical activity index, educational attainment, fasting status, and dietary score. Three degrees of freedom were used for the temperature and relative humidity splines, and four degrees of freedom per year in each study wave were used for the time trend spline. The continuous variable will be replaced by each corresponding effect modifier. 2-day, 2-week, and 8-week moving average were the exposure windows, and the IQR increase was 14.8 μg/m^3, 9.5 μg/m^3, and 8.9 μg/m^3 for 2-day, 2-week, and 8-week moving average of NO_2 throughout the entire study period, respectively.

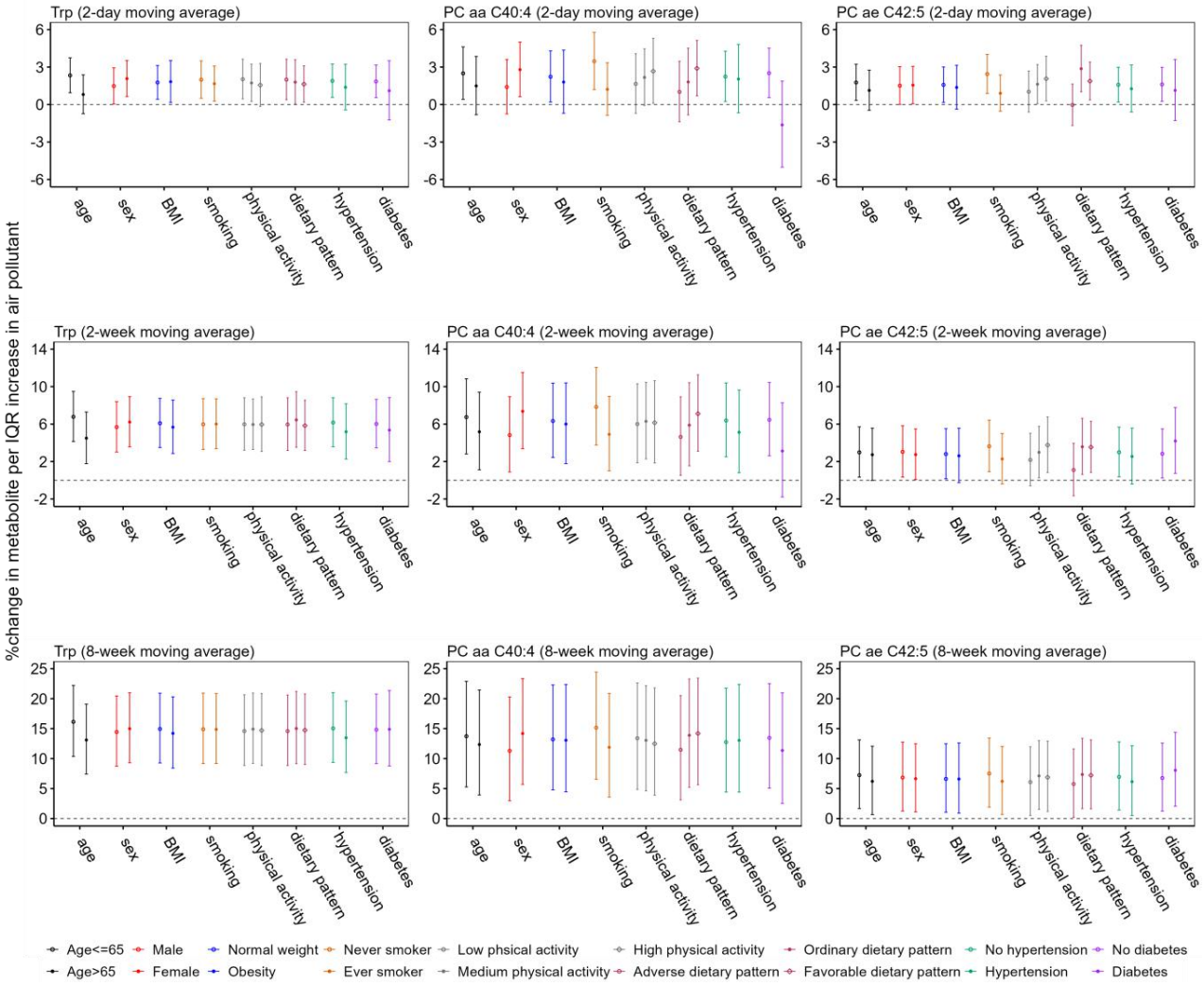


Figure S7. Percent changes (95% CIs) of metabolites (Trp, PC aa C40:4, and PC ae C42:5) per IQR increase in O_3 exposure stratified by age, sex, BMI, smoking, physical activity level, dietary pattern, hypertension, and diabetes. Results were from our main models adjusted for age, sex, body-mass index (BMI), an indicator of each study wave (KORA S4, KORA F4, or KORA FF4), day of the week, time trend, temperature, relative humidity, smoking status, alcohol consumption, physical activity index, educational attainment, fasting status, and dietary score. Three degrees of freedom were used for the temperature and relative humidity splines, and four degrees of freedom per year in each study wave were used for the time trend spline. The continuous variable will be replaced by each corresponding effect modifier. 2-day, 2-week, and 8-week moving average were the exposure windows, and the IQR increase was 35.0 μg/m^3, 30.4 μg/m^3, and 30.1 μg/m^3 for 2-day, 2-week, and 8-week moving average of O_3 throughout the entire study period, respectively.

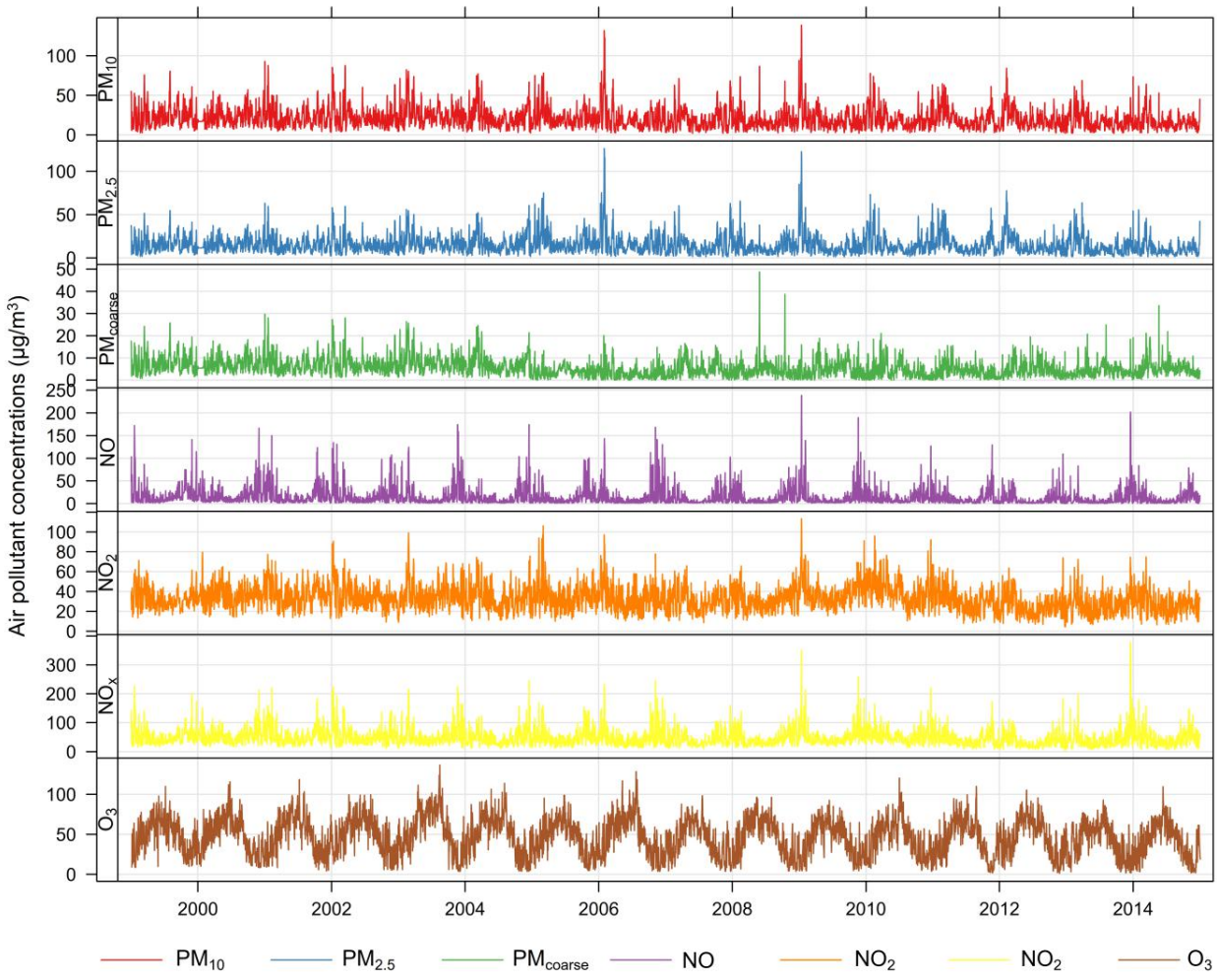


Figure S8. Time series of daily average concentrations of air pollutants from monitoring stations. The X axis shows the time period of monitoring exposures from the year 1999 to 2014. The Y axis indicates the concentrations of air pollutant exposures, and the unit is µg/m^3.

Appendix

Paper III

Title:	Long-term exposure to ambient air pollution is associated with epigenetic age acceleration
Authors:	Yueli Yao, Kathrin Wolf, Susanne Breitner, Siqi Zhang, Melanie Waldenberger, Juliane Winkelmann, Alexandra Schneider, Annette Peters
Status:	*Submitted*
Journal:	*Environmental International*
DOI:	Preprint: http://dx.doi.org/10.2139/ssrn.5255305
Impact factor:	10.3 (Journal Citation Reports®, year 2023)
Rank:	21/358 in Category Environmental Sciences Journals (Journal Citation Reports®, year 2023)

Long-term exposure to ambient air pollution is associated with epigenetic age acceleration

Yueli Yao [a,b], Kathrin Wolf [a], Susanne Breitner [a,b], Siqi Zhang [a,c], Melanie Waldenberger [a,d], Juliane Winkelmann [e,f], Alexandra Schneider [a]*, Annette Peters [a,b,g,h]*

[a] Institute of Epidemiology, Helmholtz Zentrum München GmbH, German Research Center for Environmental Health, Neuherberg, Germany

[b] Institute for Medical Information Processing, Biometry, and Epidemiology (IBE), Faculty of Medicine, LMU Munich, Pettenkofer School of Public Health, Munich, Germany

[c] Department of Environmental Health Sciences, Yale School of Public Health, New Haven, United States

[d] Research Unit Molecular Epidemiology, Helmholtz Zentrum München GmbH, German Research Center for Environmental Health, Neuherberg, Germany

[e] Institute of Neurogenomics, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany

[f] Institute of Human Genetics, Technical University of Munich, Munich, Germany

[g] German Center for Diabetes Research (DZD), Munich-Neuherberg, Germany

[h] German Centre for Cardiovascular Research (DZHK), Partner Site Munich, Munich, Germany

*Authors share last authorship

Correspondence to:

Yueli Yao, Institute of Epidemiology, Helmholtz Zentrum München GmbH, German Research Center for Environmental Health, Ingolstädter Landstr. 1, 85764 Neuherberg, Germany; Email: yueli.yao@helmholtz-munich.de.

Abstract

Background: Epigenetic aging biomarkers, predicted by selected Cytosine-phosphate-Guanine (CpG) sites, might be influenced by air pollution exposure. However, evidence from longitudinal studies is still limited.

Objectives: To determine the associations between long-term exposure to air pollution and epigenetic aging biomarkers and identify vulnerable subgroups.

Methods: Data was collected from the German population-based Cooperative Health Research in the Region of Augsburg (KORA) S4 survey (1999–2001) and two follow-up examinations (F4: 2006–08 and FF4: 2013–14). We measured DNA methylation (DNAm) in blood samples and calculated DNAm Age and DNAm-based telomere length (DNAmTL). We only included participants with at least two repeated measurements. Annual average concentrations of ultrafine particles (PNC), particulate matter (PM) less than 10 µm (PM_{10}), fine particles ($PM_{2.5}$), coarse particles (PM_{coarse}), soot ($PM_{2.5abs}$), nitrogen oxides (NO_2 and NO_x) and ozone (O_3) were estimated by land-use regression models. We applied linear mixed-effect regression models to assess the associations between air pollutants and epigenetic aging biomarkers, and further performed a limited epigenome-wide association study (EWAS) to examine whether air pollution influences individual CpGs.

Results: We included 4,105 observations from 1,651 KORA participants. Interquartile range (IQR) increases in all air pollutants except O_3 were positively associated with accelerated DNAmGrimAge and DNAmPhenoAge. Moreover, all air pollutants showed negative associations with DNAmTL. Specifically, in ever smokers, the air pollutants were positively associated with the age acceleration of DNAmHorvathAge and DNAmPhenoAge, and inversely associated with DNAmTL with the largest effect estimates observed for $PM_{2.5abs}$. We identified two exposure-related CpGs with PM_{coarse} at a Benjamini-Hochberg false discovery rate corrected p-value < 0.05 in ever smokers.

Conclusion: Our findings suggest a robust association between long-term exposure to traffic-related air pollution with epigenetic age acceleration, especially in ever smokers. These results imply that air pollution is augmenting the negative impact of smoking on biological ageing.

Keywords: Environmental exposures, epigenetic clocks, particulate matter, nitrogen oxides, smoking, sex

Introduction

Epidemiological evidence links air pollution exposure to elevated risks of aging-related diseases e.g., pulmonary, cardiovascular, and neurological diseases, and cancer.[1,2] Research identifies inflammation, oxidative stress, genetic, and epigenetic alterations—particularly DNA methylation (DNAm)—as potential hallmarks through which air pollution may influence disease development and progression.[1] DNAm-based epigenetic aging biomarkers, though derived from limited CpG sites, provide molecular insights into aging, environmental responses, and health risk prediction.[3-8] Depending on their training method, they could either accurately predict chronological age, e.g., Hannum and Horvath epigenetic clocks, or quantify age- or disease-related outcomes, e.g., PhenoAge clocks.[7] Therefore, they may serve as potential hallmarks to link air pollution and aging-related health outcomes.[1] However, evidence from longitudinal studies in cohort settings remains limited.

The Hannum clock (DNAmHannumAge) is a blood-based epigenetic clock influenced by age-related leukocyte composition and is more accurate in adults.[3] In contrast, Horvath clock (DNAmHorvathAge) is a multi-tissue epigenetic clock prediction throughout the lifespan.[4] The second-generation clocks enhance prediction of biological aging, morbidity, and mortality. For example, the Levine clock (DNAmPhenoAge) integrates chronological age with mortality-related blood markers,[5] while DNAmGrimAge combined DNAm surrogates of plasma proteins (e.g., C-reactive protein) and risk factors (e.g., smoking) to regress time-to-death.[6] DNAmSkinBloodAge improves age prediction in skin and blood samples.[8] Telomere length (TL), which shortens with cell division, is a traditional biomarker of aging but is limited by technical measurement variability.[9,10] By contrast, DNAm-based TL (DNAmTL), estimated by 140 CpG sites, providesa more robust predictor of mortality and age-related outcomes.[11]

Epidemiological studies have explored associations between long-term air pollution exposure and epigenetic aging biomarkers, though findings remain inconsistent. A cross-sectional study based on a survey of the population-based Cooperative Health Research in the Region of Augsburg (KORA F4) observed weak associations between air pollution and DNAmHorvathAge. Moreover, some examined epigenetic aging accelerations could be sex-specific.[12] Among a study of non-Hispanic white women in U.S., nitrogen dioxide (NO_2) and fine particle ($PM_{2.5}$) components were associated with acceleration of several epigenetic clocks and, though the directions of those associations varied.[13] In the Normative Aging Study (NAS), a longitudinal cohort consisting of only older males in the U.S., $PM_{2.5}$ and black carbon were associated with higher DNAmHorvathAge.[14] Several recent studies have included newer epigenetic aging biomarkers such as DNAmGrimAge and DNAmTL.[15,16] However, none of them investigated all of the above-mentioned epigenetic clocks and were limited by narrow age ranges or population representativeness (e.g., elderly adults (70–80 years) in Baranyi et al. and children (6–11 years) in Prado-Bert et al.).[15,17]

Given limitations of previous studie, more longitudinal research is needed to clarify the association between long-term air pollution and biological aging. To address this, we conducted a longitudinal analysis in the KORA cohort, examining associations between air pollution and three generations of DNAm-based aging biomarkers. We also assessed whether individual characteristics (e.g., age, sex, BMI, lifestyle factors, and pre-existing diseases) modified these associations. We hypothesized that air pollution accelerates biological aging, with susceptibility varying by individual characteristics.

Methods

Study design and participants

This longitudinal study was based on data of the population-based KORA cohort, conducted in the area of Augsburg, Germany. The fourth cross-sectional health survey of the KORA cohort (KORA S4) was conducted from October 1999 to April 2001, recruiting 4,261 participants aged 25 – 74 years with German citizenship in the city of Augsburg, Germany, and two adjacent counties. Two follow-up examinations were carried out, including 1) the first follow-up (KORA F4): 3,080 participants attended the examination between October 2006 and May 2008; 2) the second follow-up (KORA FF4): all examinations were conducted from June 2013 to September 2014 with an inclusion of 2,279 participants.

A computer-assisted personal interview, a self-administered questionnaire, and physical examinations were performed at each visit. Only participants who attended at least two visits were included in this analysis. Additionally, we excluded participants with missing data on covariates used in our main analysis (**Appendix, Figure S1**). Individual characteristics relevant in this analysis are written in the appendix (**Appendix, Text S1**).

A written informed consent was obtained from all participants. The KORA study was approved by the ethics committee of the Bavarian Chamber of Physicians (Munich, Germany).

DNA methylation data and epigenetic clocks estimates

For KORA S4 and F4 participants, DNA was extracted from whole blood, and methylation was measured by the Infinium HumanMethylation450K BeadChip. In contrast, the Infinium HumanMethylationEPIC BeadChip was used to examine the DNA methylation for KORA FF4 participants. A β-value, representing the methylation level of a given cytosine, was used for further estimates of epigenetic aging biomarkers. All epigenetic aging biomarkers were calculated by Horvath's online calculator (http://dnamage.genetics.ucla.edu/). In total, six DNAm-based aging biomarkers including DANmHorvathAge, DNAmHannumAge, DNAmPhenoAge, DNAmGrimAge, DNAmSkinBloodAge, and DNAmTL were predicted by this online calculator and included in this analysis.[3-6,8,11] There were different numbers of overlapping CpG sites between those epigenetic aging biomarkers (**Table 1**). Except for DNAmTL, differences between each epigenetic clock and chronological age were then used as age accelerations. The processing of DNAm data and prediction of epigenetic clocks are written in appendix (**Appendix, Text S2**).

Exposure assessment

Eight air pollutants were included in this analysis, including five particulate air pollutants (ultrafine particles with ≤ 100 nm in aerodynamic diameter, represented by particulate number concentration (PNC)), particulate matter (PM) with an aerodynamic diameter less than 10 μm (PM_{10}), coarse particles (PM_{coarse}), $PM_{2.5}$, and soot ($PM_{2.5abs}$)), and three gaseous air pollutants (nitrogen oxides (NO_2 and NO_x), and ozone (O_3)). Land-use regression (LUR) models were used to estimate the annual air pollution concentrations. The adjusted model-explained variance (R^2) of the LUR models ranged from 68% (PM_{coarse}) to 94% (NO_2), and the adjusted leave-one-out cross-validation R^2 was between 55% (PM_{coarse}) and 89% (NO_2), which indicated a good model fit. Participants' home addresses were then applied to the fitted models to determine individual residential exposure levels. If participants moved throughout the study period, the updated residential addresses were considered for exposure assignment. Otherwise, the same exposure

concentrations were assigned to the different visits. To account for potential confounding by road traffic noise, we assigned annual average day/night sound levels to participants' residential addresses. Details for exposure assessments are provided in the appendix (**Appendix, Text S3**).

Statistical Analyses

Descriptive analyses were conducted for participant characteristics and air pollutants. Differences across the three examination waves were tested using Kruskal-Wallis tests for continuous variables and Chi-squared tests for categorical variables. Pearson's correlation coefficients assessed relationships between chronological age, epigenetic aging biomarkers, and air pollutants.

Linear mixed-effects models with participant-specific random intercepts and batch and chip numbers as random effects were used to assess associations between repeated epigenetic aging biomarkers and air pollutants. Further covariates were selected based on previous studies.[12] ***Basic models*** adjusted for age, sex, estimated houseman cell types (**Appendix, Text S2**), and an indicator of each visit; ***behavioural models*** further added smoking, alcohol consumption (g/day), physical activity, and education; ***clinical models*** considered covariates of the basic models plus clinical variables, including body mass index (BMI), hypertension, diabetes, high-density lipoprotein (HDL), and low-density lipoprotein (LDL); ***full models*** included all covariates and were considered as our main results.

Since health effects of air pollution exposure might be closely related to those of smoking sharing some of the pathways, we were interested in exploring potential differences in the effect estimates of ever and never smokers, and performed all analyses on ever and never smokers separately. Effect modification was examined by including interaction terms between air pollutants and potential modifiers assessed at each visit, including age, sex, BMI, smoking, alcohol consumption, physical activity, education, hypertension, and type 2 diabetes. More details were presented in Appendix (**Appendix, Text S1**).

We conducted several sensitivity analyses to ensure the robustness of our findings: 1) To address potential bias from missing CpGs not covered by the Infinium HumanMethylationEPIC BeadChip (e.g., in the Hannum clock), we recalculated epigenetic aging biomarkers using only overlapping CpGs across arrays by imputing the missing CpGs; 2) We excluded estimated houseman cell types from the models since some epigenetic aging biomarkers were not affected by the white blood cells; 3) We increased the number of participants by including all participants with at least one valid visit with complete data; 4) Analyses were restricted to individuals with stable residential addresses throughout follow-up; 5) We separately included the daytime and nighttime noise averages in the full models (**Appendix, Text S3**). 6) Two-pollutant models were conducted when pollutant correlations were smaller than 0.7; 7) We conducted a limited epigenome-wide association study (EWAS) restricted to CpGs contributing to the predictions of epigenetic aging biomarkers to examine whether air pollution influences individual CpG. Ingenuity Pathway Analysis (IPA, QIAGEN Inc.) was used to identify enriched canonical pathway, and the pathway was determined if *p*-values < 0.05.

Effect estimates were presented as absolute change (together with 95% confidence intervals [95% CI]) of per interquartile range (IQR) increase in air pollutant concentrations, except for DNAmTL, which was presented as percent change from the overall mean. All analyses were done with R (version 4.1.2), with a significance threshold of *p*-values < 0.05. EWAS results were FDR-corrected using the Benjamini-Hochberg method (FDR < 0.05).

Results

Characteristics of study participants

Participants' characteristics are summarized in **Table 2**. The main analysis included 4,105 observations from 1,646 KORA participants (38.6%) who attended at least two visits with complete covariate data. Of these, 833 (50.6%) attended two visits and 813 (49.4%) attended all three.

Average age increased across KORA S4, F4, and FF4, while sex distribution (p-value = 0.78) and BMI (p-value = 0.07) remained consistent. Significant differences were observed across visits in alcohol consumption, HDL, LDL, the percentages of participants with hypertension or diabetes, and the counts of Houseman-estimated white blood cell types (p-value < 0.05). Mean values of all epigenetic aging biomarkers also varied significantly throughout the three visits (p-value < 0.01). In general, epigenetic aging biomarkers and chronological age showed moderate to high positive correlations (r = 0.48–0.83), except for DNAmTL, which showed consistently negative correlations (**Appendix**, **Figure S2**). Weak correlations were observed among Houseman-estimated white cell types (**Appendix, Figures S3**).

Characteristics of air pollutants

Annual average concentrations of $PM_{2.5}$, PM_{10}, and NO_2 at participant's residences were all below the current EU air quality standards, but higher than the WHO air quality guideline values (**Table 3**). We observed moderate to strong positive correlation among most air pollutants, except for O_3, showing negative or weak correlations with the other air pollutants. The IQR were 1.4 $\mu g/m^3$ for $PM_{2.5}$, 2.1 $\mu g/m^3$ for PM_{10}, 1.4 $\mu g/m^3$ for PM_{coarse}, $2.0\times10^3/cm^3$ for PNC, $0.3\times10^{-5}/m$ for $PM_{2.5abs}$, 7.2 $\mu g/m^3$ for NO_2, 8.8 $\mu g/m^3$ for NO_x, and 3.4 $\mu g/m^3$ for O_3. No correlation was found between air pollutants and chronological age (**Appendix**, **Figure S4**).

Associations between epigenetic aging biomarkers and long-term air pollution

DNAmTL showed robust negative associations with $PM_{2.5}$, PM_{10}, PNC, PM_{coarse}, $PM_{2.5abs}$, and NO_2 in our four regression models, and with NO_x in the basic and clinical models. The acceleration of DNAmGrimAge was positively associated with the same air pollutants, but only in the basic and clinical models (p-value < 0.05), but not in the behavioural and full models (**Figure 1**). Stepwise inclusion of behavioural covariates into the clinical model revealed that smoking status shifted the results in the behavioural and full models (**Appendix**, **Figure S5**).

Stratified analysis by smoking status (smoking-specific associations)

Among ever smokers, all air pollutants except for O_3 were associated with accelerated DNAmHorvathAge and DNAmPhenoAge (**Figure 2**). DNAmTL was negatively associated with $PM_{2.5}$, $PM_{2.5abs}$, and NO_2. DNAmGrimAge and DNAmSkinBloodAge showed stronger, though non-significant, associations in ever smokers. DNAmHannumAge effects were similar in both smoking groups. Overall, ever smokers might be more susceptible to air pollution–related epigenetic aging.

Effect modification

Figure 3 shows that associations between DNAmTL and PM_{10}, PNC, PM_{coarse}, $PM_{2.5abs}$, NO_2 and NO_x were modified by hypertension (uncorrected $p < 0.001$). No significant effect modification was observed for other modifiers, such as age, sex, obesity, alcohol consumption, educational attainment, physical activity, or diabetes. For the other five epigenetic aging biomarkers, no consistent patterns of effect modification were found (**Appendix**, **Figures S6-S10**).

211 ***Sensitivity analysis***

212 The associations between air pollution and six epigenetics clocks remained generally robust across
213 sensitivity analyses (**Appendix, Figure S11).** In detail, results were consistent after excluding missing
214 CpGs in KORA FF4, omitting adjustment for houseman-estimated white cell types, limiting to single or
215 repeat visits, or restricting to non-movers. Additional adjustment for traffic noise showed robust results or
216 slightly strengthened the effect estimates of air pollution exposure. Two-pollutant models had similar
217 estimates to single-pollutant models (**Appendix, Figure S12-S13**). The results of sensitivity analysis and
218 two-pollutant models from the stratified analysis by smoking showed comparable results to our main
219 analysis (data not shown).

220 In total, 1,253CpGs were used to estimate the five epigenetic aging biomarkers (DNAmHorvathAge,
221 DNAmHannumAge, DNAmPhenoAge, DNAmSkinBloodAge, DNAmTL). In the limited EWAS, we did
222 not observe any significant results between air pollutant exposures and these CpGs in the main analysis or
223 among never smokers (FDR-corrected *p*-value < 0.05). However, in ever smokers, PM_{coarse} was
224 significantly associated with cg12745325 (annotated to *SLC39A5* on chromosome 12) included in the
225 estimation of DNAmTL and cg24081819 (within *EPHX2* on chromosome 8) included in the prediction of
226 DNAmHorvathAge (**Table 4**). The top canonical pathway, top biofunction, and top disease identified via
227 IPA were triacylglycerol biosynthesis, cardiovascular disease, and lipid metabolism, respectively
228 (**Appendix, Table S7**).

Discussion

In this study, long-term air pollution exposure was associated with accelerated epigenetic aging. DNAmTL showed negative associations with all particulate air pollutants and NO_2, while DNAmGrimAge and DNAmPhenoAge acceleration correlated positively with these air pollutants and NO_x in clinical models. In addition, the DNAmHorvathAge and DNAmPhenoAge were accelerated by all air pollutants except for O_3 in ever smokers. The strongest associations were observed for $PM_{2.5abs}$.

Epigenetic aging biomarkers and air pollution

The acceleration of epigenetic clocks, indicative of biological aging, may reflect disrupted homeostasis and heightened susceptibility to air pollution. Ward-Caviness et al. observed that accelerated epigenetic aging (DNAmHorvath) enhanced associations between traffic-related $PM_{2.5}$ (gasoline and diesel sources) and peripheral arterial disease.[18] Similarly, their cross-sectional analysis in the KORA F4 cohort found higher annual $PM_{2.5}$ exposure with DNAmHorvathAge acceleration.[12] The NAS cohort confirmed associations between $PM_{2.5}$ exposure and DNAmHorvathAge acceleration, identifying related CpG sites associated with lung pathology, but found no association with DNAmHannumAge or DNAmPhenoAge.[14,19,20] In the U.S. Sister Study, NO_2 exposure inversely associated with DNAmHannumAge acceleration, but $PM_{2.5}$ and PM_{10} showed no significant associations study.[13] Another investigation linked lifetime air pollution exposure to DNAmHorvathAge acceleration, though the association lost significance after adjusting for multiple testing.[15] A study in children reported that indoor particulate matter and parental smoking correlated with DNAmSkinBloodAge acceleration.[17] However, our analysis observed no significant associations between long-term air pollution exposure and DNAmHorvathAge, DNAmHannumAge, DNAmPhenoAge, or DNAmSkinBloodAge.

In our study, DNAmGrimAge acceleration was positively associated with most air pollutants (except O_3) in basic and clinical models, but not after adjusting for behavioral factors (smoking status, alcohol consumption, physical activity, educational attainment). Few previous studies examined air pollution effects using DNAmGrimAge, and findings have been inconsistent. For instance, Koenigsberg et al. reported positive associations between PM_{10}, NO_2, and DNAmGrimAge, while Baranyi et al. found no association between lifetime air pollution exposure and DNAmGrimAge.[15,16] Our analyses indicated that behavioral factors, particularly smoking status, significantly attenuated the associations between air pollution and DNAmGrimAge acceleration (Appendix, Figure S5). Therefore, we assumed that effect estimates resulting from air pollution might differ between ever-smoking and never-smoking participants.

Stratified analyses among ever smokers revealed positive associations of DNAmHorvathAge and DNAmPhenoAge acceleration with $PM_{2.5}$, PM_{coarse}, and NO_2, consistent with prior research.[12,14,18] Although DNAmHorvathAge was initially developed to predict chronological age, accelerated DNAmHorvathAge in human liver tissue has been associated with increased BMI and higher risk of mortality.[21,22] Compared to other DNAm aging biomarkers, DNAmPhenoAge has a stronger positive correlation with smoking status and better accuracy in predicting lifespan and morbidity outcomes such as Alzheimer's disease and cancers Furthermore, the CpGs in DNAmPhenoAge are involved in inflammatory response and immune cells.[5] Additionally, the top enriched pathway we identified by the exposure-related CpGs in ever smokers was mainly involved in lipid metabolism and cardiovascular disease, further indicating potential health risks from long-term air pollution exposure.

TL primarily varies between rather than within individuals, and shorter leukocyte TL is associated with cardiovascular disease and reduced lifespan.[11] However, its measurement is sensitive to technical

confounding factors such as different DNA extraction and measuring methods.[10] DNAmTL is derived from leukocyte telomere length but showed better predictive ability than leukocyte TL for coronary heart disease, heart failure, and mortality outcomes.[11] In our study, DNAmTL was negatively associated with long-term exposure to particulate air pollutants ($PM_{2.5}$, PM_{10}, PM_{coarse}, and $PM_{2.5abs}$) and NO_2 from our fully covariates-adjusted models in overall participants and ever smokers. Consistent with our findings, Ward-Caviness et al. previously observed an inverse relationship between black carbon (equal to $PM_{2.5abs}$) and TL-based age acceleration in males using the KORA F4 methylation data, which was replicated in the NAS cohort as well.[12] Additionally, , a life-course study similarly reported shorter DNAmTL was associated with increased $PM_{2.5}$ and NO_2 exposure.[15] The results from the above studies support our findings in this longitudinal study, suggesting that DNAmTL could be a sensitive biomarker for detecting adverse health outcomes related to air pollution exposure.

Variability of epigenetic aging biomarkers in response to air pollution

Epigenetic aging biomarkers based on DNAm generally outperform transcriptomic, proteomic, metabolomic predictors, and telomere length in predicting chronological age and health outcomes.[23] However, responses to air pollution exposure vary among these epigenetic aging biomarkers due to differences in their training datasets, calibration methods, and statistical approaches, which leads to inconsistent findings across studies. For example, previous research reported differing associations between epigenetic age acceleration and air pollutants: DNAmHorvathAge and intrinsic/extrinsic epigenetic age accelerations were positively associated with NO_x and black carbon exposure in females, but inversely associated with PM_{10} in males, but not with $PM_{2.5}$ in both females and males.[12] Only DNAmHannumAge acceleration was inversely associated with NO_2 but not the other examined age accelerations, and no significant result from $PM_{2.5}$ and PM_{10}. The examined age accelerations even responded differently to $PM_{2.5}$ component clusters.[13] Additionally, findings from the NAS cohort indicated that DNAmHorvathAge acceleration correlated with $PM_{2.5}$ exposure, whereas DNAmPhenoAge acceleration correlated specifically with certain $PM_{2.5}$ components (e.g., calcium, lead), highlighting the complexity and variability of epigenetic aging biomarkers in air pollution research.[14,19,20]

Susceptibility factors of air pollution effects

Behavioral, lifestyle, and health-related factors may influence DNAm-based aging biomarkers. Smoking, in particular, has been consistently linked to accelerated biological aging: Lu et al. found that DNAmTL shortened by approximately 0.02 kilobases per smoking pack-year, while Cardenas et al. reported that former smokers exhibited accelerated DNAmHorvathAge and shorter DNAmTL compared to non-smokers.[11,24] Our results similarly showed sensitivity to smoking adjustment, revealing stronger negative associations between air pollution exposure and DNAmTL among ever-smokers. In addition, DNAmHorvathAge and DNAmPhenoAge showed a stronger association with all air pollutants in ever smokers compared to participants who never smoked. Epigenetic modifications, particularly DNAm, are associated with aging and age-related chronic diseases, including cancer, cardiovascular disease, and diabetes. Individuals with chronic diseases might have a faster aging speed than healthy individuals.[25] For example, the NAS cohort found that participants with coronary heart disease, hypertension, or lifetime cancer diagnoses had higher mean epigenetic ages when assessing long-term air pollution exposure ($PM_{2.5}$ and black carbon).[14] In contrast, our study observed inverse associations: DNAmHannumAge acceleration and shorter DNAmTL were more pronounced in non-hypertensive compared to hypertensive participants, specifically among never smokers; however, these differences were attenuated among ever smokers. A potential explanation might be unknown confounding factors such as anti-hypertensive medication. One

prior study reported anti-hypertensive medications were associated with greater DNAmHorvathAge acceleration,[26] while opposite results were found in another study,[27] highlighting the complexity of these interactions. Regarding differences between women and men, some epigenetic aging biomarkers might be sex-specific, and females might experience increased aging speed after menopause.[12,28]

Enriched pathway with exposure-related CpG

Some studies have examined associations between air pollution exposure and DNAm at single CpG site, identifying significant exposure-related CpGs. For example, Plusquin et al. found that long-term NO_2 and NO_x exposure altered DNAm levels at multiple CpGs linked to immune system pathways when integrated with transcriptomic data.[29] Similarly, White et al. conducted a targeted EWAS focused on CpGs from DNAmHorvathAge, DNAmHannumAge, and DNAmPhenoAge, reporting that cg22920873, annotated by the *C7orf55* gene on chromosome 7, was associated with annual PM_{10} exposure and identified pathways involving developmental processes and cell communication.[13] Eze et al., using a candidate pathway analysis, reported associations between $PM_{2.5}$ exposure and enriched pathways related to inflammation (C-reactive protein), BMI, and renal function.[30] In our study, CpG-annotated genes associated with air pollution exposure were mainly involved in lipid metabolism and cardiovascular disease pathways. However, neither identified CpG sites nor enriched biological pathways out of those studies and our study were consistent, more studies are needed to investigate the heterogenicity across diverse populations and exposures.

Strengths and limitations

To our knowledge, this is the first longitudinal study to explore the associations between long-term ambient air pollution exposure and multiple generations of DNAm-based epigenetic clocks, DNAm-based TL, and multiple air pollutants. Compared with previous studies, our study has the largest sample size and a broader range of air pollutants analyzed, enhancing its robustness and generalizability.[12-18,20] Moreover, the well-characterized KORA cohort has standardized and comprehensive methods data collection, improving the reliability of our findings. Our longitudinal study design, incorporating repeated biomarker measurements, strengthened statistical power, reduced potential residual confounding from unmeasured factors, and provided analytical improvement. Additionally, this study has the strength to assess susceptibility from both external and intrinsic factors, particularly behavioral and lifestyle factors known to affect epigenetic aging.

Our study also has some limitations. The air pollutants concentrations were estimated using spatial models from data collected in 2014–2015. Although spatial contrasts in pollutant levels have been shown previously to remain stable over time, this estimation approach might introduce some historical exposure misclassification. To minimize such misclassification, we conducted sensitivity analyses restricted to non-movers (participants who did not relocate during the study), and the robust findings among this subgroup supported our exposure assessment methodology.

Conclusions

In conclusion, our study suggested that long-term exposure to ambient air pollution is associated with epigenetic aging biomarkers. This was particularly the case in current and former smokers, who might be a susceptible population group for air pollution exposure. The study indicates that air pollution is potentially contributing to immunosenescence and thereby strengthens the evidence that chronic exposure to low-level air pollution impacts multiple non-communicable diseases such as cardiometabolic and pulmonary diseases

Declaration of Interest Statement

We have no conflict of interest to declare.

Acknowledgments

We thank all participants for their long-term commitment to the KORA study, the staff for data collection and research data management and the members of the KORA Study Group (https://www.helmholtz-munich.de/en/epi/cohort/kora) who are responsible for the design and conduct of the study. We thank Nadine Lindemann, Sonja Kunze and Eva Reischl for profiling of KORA methylation data.

Funding

This work was supported by a scholarship under the State Scholarship Fund by the China Scholarship Council (File No. 201906180003). The KORA study was initiated and financed by the Helmholtz Zentrum München – German Research Center for Environmental Health, which is funded by the German Federal Ministry of Education and Research (BMBF) and by the State of Bavaria. Data collection in the KORA study is done in cooperation with the University Hospital of Augsburg. This work was also supported by the EXPANSE project which is funded by the European Union's Horizon 2020 research and innovation program under grant agreement No. 874627.

Author Statement

All authors contributed to the study design and conceptualization. Yueli Yao verified and analyzed the underlying data, with support from Kathrin Wolf, Susanne Breitner, Siqi Zhang, and Alexandra Schneider in developing the analysis plan. Yueli Yao drafted the manuscript, and all authors contributed to the interpretation of the findings. All co-authors—including Kathrin Wolf, Susanne Breitner, Alexandra Schneider, Siqi Zhang, Melanie Waldenberger, Juliane Winkelmann, and Annette Peters—critically revised the manuscript and approved the final version. All authors had full access to all study data and take responsibility for the integrity and accuracy of the data analysis. Alexandra Schneider and Annette Peters led the project and had final responsibility for the decision to submit the manuscript for publication.

References

1. Peters, A., T.S. Nawrot, and A.A. Baccarelli, Hallmarks of environmental insults. *Cell* 2021. **184(6)**: 1455-1468.
2. Kasdagli, M.I., K. Katsouyanni, K. de Hoogh, P. Lagiou, and E. Samoli, Investigating the association between long-term exposure to air pollution and greenness with mortality from neurological, cardio-metabolic and chronic obstructive pulmonary diseases in Greece. *Environ Pollut* 2022. **292(Pt B)**: 118372.
3. Horvath, S., DNA methylation age of human tissues and cell types. *Genome Biol.* 2013. **14**: R115.
4. Hannum, G., J. Guinney, L. Zhao, L. Zhang, G. Hughes, S. Sadda, et al., Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol Cell* 2013. **49(2)**: 359-367.
5. Levine, Morgan E., Ake T. Lu, A. Quach, Brian H.C. Chen, Themistocles L. Assimes, S. Bandinelli, et al., An epigenetic biomarker of aging for lifespan and healthspan. *Aging (Albany NY)* 2018. **10**: 573-591.
6. Lu, A., A. Quach, J. Wilson, A. Reiner, A. Aviv, K. Raj, et al., DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging (Albany NY)* 2019. **11**: 303-327.
7. Horvath, S. and K. Raj, DNA methylation-based biomarkers and the epigenetic clock theory of ageing. *Nat Rev Genet* 2018. **19(6)**: 371-384.
8. Horvath, S., J. Oshima, G.M. Martin, A.T. Lu, A. Quach, H. Cohen, et al., Epigenetic clock for skin and blood cells applied to Hutchinson Gilford Progeria Syndrome and ex vivo studies. *Aging (Albany NY)* 2018. **10(7)**: 1758-1775.
9. Nordfjäll, K., U. Svenson, K.-F. Norrback, R. Adolfsson, and G. Roos, Large-scale parent–child comparison confirms a strong paternal influence on telomere length. *European Journal of Human Genetics* 2010. **18(3)**: 385-389.
10. Denham, J., F.Z. Marques, and F.J. Charchar, Leukocyte telomere length variation due to DNA extraction method. *BMC Research Notes* 2014. **7(1)**: 877.
11. Lu, A.T., A. Seeboth, P.C. Tsai, D. Sun, A. Quach, A.P. Reiner, et al., DNA methylation-based estimator of telomere length. *Aging (Albany NY)* 2019. **11(16)**: 5895-5923.
12. Ward-Caviness, C.K., J.C. Nwanaji-Enwerem, K. Wolf, S. Wahl, E. Colicino, L. Trevisi, et al., Long-term exposure to air pollution is associated with biological aging. *Oncotarget* 2016. **7(46)**: 74510-74525.
13. White, A.J., J.K. Kresovich, J.P. Keller, Z. Xu, J.D. Kaufman, C.R. Weinberg, et al., Air pollution, particulate matter composition and methylation-based biologic age. *Environ Int* 2019. **132**: 105071.
14. Nwanaji-Enwerem, J.C., E. Colicino, L. Trevisi, I. Kloog, A.C. Just, J. Shen, et al., Long-term ambient particle exposures and blood DNA methylation age: findings from the VA normative aging study. *Environ Epigenet* 2016. **2(2)**.
15. Baranyi, G., I.J. Deary, D.L. McCartney, S.E. Harris, N. Shortt, S. Reis, et al., Life-course exposure to air pollution and biological ageing in the Lothian Birth Cohort 1936. 2022.
16. Koenigsberg, S.H., C.-J. Chang, J. Ish, Z. Xu, J.K. Kresovich, K.G. Lawrence, et al., Air pollution and epigenetic aging among Black and White women in the US. *Environment International* 2023. **181**: 108270.
17. de Prado-Bert, P., C. Ruiz-Arenas, M. Vives-Usano, S. Andrusaityte, S. Cadiou, A. Carracedo, et al., The early-life exposome and epigenetic age acceleration in children. *Environ Int* 2021. **155**: 106683.
18. Ward-Caviness, C.K., A.G. Russell, A.M. Weaver, E. Slawsky, R. Dhingra, L.C. Kwee, et al., Accelerated epigenetic age as a biomarker of cardiovascular sensitivity to traffic-related air pollution. *Aging (Albany NY)* 2020. **12(23)**: 24141-24155.
19. Wang, C., P. Koutrakis, X. Gao, A. Baccarelli, and J. Schwartz, Associations of annual ambient PM2.5 components with DNAm PhenoAge acceleration in elderly men: The Normative Aging Study. *Environ Pollut* 2020. **258**: 113690.

20. Nwanaji-Enwerem, J.C., L. Dai, E. Colicino, Y. Oulhote, Q. Di, I. Kloog, et al., 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* 2017. **102**: 57-65.
21. Marioni, R.E., S. Shah, A.F. McRae, B.H. Chen, E. Colicino, S.E. Harris, et al., DNA methylation age of blood predicts all-cause mortality in later life. *Genome Biol* 2015. **16(1)**: 25.
22. Horvath, S., W. Erhart, M. Brosch, O. Ammerpohl, W. von Schönfels, M. Ahrens, et al., Obesity accelerates epigenetic aging of human liver. *Proc Natl Acad Sci U S A* 2014. **111(43)**: 15538-43.
23. Jylhävä, J., N.L. Pedersen, and S. Hägg, Biological Age Predictors. *EBioMedicine* 2017. **21**: 29-36.
24. Cardenas, A., S. Ecker, R.P. Fadadu, K. Huen, A. Orozco, L.M. McEwen, et al., Epigenome-wide association study and epigenetic age acceleration associated with cigarette smoking among Costa Rican adults. *Scientific Reports* 2022. **12(1)**: 4277.
25. Oblak, L., J. van der Zaag, A.T. Higgins-Chen, M.E. Levine, and M.P. Boks, A systematic review of biological, social and environmental factors associated with epigenetic clock acceleration. *Ageing Res Rev* 2021. **69**: 101348.
26. Gao, X., E. Colicino, J. Shen, A.C. Just, J.C. Nwanaji-Enwerem, C. Wang, et al., Accelerated DNA methylation age and the use of antihypertensive medication among older adults. *Aging (Albany NY)* 2018. **10(11)**: 3210-3228.
27. Kho, M., Y.Z. Wang, D. Chaar, W. Zhao, S.M. Ratliff, T.H. Mosley, et al., Accelerated DNA methylation age and medication use among African Americans. *Aging (Albany NY)* 2021. **13(11)**: 14604-14629.
28. Levine, M.E., A.T. Lu, B.H. Chen, D.G. Hernandez, A.B. Singleton, L. Ferrucci, et al., Menopause accelerates biological aging. *Proc Natl Acad Sci U S A* 2016. **113(33)**: 9327-32.
29. Plusquin, M., F. Guida, S. Polidoro, R. Vermeulen, O. Raaschou-Nielsen, G. Campanella, et al., DNA methylation and exposure to ambient air pollution in two prospective cohorts. *Environment International* 2017. **108**: 127-136.
30. Eze, I.C., A. Jeong, E. Schaffner, F.I. Rezwan, A. Ghantous, M. Foraster, et al., Genome-Wide DNA Methylation in Peripheral Blood and Long-Term Exposure to Source-Specific Transportation Noise and Air Pollution: The SAPALDIA Study. *Environmental Health Perspectives* 2020. **128(6)**: 067003.

458 **Table 1** Number of overlapping CpGs between the epigenetic aging biomarkers. CpGs for DNAmGrimAge
459 cannot be determined since it is calculated by a two-stage method.

	DNAmHorvathAge	DNAmHannumAge	DNAmPhenoAge	DNAmSkinBloodAge
DNAmHannumAge	6			
DNAmPhenoAge	41	6		
DNAmSkinBloodAge	60	5	58	
DNAmTL	0	45	1	5

460

Table 2 Descriptive statistics of participant characteristics by study wave (total number of observations: N=4105).

	S4 (N=1481)	F4 (N=1599)	FF4 (N=1025)	
Variable	**Mean ± SD / N (%)**	**Mean ± SD / N (%)**	**Mean ± SD / N (%)**	***p*–value**
Age (years)	53.9 ± 8.9	60.7 ± 8.9	64.7 ± 8.1	< 0.001
Sex (male)	736 (49.7)	783 (49.0)	495 (48.3)	0.78
Education				0.36
Primary school	873 (59.0)	919 (57.5)	567 (55.3)	
High school	338 (22.8)	368 (23.0)	240 (23.4)	
College	270 (18.2)	312 (19.5)	218 (21.3)	
BMI (kg/m^2)	27.7 ± 4.5	28.1 ± 4.8	28.2 ± 5.0	0.066
Alcohol consumption (g/day)	17.1 ± 22.1	15.7 ± 20.7	15.6 ± 20.5	0.036
Smoking status				< 0.001
Never smoker	633 (42.7)	673 (42.1)	421 (41.1)	
Current smoker	554 (37.4)	697 (43.6)	473 (46.1)	
Former smoker	294 (19.9)	229 (14.3)	131 (12.8)	
Physical activity				< 0.001
Low	483 (32.6)	492 (30.8)	279 (27.2)	
Medium	708 (47.8)	697 (43.6)	453 (44.2)	
High	290 (19.6)	410 (25.6)	293 (28.6)	
Hypertension (yes)	620 (41.9)	719 (45.0)	483 (47.1)	0.028
Diabetes (yes)	51 (3.4)	142 (8.9)	130 (12.7)	< 0.001
HDL cholesterol (mmol/l)	1.5 ± 0.4	1.5 ± 0.4	1.7 ± 0.5	< 0.001
LDL cholesterol (mmol/l)	3.7 ± 1.1	3.6 ± 0.9	3.6 ± 0.9	< 0.001
DNAmHorvathAge (years)[a]	55.2 ± 8.3 (60.8)	59.1 ± 7.5 (38.1)	65.0 ± 6.1 (54.0)	< 0.001
DNAmHannumAge (years)[a]	58.4 ± 9.3 (83.6)	70.4 ± 10.0 (96.5)	53.8 ± 7.7 (1.2)	< 0.001
DNAmPhenoAge (years)[a]	49.5 ± 11.0 (25.0)	54.0 ± 12.5 (21.1)	52.6 ± 9.7 (3.2)	< 0.001
DNAmGrimAge (years)[a]	55.3 ± 8.6 (56.9)	61.9 ± 8.8 (52.8)	64.0 ± 7.8 (36.8)	< 0.001
DNAmSkinBloodAge (years)[a]	54.2 ± 9.3 (53.4)	61.8 ± 9.2 (60.2)	59.5 ± 7.7 (7.4)	< 0.001
DNAmTL	7.1 ± 0.3	6.7 ± 0.3	6.8 ± 0.2	< 0.001
CD8 T cells	0.11 ± 0.05	0.1 ± 0.05	0.05 ± 0.04	< 0.001
CD4 T cells	0.22 ± 0.05	0.24 ± 0.05	0.19 ± 0.06	< 0.001
Natural killer cells	0.03 ± 0.02	0.06 ± 0.03	0.07 ± 0.04	< 0.001
B cells	0.06 ± 0.02	0.07 ± 0.02	0.05 ± 0.03	< 0.001
Monocytes	0.11 ± 0.02	0.1 ± 0.02	0.07 ± 0.02	< 0.001
Daytime noise (dB(A))[b]	54.84 ± 6.61 (0.7%)	54.68 ± 6.61 (0.6%)	54.36 ± 6.52 (6.9%)	0.38
Nighttime noise (dB(A))[b]	45.81 ± 6.41 (0.7%)	45.64 ± 6.41 (0.6%)	45.31 ± 6.32 (6.9%)	0.27

KORA = Cooperative Health Research in the Region of Augsburg; S4 = fourth cross-sectional health survey of the KORA cohort; F4 = first follow-up examination of KORA S4; FF4 = second follow-up examination of KORA S4; BMI = body mass index; HDL = high density lipoprotein; LDL= low density lipoprotein;

Physical activity was defined according to the exercise time per week: Low = almost or no sporting activity, Medium = regular/ irregular approx. 1 hour per week, High = regularly more than 2 hours per week.

Of 1646 participants in total, 833 attended two examinations, and 813 attended three examinations.

[a] Numbers in brackets indicate the percent of persons with positive age acceleration.

[b] Numbers in brackets indicate the missing percent of noise data in each examination.

p-value was based on the Kruskal-Wallis test for continuous variables, and Pearson's Chi-squared test for categorical variables.

Table 3 Descriptive statistics and Spearman correlation coefficients of air pollution concentrations (N=1646).

				Spearman correlation coefficients							
Pollutant	**Mean ± SD**	**Range**	**IQR**	**$PM_{2.5}$**	**PM_{10}**	**PNC**	**PM_{Coarse}**	**$PM_{2.5abs}$**	**O_3**	**NO_2**	**NO_x**
$PM_{2.5}$ ($\mu g/m^3$)	11.8 ± 1.0	8.4 - 14.3	1.4	1							
PM_{10} ($\mu g/m^3$)	16.6 ± 1.6	12.7 - 22.3	2.1	0.53	1						
PNC ($10^3/cm^3$)	7.3 ± 1.9	3.3 - 15.0	2.0	0.65	0.82	1					
PM_{coarse} ($\mu g/m^3$)	5.0 ± 1.0	2.6 - 8.5	1.4	0.57	0.78	0.76	1				
$PM_{2.5abs}$ ($10^{-5}/m$)	1.2 ± 0.2	0.8 - 1.9	0.3	0.62	0.78	0.78	0.81	1			
O_3 ($\mu g/m^3$)	39.1 ± 2.4	31.3 - 46.2	3.4	-0.19	0.01	-0.07	0.09	-0.15	1		
NO_2 ($\mu g/m^3$)	14.3 ± 4.6	6.9 - 28.2	7.2	0.72	0.72	0.77	0.83	0.87	-0.21	1	
NO_x ($\mu g/m^3$)	22.0 ± 7.6	4.0 - 50.5	8.8	0.76	0.75	0.90	0.76	0.74	-0.10	0.83	1

Of the total of 1646 participants, only 25 participants moved between the examinations. For participants who changed residence, the updated residential addresses were used for exposure assignment of the follow-up visits. Otherwise, the same exposure levels were assigned to the follow-up visits. $PM_{2.5}$ = particulate matter with an aerodynamic diameter less than or equal to 2.5 μm; PM_{10} = particulate matter with an aerodynamic diameter less than or equal to 10 μm; PNC = particle number concentration; PM_{coarse} = particulate matter with an aerodynamic diameter of 2.5-10 μm; $PM_{2.5abs}$ = $PM_{2.5}$ absorbance; PNC = particle number concentration; NO_2 = nitrogen dioxide; NO_x = nitrogen oxide; O_3 = ozone. Pearson's correlation coefficients were calculated to determine the correlations between air pollutants. **EU air quality standards**: 25 $\mu g/m^3$, 40 $\mu g/m^3$ for $PM_{2.5}$, PM_{10}, and NO_2, respectively. **WHO air quality guideline**: 5 $\mu g/m^3$, 10 $\mu g/m^3$, and 10 $\mu g/m^3$ for $PM_{2.5}$, PM_{10}, and NO_2, respectively.

485 **Table 4** Results from epigenome-wide association study (EWAS) restricted to CpG sites involved in
486 predicting the respective epigenetic clocks with PM_{coarse} (ever smokers).

CpG	CHR	Gene	Beta (mean ± SD)	Regression coefficient (β)	*p*-value	FDR	Epigenetic clocks
cg12745325	12	SLC39A5	0.64 ± 0.06	-0.005	6.6×10^{-5}	0.04	DNAmTL
cg24081819	8	EPHX2	0.10 ± 0.03	0.003	7.1×10^{-5}	0.04	DNAmHorvathAge

487 Linear mixed-effects models, adjusted for age, sex, an indicator of each study waves visit (KORA S4,
488 KORA F4, or KORA FF4), Houseman imputed cells, batch, educational attainment, alcohol consumption,
489 physical activity, BMI, hypertension, diabetes, HDL, and LDL.

490 CHR: chromosome; FDR: Benjamini-Hochberg false discovery rate.

491 DNA methylation in the 0–1 range can be divided in low-, moderate- and high-methylation with ranges [0–
492 0.35], [0.35–0.65] and [0.65–1] respectively. Covariates include those from the full model (see Figure 1).

493 Linear mixed-effects models were used with random participant-specific intercepts to examine the
494 associations between repeated epigenetic aging biomarkers and air pollutants. Batch and chip numbers were
495 included as random effects in the models to control the potential technical variations. Further covariates
496 were selected based on previous studies.

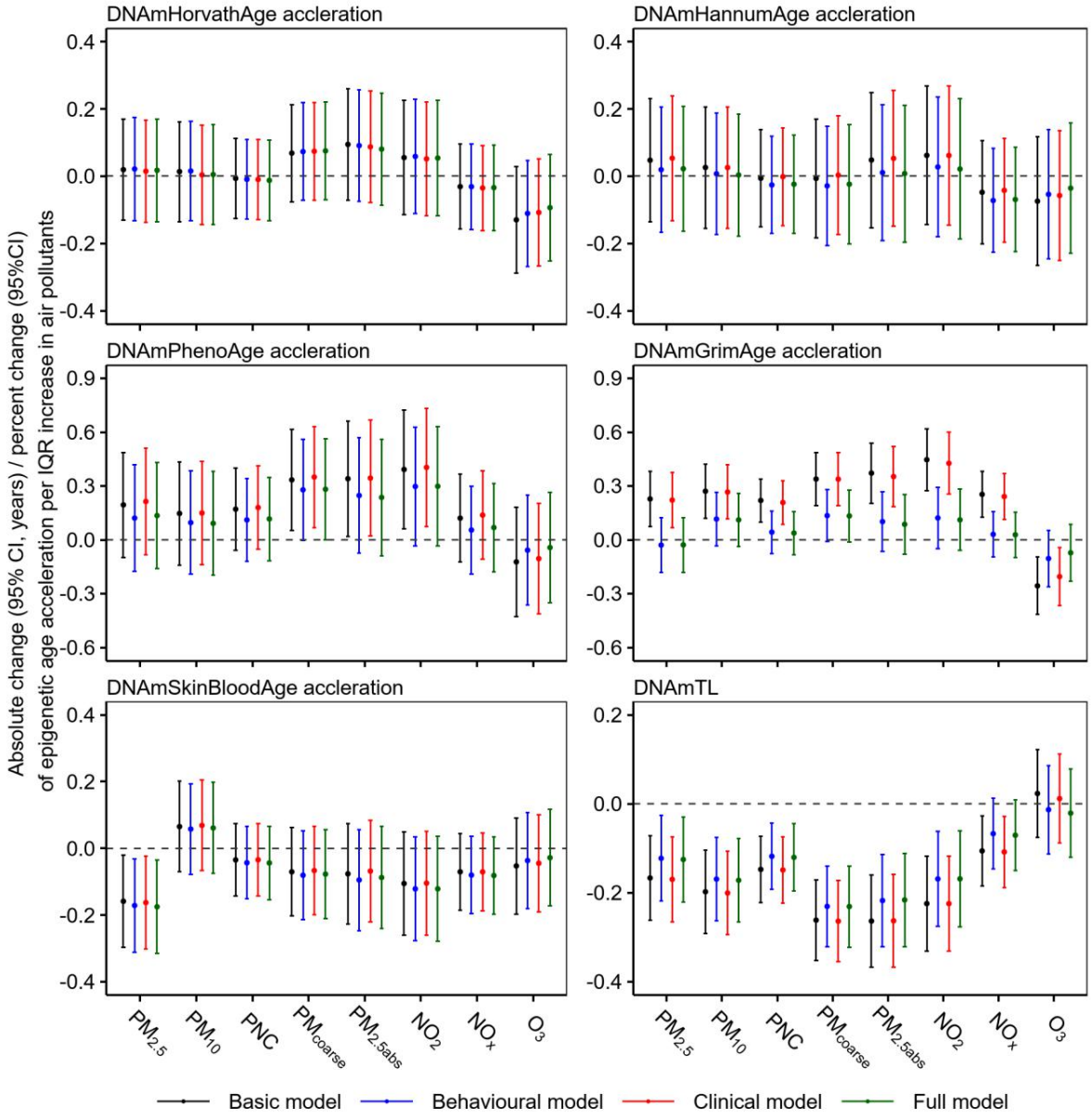


Figure 1. Absolute change (95% CI, years) of epigenetic age acceleration per IQR increase in air pollutant concentrations with basic, behavioral, clinical, and full covariate adjustment. For DNAmTL, the percent change (95% CI) is displayed. Covariate-adjusted linear mixed-effect regression models were used. Basic model: basic models were adjusted for age, sex, an indicator of each study wave (KORA S4, KORA F4, or KORA FF4), houseman estimated cells, batch and chip; Behavioral model: Basic model further adjusted for educational attainment, smoking status, alcohol consumption, physical activity; Clinical model: Basic model further adjusted for BMI, hypertension, diabetes, HDL, and LDL. Full model: all covariates from the behavioral and clinical model combined. $PM_{2.5}$ = particulate matter with an aerodynamic diameter less than or equal to 2.5 µm; PM_{10} = particulate matter with an aerodynamic diameter less than or equal to 10 µm; PNC = particle number concentration; PM_{coarse} = particulate matter with an aerodynamic diameter of 2.5-10 µm; $PM_{2.5abs}$ = $PM_{2.5}$ absorbance; NO_2 = nitrogen dioxide; NO_x = nitrogen oxide; O_3 = ozone. An IQR increase was 1.4 µg/m^3 for $PM_{2.5}$, 2.1 µg/m^3 for PM_{10}, 2.0×10^3/cm^3 for PNC, 1.4 µg/m^3 for PM_{coarse}, 0.3×10^{-5}/m for $PM_{2.5abs}$, 7.2 µg/m^3 for NO_2, 8.8 µg/m^3 for NO_x, and 3.4 µg/m^3 for O_3.

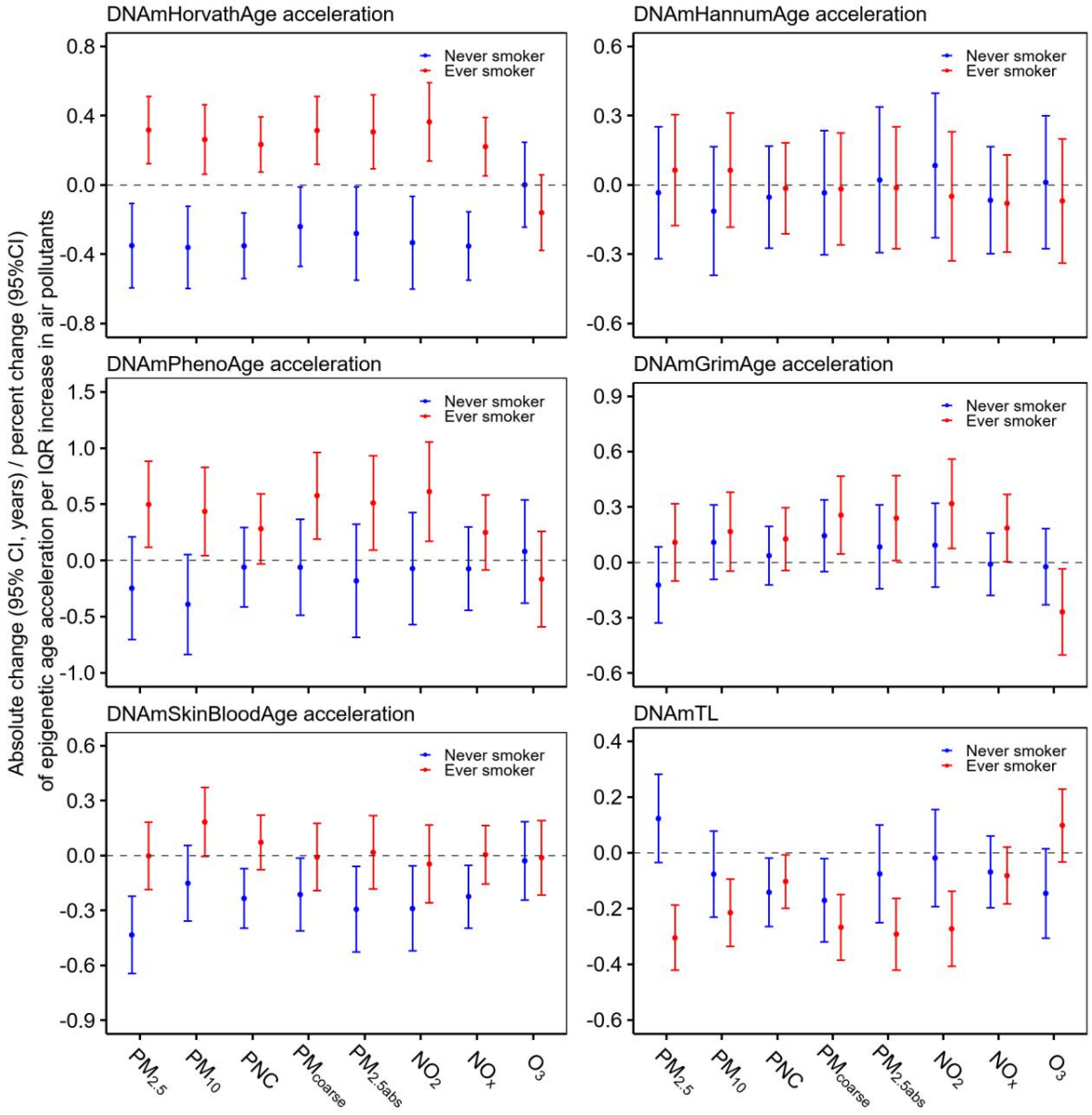


Figure 2. Absolute change (95% CI, years) / percent change (95%CI) of epigenetic aging biomarkers per IQR increase in air pollutant concentrations stratified by smoking status (percent change only for DNAmTL). Covariate-adjusted linear mixed-effect regression models were used. Results from the full model adjusting for age, sex, an indicator of each study waves visit (KORA S4, KORA F4, or KORA FF4), Houseman imputed cells, batch, educational attainment, alcohol consumption, physical activity, BMI, hypertension, diabetes, HDL, and LDL. Never smoking = participants who never smoked; Ever smoking = current or former smokers. $PM_{2.5}$ = particulate matter with an aerodynamic diameter less than or equal to 2.5 µm; PM_{10} = particulate matter with an aerodynamic diameter less than or equal to 10 µm; PNC = particle number concentration; PM_{coarse} = particulate matter with an aerodynamic diameter of 2.5-10 µm; $PM_{2.5abs}$ = $PM_{2.5}$ absorbance; NO_2 = nitrogen dioxide; NO_x = nitrogen oxide; O_3 = ozone. An IQR increase was 1.4 µg/m^3 for $PM_{2.5}$, 2.1 µg/m^3 for PM_{10}, 2.0×10^3/cm^3 for PNC, 1.4 µg/m^3 for PM_{coarse}, 0.3×10^{-5}/m for $PM_{2.5abs}$, 7.2 µg/m^3 for NO_2, 8.8 µg/m^3 for NO_x, and 3.4 µg/m^3 for O_3.

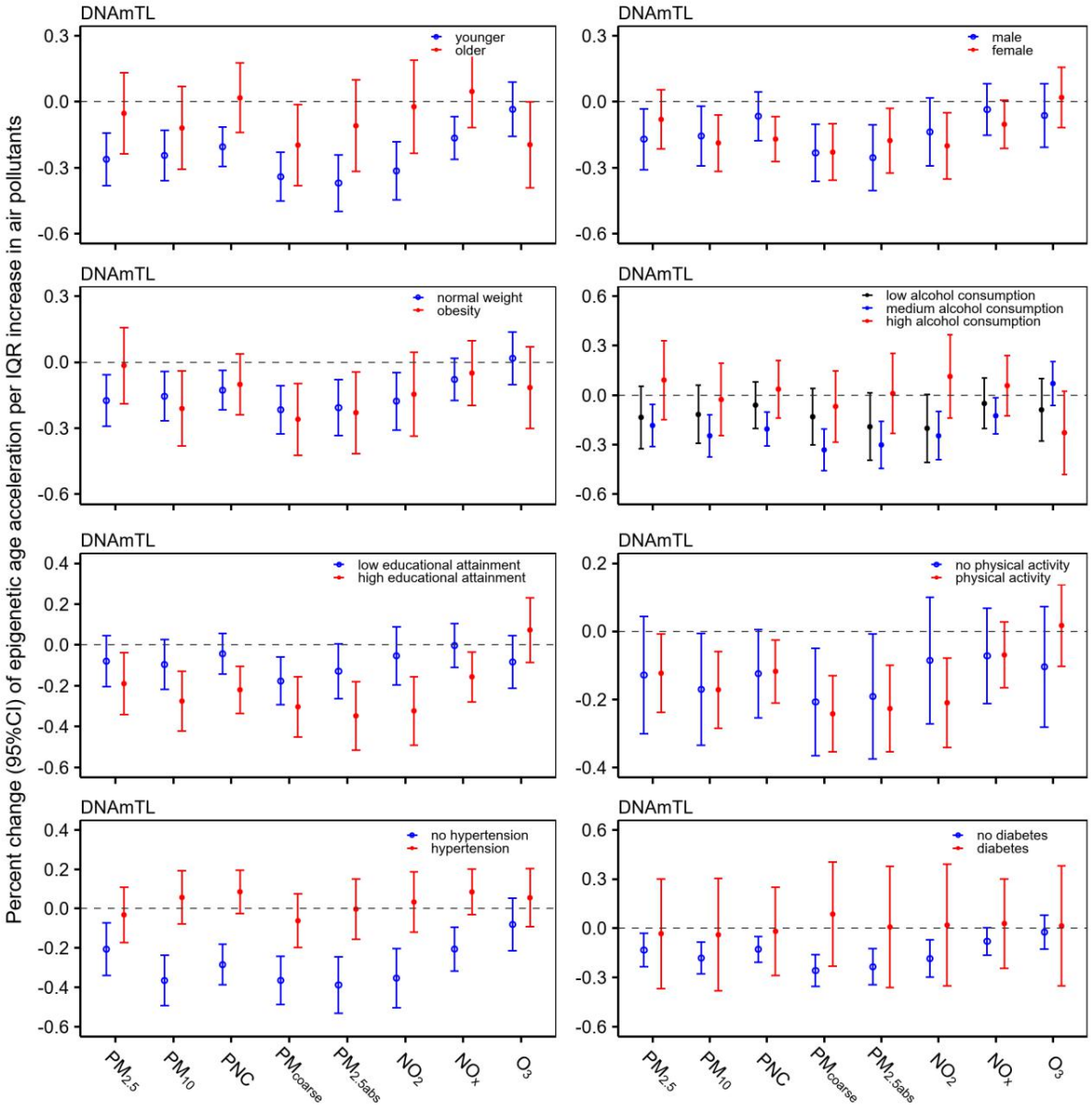


Figure 3. Percent change (95% CI) in DNAmTL per IQR increase in air pollutant concentrations stratified by categorized age group (≤ 65 vs > 65), sex, categorized BMI group, categorized group of alcohol consumption, educational attainment, physical activity, and history of diseases (hypertension, diabetes). Covariate-adjusted linear mixed-effect regression models were used. Results were from our full model adjusted by age, sex, an indicator of each study waves visit (KORA S4, KORA F4, or KORA FF4), houseman estimated cells, batch, chip, educational attainment, smoking status, alcohol consumption, physical activity, BMI, hypertension, diabetes, HDL, and LDL. $PM_{2.5}$ = particulate matter with an aerodynamic diameter less than or equal to 2.5 μm; PM_{10} = particulate matter with an aerodynamic diameter less than or equal to 10 μm; PNC = particle number concentration; PM_{coarse} = particulate matter with an aerodynamic diameter of 2.5-10 μm; $PM_{2.5abs}$ = $PM_{2.5}$ absorbance; NO_2 = nitrogen dioxide; NO_x = nitrogen oxide; O_3 = ozone. An IQR increase was 1.4 $\mu g/m^3$ for $PM_{2.5}$, 2.1 $\mu g/m^3$ for PM_{10}, $2.0\times10^3/cm^3$ for PNC, 1.4 $\mu g/m^3$ for PM_{coarse}, $0.3\times10^{-5}/m$ for $PM_{2.5abs}$, 7.2 $\mu g/m^3$ for NO_2, 8.8 $\mu g/m^3$ for NO_x, and 3.4 $\mu g/m^3$ for O_3.

Appendix

Long-term exposure to ambient air pollution is associated with epigenetic age acceleration

Yueli Yao [a,b], Kathrin Wolf [a], Susanne Breitner [a,b], Siqi Zhang [a,c], Melanie Waldenberger [a,d], Juliane Winkelmann [e,f], Alexandra Schneider [a]*, Annette Peters [a,b,g,h]*

[a] Institute of Epidemiology, Helmholtz Zentrum München GmbH, German Research Center for Environmental Health, Neuherberg, Germany

[b] Institute for Medical Information Processing, Biometry, and Epidemiology (IBE), Faculty of Medicine, LMU Munich, Pettenkofer School of Public Health, Munich, Germany

[c] Department of Environmental Health Sciences, Yale School of Public Health, New Haven, United States

[d] Research Unit Molecular Epidemiology, Helmholtz Zentrum München GmbH, German Research Center for Environmental Health, Neuherberg, Germany

[e] Institute of Neurogenomics, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany

[f] Institute of Human Genetics, Technical University of Munich, Munich, Germany

[g] German Center for Diabetes Research (DZD), Munich-Neuherberg, Germany

[h] German Centre for Cardiovascular Research (DZHK), Partner Site Munich, Munich, Germany

*Authors share last authorship

Correspondence to:

Yueli Yao, Institute of Epidemiology, Helmholtz Zentrum München GmbH, German Research Center for Environmental Health, Ingolstädter Landstr. 1, 85764 Neuherberg, Germany; Email: yueli.yao@helmholtz-munich.de.

Contents

Figure S10. Absolute changes (95% CI, years) in DNAmSkinBloodAge acceleration per IQR increase in air pollutant concentrations stratified by categorized age group (≤ 65 vs > 65), sex, categorized BMI group, categorized group of alcohol consumption, educational attainment, physical activity, and history of diseases (hypertension, diabetes).

Figure S11. Absolute change (95% CI, years) of epigenetic age acceleration per IQR increase in air pollutant concentrations with basic, behavioral, clinical, and full covariate adjustment. For DNAmTL, the percent change (95% CI) is displayed.

Figure S12. Comparison between single and two-pollutant models after additional inclusion of $PM_{2.5}$.

Figure S13. Comparison between single and two-pollutant models after additional inclusion of O_3.

Text S1. Definitions of individual characteristics

The continuous body mass index (BMI) was calculated as weight divided by height squared (kg/m^2) and was further categorized for effect modification analysis into normal weight (≤ 30 kg/m^2) and obesity ($>$ 30 kg/m^2). Educational attainment was included in the main analysis using three categories: primary school, high school, and college. For effect modification analysis, education was recategorized into high (higher than secondary school) and low (secondary school or less). Physical activity was categorized based on the amount of time spent on physical exercise per week, with participants classified as "no" (no or almost no physical exercise) or "yes" (regular or irregular activity of approximately one hour or more than two hours per week). Alcohol consumption was categorized into three groups: no consumption (0 g/day), moderate consumption (0.1–39.9 g/day for men and 0.1–19.9 g/day for women), and high consumption ($\geq$ 40 g/day for men and $\geq$ 20 g/day for women). Smoking status was initially categorized as smokers (regular or irregular), former smokers (ex-smokers), and never smokers; for the purpose of interaction analysis, it was recategorized into ever smokers (including regular, irregular, and former smokers) and never smokers. Diabetes and hypertension were defined as self-reported diagnoses that were confirmed by a physician or documented in medical records.

Text S2. Processing of DNA methylation data and prediction of epigenetic clocks

The bisulfite conversion and genome-wide methylation assessment were performed as previously described.[1] Further quality control and pre-processing of the data were performed on the raw methylation data of KORA S4, F4, and FF4, following the CPACOR pipeline,[2] starting with the exclusion of single-nucleotide polymorphism markers, background correction using the R package minfi,[3] and subsequently setting probes to NA if the signals had a detection p-value of > 0.01 or were summarized from ≤ 3 functional beads. Quantile normalization was performed on the signal intensity values, divided into categories by probe type and colour channel. A β-value, representing the methylation level of a given cytosine, was calculated by the ratio of the methylated signal intensity to the sum of the methylated and unmethylated signal intensities and used for further estimates of epigenetic aging biomarkers.

The Horvath clock was trained in 51 different tissue types, including 353 CpGs.[4] The Hannum and Levine clocks were developed and trained from whole blood samples consisting of 71 and 513 CpG sites, respectively.[5,6] The skin & blood clock was calculated by 391 CpGs and trained on different tissue and cell types, e.g., endothelial cells, skin, as well as blood.[7] The DNAmTL estimator was predicted by 140 CpG sites based on Lu's method.[8] In addition, the white blood cell proportions (monocytes, granulocytes, natural killer cells, CD4 T cells, CD8 T cells, and B cells) were estimated by the Houseman method.[9]

Reference:

1. Zeilinger, S, Kuhnel, B, Klopp, N, Baurecht, H, Kleinschmidt, A, Gieger, C, et al. 2013. Tobacco smoking leads to extensive genome-wide changes in DNA methylation. PLoS One. 8(5): e63812. doi:10.1371/journal.pone.0063812.
2. Lehne, B, Drong, AW, Loh, M, Zhang, W, Scott, WR, Tan, ST, et al. 2015. A coherent approach for analysis of the Illumina HumanMethylation450 BeadChip improves data quality and performance in epigenome-wide association studies. Genome Biol. 16(1): 37. 25853392, 10.1186/s13059-015-0600-x.
3. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, Irizarry RA. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics. 2014;30(10):1363-9. doi: 10.1093/bioinformatics/btu049.
4. Horvath, S. 2013. DNA methylation age of human tissues and cell types. Genome Biol.14: R115. doi:10.1186/gb-2013-14-10-r115.
5. Hannum, G, Guinney, J, Zhao, L, Zhang, L, Hughes, G, Sadda, S, et al. 2013. Genome-wide methylation profiles reveal quantitative views of human aging rates. Mol Cell. 49(2): 359-367. doi:10.1016/j.molcel.2012.10.016.
6. Levine, Morgan E, Lu, Ake T, Quach, A, Chen, Brian HC, Assimes, Themistocles L, Bandinelli, S, et al. 2018. An epigenetic biomarker of aging for lifespan and healthspan. Aging (Albany NY). 10: 573-591. doi:10.18632/aging.101414.
7. Horvath, S, Oshima, J, Martin, GM, Lu, AT, Quach, A, Cohen, H, et al. 2018. Epigenetic clock for skin and blood cells applied to Hutchinson Gilford Progeria Syndrome and ex vivo studies. Aging (Albany NY). 10(7): 1758-1775. doi:10.18632/aging.101508.
8. Lu, AT, Seeboth, A, Tsai, PC, Sun, D, Quach, A, Reiner, AP, et al. 2019. DNA methylation-based estimator of telomere length. Aging (Albany NY). 11(16): 5895-5923. doi:10.18632/aging.102173.
9. Houseman, EA, Accomando, WP, Koestler, DC, Christensen, BC, Marsit, CJ, Nelson, HH, et al. 2012. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinformatics. 13(1): 86. doi:10.1186/1471-2105-13-86.

Text S3. Exposure assessment

Air pollution

The land-use regression (LUR) models used to estimate annual air pollutant concentrations were developed and validated following the methodology outlined by Wolf et al.[10] Model performance was evaluated using leave-one-out cross-validation (LOOCV), a robust method for assessing predictive accuracy.

To capture seasonal variation in pollutant levels, air pollution measurements were conducted at 20 monitoring sites within the KORA study area. These measurements took place in three bi-weekly campaigns between March 2014 and April 2015, each representing a different seasonal condition—warm, cold, and intermediate. In parallel, continuous monitoring was conducted at a designated reference site throughout the entire study period. This reference data was used to adjust for temporal variability, ensuring that seasonal fluctuations did not bias the annual average estimates.

Annual average concentrations of various air pollutants were subsequently calculated at each monitoring site. These values served as the dependent variables in the LUR model, which was developed by regressing the measured pollutant concentrations against a comprehensive set of geographic information system (GIS)-based spatial predictors.[10] The spatial predictors included detailed land-use characteristics (such as the proportion of residential, industrial, commercial, transportation infrastructure, urban green spaces, and water bodies), population and household density, building density, topographic features, and geographic coordinates. Additionally, traffic-related variables were incorporated, including total traffic load within defined buffer zones, traffic intensity on the nearest major roads, and the intensity of heavy-duty vehicle traffic.[10]

Noise

Long-term traffic noise exposure was estimated using the noise- and air-pollution information system (http://www.laermkarten.de/augsburg/), developed by ACCON GmbH. This system incorporates detailed three-dimensional ground model on roads and buildings, with rural road networks derived from Google Earth and OpenStreetMap. Traffic data were obtained from local government sources including the Bavarian Ministry of the Interior, Building and Transport, the digital street map of Augsburg, several traffic censuses and surveys. Noise levels were modeled at four meters above ground for 2009 (urban) or 2000–2011 (rural). Maximum annual A-weighted equivalent continuous sound pressure levels [dB(A) Leq] for the full day (24 h) and nighttime (22:00–06:00) were estimated for each participant's residential address. If the address unavailable, the noise level from the nearest available building was assigned.[11]

Reference:

10. Wolf, K., J. Cyrys, T. Harciníková, J. Gu, T. Kusch, R. Hampel, et al., Land use regression modeling of ultrafine particles, ozone, nitrogen oxides and markers of particulate matter pollution in Augsburg, Germany. Sci Total Environ 2017. 579: 1531-1540. doi: 10.1016/j.scitotenv.2016.11.160
11. Pitchika, A, Hampel, R, Wolf, K, Kraus, U, Cyrys, J, Babisch, W, et al. 2017. Long-term associations of modeled and self-reported measures of exposure to air pollution and noise at residence on prevalent hypertension and blood pressure. Science of The Total Environment. 593-594: 337-346. doi: 10.1016/j.scitotenv.2017.03.156.

Table S1 Absolute/percent changes (95% CI) in epigenetic aging biomarkers per IQR increase in air pollutants.

Epigenetic aging biomarkers	Exposure	Basic	Behavioral	Clinical	Full
Acceleration of DNAmGrimAge	$PM_{2.5}$	0.23 (0.08, 0.38); $p = 0.01$	---	0.22 (0.07, 0.38); $p = 0.02$	---
	PM_{10}	0.27 (0.12, 0.42); $p = 0.01$	---	0.27 (0.12, 0.42); $p = 0.002$	---
	PNC	0.22 (0.1, 0.34); $p = 0.002$	---	0.21 (0.09, 0.33); $p = 0.003$	---
	PM_{coarse}	0.34 (0.19, 0.49); $p = 8.0\times10^{-5}$	---	0.34 (0.19, 0.49); $p = 9.0\times10^{-5}$	---
	$PM_{2.5abs}$	0.37 (0.2, 0.54); $p = 0.0001$	---	0.35 (0.18, 0.52); $p = 0.0003$	---
	NO_2	0.45 (0.27, 0.62); $p = 1.0\times10^{-5}$	---	0.43 (0.25, 0.6); $p = 2.0\times10^{-5}$	---
	NO_x	0.25 (0.13, 0.38); $p = 0.001$	---	0.24 (0.11, 0.37); $p = 0.001$	---
DNAmTL	$PM_{2.5}$	-0.2 (-0.3, -0.1); $p = 0.0007$	-0.1 (-0.2, 0); $p = 0.02$	-0.2 (-0.3, -0.1); $p = 0.0007$	-0.1 (-0.2, 0); $p = 0.01$
	PM_{10}	-0.2 (-0.3, -0.1); $p = 0.00008$	-0.2 (-0.3, -0.1); $p = 0.001$	-0.2 (-0.3, -0.1); $p = 7.0\times10^{-5}$	-0.2 (-0.3, 0); $p = 0.0009$
	PNC	-0.1 (-0.2, -0.1); $p = 0.0002$	-0.1 (-0.2, 0); $p = 0.003$	-0.1 (-0.2, -0.1); $p = 0.0002$	-0.1 (-0.2, 0); $p = 0.03$
	PM_{coarse}	-0.3 (-0.4, -0.2); $p = 1.4\times10^{-7}$	-0.2 (-0.3, -0.1); $p = 5.9\times10^{-6}$	-0.3 (-0.4, -0.2); $p = 1.3\times10^{-7}$	-0.2 (-0.3, -0.1); $p = 6.0\times10^{-6}$
	$PM_{2.5abs}$	-0.3 (-0.4, -0.2); $p = 2.7\times10^{-6}$	-0.2 (-0.3, -0.1); $p = 0.0002$	-0.3 (-0.4, -0.2); $p = 3.4\times10^{-6}$	-0.2 (-0.3, -0.1); $p = 0.0002$
	NO_2	-0.2 (-0.3, -0.1); $p = 7.6\times10^{-5}$	-0.2 (-0.3, -0.1); $p = 0.003$	-0.2 (-0.3, -0.1); $p = 7.9\times10^{-5}$	-0.2 (-0.3, -0.1); $p = 0.003$
	NO_x	-0.1 (-0.2, 0); $p = 0.01$	---	-0.1 (-0.2, 0); $p = 0.01$	---

Covariate-adjusted linear mixed-effect regression models were used. The effect estimates of acceleration of DNAmGrimAge and DNAmTL were presented as absolute change (together 95% CI) and percent change, respectively. Basic model: adjusted for age, sex, an indicator of each study wave (KORA S4, KORA F4, or KORA FF4), houseman estimated cells, batch, and chip; Behavioral model: further adjusted for educational attainment, smoking status, alcohol consumption, physical activity into basic model; Clinical model: additionally included BMI, hypertension, diabetes, HDL, and LDL into the basic models. Full model: all confounders used in the other three models. $PM_{2.5}$ = particulate matter with an aerodynamic diameter less than or equal to 2.5 μm; PM_{10} = particulate matter with an aerodynamic diameter less than or equal to 10 μm; PNC = particle number concentration; PM_{coarse} = particulate matter with an aerodynamic diameter of 2.5-10 μm; $PM_{2.5abs}$ = $PM_{2.5}$ absorbance; NO_2 = nitrogen dioxide; NO_x = nitrogen oxide. An IQR increase was 1.4 $\mu g/m^3$ for $PM_{2.5}$, 2.1 $\mu g/m^3$ for PM_{10}, $2.0\times10^3/cm^3$ for PNC, 1.4 $\mu g/m^3$ for PM_{coarse}, $0.3\times10^{-5}/m$ for $PM_{2.5abs}$, 7.2 $\mu g/m^3$ for NO_2, and 8.8 $\mu g/m^3$ for NO_x.

Table S2 Effect estimates (95% CI) of the confounders used in the full models for each epigenetic aging biomarker.

	Air pollution	age	sex	Smoking		alcohol	Physical activity	Education level		BMI	Hypertension	diabetes	HDL	LDL
				Former smoker	Current smoker			High school	College					
Acceleration of DNAmHorvathAge	$PM_{2.5}$	-0.32 (-0.34, -0.31)	-0.85 (-1.13, -0.57)	-0.3 (-0.55, -0.06)	-0.25 (-0.6, 0.09)	0 (0, 0.01)	-0.12 (-0.37, 0.13)	0.4 (0.13, 0.67)	-0.13 (-0.43, 0.18)	0.05 (0.02, 0.07)	0.12 (-0.12, 0.36)	-0.22 (-0.65, 0.22)	-0.34 (-0.65, -0.04)	0.14 (0.03, 0.26)
	PM_{10}	-0.32 (-0.34, -0.31)	-0.85 (-1.13, -0.57)	-0.3 (-0.54, -0.05)	-0.25 (-0.59, 0.09)	0 (0, 0.01)	-0.12 (-0.37, 0.13)	0.4 (0.13, 0.67)	-0.13 (-0.43, 0.18)	0.05 (0.02, 0.07)	0.12 (-0.12, 0.36)	-0.22 (-0.65, 0.22)	-0.34 (-0.65, -0.04)	0.14 (0.03, 0.26)
	PNC	-0.32 (-0.34, -0.31)	-0.85 (-1.13, -0.57)	-0.3 (-0.54, -0.05)	-0.25 (-0.59, 0.09)	0 (0, 0.01)	-0.12 (-0.37, 0.13)	0.4 (0.13, 0.67)	-0.13 (-0.43, 0.18)	0.05 (0.02, 0.07)	0.12 (-0.12, 0.36)	-0.22 (-0.65, 0.22)	-0.35 (-0.65, -0.04)	0.14 (0.03, 0.26)
	PM_{coarse}	-0.33 (-0.34, -0.31)	-0.86 (-1.14, -0.58)	-0.31 (-0.55, -0.06)	-0.27 (-0.61, 0.07)	0 (0, 0.01)	-0.12 (-0.37, 0.13)	0.39 (0.12, 0.67)	-0.14 (-0.44, 0.17)	0.05 (0.02, 0.07)	0.12 (-0.12, 0.36)	-0.22 (-0.65, 0.21)	-0.34 (-0.65, -0.04)	0.14 (0.03, 0.26)
	$PM_{2.5abs}$	-0.33 (-0.34, -0.31)	-0.86 (-1.14, -0.58)	-0.31 (-0.55, -0.06)	-0.26 (-0.61, 0.08)	0 (0, 0.01)	-0.12 (-0.37, 0.13)	0.39 (0.12, 0.67)	-0.13 (-0.43, 0.17)	0.05 (0.02, 0.07)	0.12 (-0.12, 0.36)	-0.22 (-0.65, 0.21)	-0.34 (-0.65, -0.03)	0.14 (0.03, 0.26)
	NO_2	-0.32 (-0.34, -0.31)	-0.86 (-1.14, -0.57)	-0.3 (-0.55, -0.06)	-0.26 (-0.6, 0.08)	0 (0, 0.01)	-0.12 (-0.37, 0.13)	0.4 (0.12, 0.67)	-0.13 (-0.43, 0.18)	0.05 (0.02, 0.07)	0.12 (-0.12, 0.36)	-0.22 (-0.65, 0.22)	-0.34 (-0.65, -0.03)	0.14 (0.03, 0.26)
	NO_x	-0.32 (-0.34, -0.31)	-0.85 (-1.13, -0.57)	-0.29 (-0.54, -0.05)	-0.24 (-0.59, 0.1)	0 (0, 0.01)	-0.12 (-0.37, 0.13)	0.4 (0.13, 0.67)	-0.13 (-0.43, 0.18)	0.05 (0.02, 0.07)	0.12 (-0.12, 0.36)	-0.22 (-0.65, 0.22)	-0.35 (-0.66, -0.04)	0.14 (0.03, 0.26)
	O_3	-0.32 (-0.34, -0.31)	-0.85 (-1.13, -0.57)	-0.29 (-0.54, -0.05)	-0.25 (-0.59, 0.09)	0 (0, 0.01)	-0.12 (-0.37, 0.13)	0.4 (0.13, 0.68)	-0.12 (-0.42, 0.19)	0.05 (0.02, 0.07)	0.12 (-0.12, 0.36)	-0.22 (-0.65, 0.22)	-0.34 (-0.64, -0.03)	0.14 (0.03, 0.26)
Acceleration of DNAmHannumAge	$PM_{2.5}$	-0.16 (-0.18, -0.15)	-1.73 (-2.07, -1.39)	0.09 (-0.21, 0.39)	0.47 (0.06, 0.89)	0.01 (0.01, 0.02)	-0.11 (-0.41, 0.19)	0.05 (-0.28, 0.39)	-0.32 (-0.69, 0.05)	0.05 (0.02, 0.09)	-0.04 (-0.34, 0.25)	-0.58 (-1.11, -0.05)	-0.21 (-0.59, 0.16)	-0.18 (-0.32, -0.03)
	PM_{10}	-0.16 (-0.18, -0.15)	-1.73 (-2.07, -1.39)	0.09 (-0.2, 0.39)	0.48 (0.06, 0.89)	0.01 (0.01, 0.02)	-0.11 (-0.41, 0.19)	0.05 (-0.28, 0.39)	-0.32 (-0.69, 0.05)	0.05 (0.02, 0.09)	-0.04 (-0.34, 0.25)	-0.58 (-1.11, -0.05)	-0.21 (-0.59, 0.16)	-0.18 (-0.32, -0.03)
	PNC	-0.16 (-0.18, -0.15)	-1.73 (-2.07, -1.38)	0.1 (-0.2, 0.39)	0.48 (0.07, 0.9)	0.01 (0.01, 0.02)	-0.11 (-0.41, 0.19)	0.06 (-0.28, 0.39)	-0.32 (-0.69, 0.05)	0.05 (0.02, 0.09)	-0.04 (-0.34, 0.25)	-0.58 (-1.1, -0.05)	-0.22 (-0.59, 0.16)	-0.18 (-0.32, -0.03)
	PM_{coarse}	-0.16 (-0.18, -0.15)	-1.73 (-2.07, -1.38)	0.1 (-0.2, 0.39)	0.48 (0.07, 0.9)	0.01 (0.01, 0.02)	-0.11 (-0.41, 0.19)	0.06 (-0.28, 0.39)	-0.32 (-0.69, 0.06)	0.05 (0.02, 0.09)	-0.04 (-0.34, 0.25)	-0.58 (-1.1, -0.05)	-0.22 (-0.59, 0.16)	-0.18 (-0.32, -0.03)
	$PM_{2.5abs}$	-0.16 (-0.18, -0.15)	-1.73 (-2.07, -1.39)	0.09 (-0.2, 0.39)	0.48 (0.06, 0.89)	0.01 (0.01, 0.02)	-0.11 (-0.41, 0.19)	0.05 (-0.28, 0.39)	-0.32 (-0.69, 0.05)	0.05 (0.02, 0.09)	-0.04 (-0.34, 0.25)	-0.58 (-1.11, -0.05)	-0.21 (-0.59, 0.16)	-0.18 (-0.32, -0.03)
	NO_2	-0.16 (-0.18, -0.15)	-1.73 (-2.07, -1.39)	0.09 (-0.21, 0.39)	0.47 (0.06, 0.89)	0.01 (0.01, 0.02)	-0.11 (-0.41, 0.19)	0.05 (-0.28, 0.39)	-0.32 (-0.69, 0.05)	0.05 (0.02, 0.09)	-0.04 (-0.33, 0.25)	-0.58 (-1.1, -0.05)	-0.21 (-0.59, 0.16)	-0.18 (-0.32, -0.03)
	NO_x	-0.16 (-0.18, -0.15)	-1.72 (-2.06, -1.38)	0.1 (-0.2, 0.4)	0.49 (0.08, 0.91)	0.01 (0.01, 0.02)	-0.11 (-0.41, 0.19)	0.06 (-0.28, 0.39)	-0.32 (-0.69, 0.05)	0.05 (0.02, 0.08)	-0.05 (-0.34, 0.25)	-0.58 (-1.1, -0.05)	-0.22 (-0.6, 0.15)	-0.17 (-0.32, -0.03)
	O_3	-0.16 (-0.18, -0.15)	-1.73 (-2.07, -1.39)	0.1 (-0.2, 0.39)	0.48 (0.06, 0.89)	0.01 (0.01, 0.02)	-0.11 (-0.41, 0.2)	0.06 (-0.28, 0.39)	-0.32 (-0.69, 0.06)	0.05 (0.02, 0.09)	-0.04 (-0.34, 0.25)	-0.58 (-1.11, -0.05)	-0.21 (-0.59, 0.16)	-0.18 (-0.32, -0.03)
Acceleration of DNAmPhneoAge	$PM_{2.5}$	-0.11 (-0.14, -0.08)	-0.75 (-1.29, -0.2)	0.22 (-0.26, 0.69)	1.37 (0.71, 2.04)	0.03 (0.01, 0.04)	-0.46 (-0.94, 0.02)	0.27 (-0.26, 0.8)	-0.7 (-1.28, -0.11)	0.1 (0.05, 0.15)	0.52 (0.05, 0.98)	0.37 (-0.46, 1.21)	-0.2 (-0.79, 0.4)	0 (-0.23, 0.22)

	PM_{10}	-0.11 (-0.14, -0.08)	-0.74 (-1.28, -0.2)	0.23 (-0.25, 0.7)	1.39 (0.73, 2.05)	0.03 (0.01, 0.04)	-0.46 (-0.95, 0.02)	0.27 (-0.26, 0.8)	-0.7 (-1.29, -0.11)	0.1 (0.05, 0.15)	0.51 (0.05, 0.98)	0.37 (-0.47, 1.21)	-0.2 (-0.8, 0.39)	0 (-0.23, 0.22)
	PNC	-0.11 (-0.14, -0.08)	-0.74 (-1.29, -0.2)	0.22 (-0.25, 0.7)	1.37 (0.71, 2.04)	0.03 (0.01, 0.04)	-0.46 (-0.94, 0.02)	0.27 (-0.27, 0.8)	-0.7 (-1.29, -0.11)	0.1 (0.05, 0.15)	0.51 (0.05, 0.98)	0.37 (-0.47, 1.2)	-0.2 (-0.79, 0.4)	0 (-0.23, 0.22)
	PM_{coarse}	-0.11 (-0.14, -0.08)	-0.76 (-1.3, -0.22)	0.21 (-0.27, 0.68)	1.34 (0.67, 2)	0.03 (0.01, 0.04)	-0.46 (-0.94, 0.02)	0.25 (-0.29, 0.78)	-0.74 (-1.33, -0.15)	0.1 (0.05, 0.15)	0.51 (0.04, 0.98)	0.36 (-0.48, 1.2)	-0.2 (-0.8, 0.39)	0 (-0.23, 0.22)
	$PM_{2.5abs}$	-0.11 (-0.14, -0.08)	-0.75 (-1.29, -0.21)	0.21 (-0.26, 0.69)	1.36 (0.7, 2.02)	0.03 (0.01, 0.04)	-0.46 (-0.94, 0.02)	0.25 (-0.28, 0.79)	-0.71 (-1.3, -0.12)	0.1 (0.05, 0.15)	0.51 (0.05, 0.98)	0.36 (-0.47, 1.2)	-0.19 (-0.78, 0.41)	0 (-0.23, 0.22)
	NO_2	-0.11 (-0.14, -0.08)	-0.76 (-1.3, -0.21)	0.2 (-0.27, 0.68)	1.34 (0.68, 2.01)	0.03 (0.01, 0.04)	-0.46 (-0.94, 0.3)	0.26 (-0.27, 0.8)	-0.71 (-1.3, -0.12)	0.1 (0.05, 0.15)	0.52 (0.05, 0.98)	0.37 (-0.46, 1.21)	-0.2 (-0.8, 0.4)	0 (-0.23, 0.22)
	NO_x	-0.11 (-0.14, -0.08)	-0.74 (-1.28, -0.2)	0.23 (-0.25, 0.7)	1.39 (0.72, 2.05)	0.03 (0.01, 0.04)	-0.46 (-0.95, 0.02)	0.27 (-0.26, 0.81)	-0.69 (-1.28, -0.11)	0.1 (0.05, 0.15)	0.51 (0.05, 0.98)	0.37 (-0.47, 1.21)	-0.2 (-0.8, 0.4)	0 (-0.23, 0.22)
	O_3	-0.11 (-0.14, -0.08)	-0.73 (-1.28, -0.19)	0.24 (-0.24, 0.71)	1.4 (0.74, 2.06)	0.03 (0.01, 0.04)	-0.46 (-0.94, 0.02)	0.28 (-0.26, 0.81)	-0.69 (-1.28, -0.1)	0.1 (0.05, 0.15)	0.51 (0.05, 0.98)	0.37 (-0.46, 1.21)	-0.18 (-0.78, 0.41)	0 (-0.23, 0.23)
Aacceleration of DNAmGrimAge	$PM_{2.5}$	-0.23 (-0.24, -0.22)	-2.36 (-2.64, -2.08)	2.01 (1.77, 2.25)	5.91 (5.57, 6.25)	0.01 (0, 0.01)	-0.52 (-0.77, -0.27)	-0.27 (-0.55, 0)	-0.86 (-1.16, -0.56)	0.07 (0.05, 0.1)	-0.03 (-0.27, 0.21)	0.27 (-0.16, 0.7)	-0.46 (-0.77, -0.15)	-0.1 (-0.22, 0.02)
	PM_{10}	-0.23 (-0.24, -0.22)	-2.37 (-2.65, -2.09)	2 (1.75, 2.24)	5.88 (5.54, 6.22)	0.01 (0, 0.01)	-0.52 (-0.77, -0.27)	-0.28 (-0.55, -0.01)	-0.86 (-1.17, -0.56)	0.07 (0.05, 0.1)	-0.02 (-0.26, 0.22)	0.27 (-0.16, 0.7)	-0.46 (-0.76, -0.15)	-0.1 (-0.22, 0.02)
	PNC	-0.23 (-0.24, -0.22)	-2.37 (-2.65, -2.09)	2 (1.76, 2.25)	5.89 (5.55, 6.23)	0.01 (0, 0.01)	-0.52 (-0.77, -0.27)	-0.28 (-0.55, 0)	-0.86 (-1.16, -0.56)	0.07 (0.05, 0.1)	-0.03 (-0.26, 0.21)	0.27 (-0.16, 0.7)	-0.46 (-0.76, -0.15)	-0.1 (-0.22, 0.02)
	PM_{coarse}	-0.23 (-0.24, -0.22)	-2.38 (-2.66, -2.1)	1.99 (1.75, 2.24)	5.87 (5.53, 6.21)	0.01 (0, 0.01)	-0.52 (-0.76, -0.27)	-0.29 (-0.56, -0.01)	-0.88 (-1.19, -0.58)	0.07 (0.05, 0.1)	-0.03 (-0.27, 0.21)	0.26 (-0.17, 0.69)	-0.46 (-0.76, -0.15)	-0.1 (-0.22, 0.02)
	$PM_{2.5abs}$	-0.23 (-0.24, -0.22)	-2.37 (-2.65, -2.09)	2 (1.76, 2.24)	5.89 (5.55, 6.23)	0.01 (0, 0.01)	-0.52 (-0.77, -0.27)	-0.28 (-0.55, -0.01)	-0.87 (-1.17, -0.56)	0.07 (0.05, 0.1)	-0.02 (-0.26, 0.21)	0.26 (-0.17, 0.7)	-0.45 (-0.76, -0.14)	-0.1 (-0.22, 0.02)
	NO_2	-0.23 (-0.24, -0.22)	-2.37 (-2.65, -2.09)	2 (1.75, 2.24)	5.88 (5.54, 6.22)	0.01 (0, 0.01)	-0.52 (-0.76, -0.27)	-0.28 (-0.55, 0)	-0.87 (-1.17, -0.56)	0.07 (0.05, 0.1)	-0.02 (-0.26, 0.22)	0.27 (-0.16, 0.7)	-0.45 (-0.76, -0.14)	-0.1 (-0.22, 0.02)
	NO_x	-0.23 (-0.24, -0.22)	-2.37 (-2.65, -2.09)	2 (1.76, 2.25)	5.89 (5.55, 6.23)	0.01 (0, 0.01)	-0.52 (-0.77, -0.27)	-0.27 (-0.55, 0)	-0.86 (-1.16, -0.56)	0.07 (0.05, 0.1)	-0.02 (-0.26, 0.21)	0.27 (-0.16, 0.7)	-0.46 (-0.76, -0.15)	-0.1 (-0.22, 0.02)
	O_3	-0.23 (-0.24, -0.22)	-2.37 (-2.65, -2.09)	2.01 (1.77, 2.25)	5.9 (5.56, 6.24)	0.01 (0, 0.01)	-0.52 (-0.76, -0.27)	-0.27 (-0.54, 0)	-0.85 (-1.16, -0.55)	0.07 (0.05, 0.1)	-0.02 (-0.26, 0.22)	0.27 (-0.16, 0.7)	-0.45 (-0.76, -0.14)	-0.1 (-0.22, 0.02)
Acceleration of DNAmSkinBloodAge	$PM_{2.5}$	-0.09 (-0.1, -0.08)	-0.49 (-0.75, -0.23)	0.15 (-0.08, 0.37)	-0.02 (-0.33, 0.29)	0 (0, 0.01)	-0.16 (-0.39, 0.07)	0.1 (-0.15, 0.35)	-0.21 (-0.49, 0.06)	0.01 (-0.01, 0.04)	-0.09 (-0.31, 0.13)	-0.36 (-0.75, 0.04)	0.05 (-0.24, 0.33)	0.08 (-0.03, 0.18)
	PM_{10}	-0.09 (-0.1, -0.08)	-0.51 (-0.76, -0.25)	0.12 (-0.1, 0.34)	-0.06 (-0.38, 0.25)	0 (0, 0.01)	-0.16 (-0.39, 0.07)	0.09 (-0.16, 0.34)	-0.21 (-0.49, 0.06)	0.01 (-0.01, 0.04)	-0.08 (-0.3, 0.13)	-0.36 (-0.75, 0.04)	0.06 (-0.22, 0.34)	0.07 (-0.03, 0.18)
	PNC	-0.09 (-0.1, -0.08)	-0.5 (-0.76, -0.24)	0.13 (-0.09, 0.35)	-0.04 (-0.36, 0.27)	0 (0, 0.01)	-0.16 (-0.39, 0.07)	0.09 (-0.16, 0.35)	-0.21 (-0.49, 0.07)	0.01 (-0.01, 0.04)	-0.09 (-0.31, 0.13)	-0.35 (-0.75, 0.04)	0.05 (-0.23, 0.33)	0.07 (-0.03, 0.18)
	PM_{coarse}	-0.09 (-0.1, -0.08)	-0.5 (-0.75, -0.24)	0.13 (-0.09, 0.36)	-0.04 (-0.35, 0.28)	0 (0, 0.01)	-0.16 (-0.39, 0.07)	0.1 (-0.15, 0.35)	-0.2 (-0.48, 0.08)	0.01 (-0.01, 0.04)	-0.09 (-0.31, 0.13)	-0.35 (-0.75, 0.04)	0.05 (-0.23, 0.34)	0.07 (-0.03, 0.18)
	$PM_{2.5abs}$	-0.09	-0.5	0.13	-0.04	0 (0, 0.01)	-0.16	0.1	-0.21	0.01	-0.09	-0.35	0.05	0.07

		(-0.1, -0.08)	(-0.75, -0.24)	(-0.09, 0.36)	(-0.35, 0.27)		(-0.39, 0.07)	(-0.15, 0.35)	(-0.49, 0.07)	(-0.01, 0.04)	(-0.31, 0.13)	(-0.75, 0.04)	(-0.23, 0.33)	(-0.03, 0.18)
	NO_2	-0.09 (-0.1, -0.08)	-0.49 (-0.75, -0.24)	0.14 (-0.09, 0.36)	-0.03 (-0.34, 0.28)	0 (0, 0.01)	-0.16 (-0.39, 0.07)	0.1 (-0.16, 0.35)	-0.21 (-0.49, 0.07)	0.01 (-0.01, 0.04)	-0.08 (-0.3, 0.14)	-0.36 (-0.75, 0.04)	0.06 (-0.22, 0.34)	0.07 (-0.03, 0.18)
	NO_x	-0.09 (-0.1, -0.08)	-0.5 (-0.76, -0.25)	0.13 (-0.09, 0.36)	-0.04 (-0.35, 0.28)	0 (0, 0.01)	-0.16 (-0.39, 0.07)	0.09 (-0.16, 0.34)	-0.22 (-0.49, 0.06)	0.01 (-0.01, 0.04)	-0.09 (-0.31, 0.13)	-0.35 (-0.75, 0.04)	0.05 (-0.23, 0.33)	0.08 (-0.03, 0.18)
	O_3	-0.09 (-0.1, -0.08)	-0.5 (-0.75, -0.24)	0.13 (-0.1, 0.35)	-0.05 (-0.37, 0.26)	0 (0, 0.01)	-0.16 (-0.38, 0.07)	0.09 (-0.16, 0.34)	-0.21 (-0.49, 0.07)	0.01 (-0.01, 0.04)	-0.09 (-0.31, 0.13)	-0.36 (-0.75, 0.04)	0.05 (-0.24, 0.33)	0.07 (-0.03, 0.18)
DNAmTL	$PM_{2.5}$	-0.21 (-0.22, -0.21)	1.27 (1.1, 1.45)	-0.4 (-0.56, -0.25)	-0.98 (-1.19, -0.76)	0 (-0.01, 0)	0.21 (0.05, 0.36)	0.2 (0.03, 0.37)	0.4 (0.21, 0.59)	-0.02 (-0.03, 0)	-0.06 (-0.22, 0.09)	-0.18 (-0.45, 0.09)	0.1 (-0.1, 0.29)	0.07 (0, 0.14)
	PM_{10}	-0.21 (-0.22, -0.21)	1.27 (1.1, 1.45)	-0.4 (-0.56, -0.25)	-0.98 (-1.19, -0.76)	0 (-0.01, 0)	0.21 (0.05, 0.36)	0.21 (0.03, 0.38)	0.4 (0.21, 0.59)	-0.02 (-0.03, 0)	-0.06 (-0.21, 0.09)	-0.18 (-0.45, 0.09)	0.1 (-0.1, 0.29)	0.07 (0, 0.14)
	PNC	-0.21 (-0.22, -0.21)	1.27 (1.09, 1.45)	-0.4 (-0.56, -0.25)	-0.98 (-1.19, -0.76)	0 (-0.01, 0)	0.21 (0.05, 0.36)	0.21 (0.03, 0.38)	0.4 (0.21, 0.59)	-0.02 (-0.03, 0)	-0.06 (-0.22, 0.09)	-0.18 (-0.45, 0.09)	0.09 (-0.1, 0.29)	0.07 (0, 0.14)
	PM_{coarse}	-0.21 (-0.22, -0.21)	1.28 (1.11, 1.46)	-0.39 (-0.55, -0.24)	-0.96 (-1.17, -0.74)	0 (-0.01, 0)	0.2 (0.05, 0.36)	0.22 (0.04, 0.39)	0.43 (0.24, 0.62)	-0.02 (-0.03, 0)	-0.06 (-0.21, 0.09)	-0.17 (-0.44, 0.1)	0.1 (-0.09, 0.29)	0.07 (0, 0.14)
	$PM_{2.5abs}$	-0.21 (-0.22, -0.21)	1.28 (1.1, 1.45)	-0.4 (-0.55, -0.25)	-0.97 (-1.18, -0.75)	0 (-0.01, 0)	0.21 (0.05, 0.36)	0.21 (0.04, 0.39)	0.41 (0.22, 0.6)	-0.02 (-0.03, 0)	-0.06 (-0.21, 0.09)	-0.17 (-0.44, 0.1)	0.09 (-0.11, 0.28)	0.07 (0, 0.14)
	NO_2	-0.21 (-0.22, -0.21)	1.26 (1.09, 1.44)	-0.4 (-0.55, -0.25)	-0.97 (-1.19, -0.76)	0 (-0.01, 0)	0.2 (0.05, 0.36)	0.2 (0.03, 0.37)	0.41 (0.21, 0.6)	-0.02 (-0.03, 0)	-0.06 (-0.21, 0.09)	-0.18 (-0.45, 0.09)	0.1 (-0.09, 0.3)	0.07 (0, 0.14)
	NO_x	-0.21 (-0.22, -0.21)	1.27 (1.09, 1.45)	-0.41 (-0.56, -0.26)	-0.99 (-1.2, -0.77)	0 (-0.01, 0)	0.21 (0.05, 0.36)	0.2 (0.02, 0.37)	0.4 (0.21, 0.59)	-0.02 (-0.03, 0)	-0.06 (-0.21, 0.09)	-0.18 (-0.45, 0.09)	0.1 (-0.1, 0.29)	0.07 (0, 0.14)
	O_3	-0.21 (-0.22, -0.21)	1.28 (1.1, 1.45)	-0.42 (-0.57, -0.26)	-1 (-1.22, -0.79)	0 (-0.01, 0)	0.21 (0.05, 0.37)	0.2 (0.02, 0.37)	0.4 (0.21, 0.59)	-0.02 (-0.03, 0)	-0.06 (-0.21, 0.09)	-0.18 (-0.45, 0.09)	0.09 (-0.1, 0.28)	0.07 (0, 0.14)

Covariate-adjusted linear mixed-effect regression models were used. Results were from our full model adjusted by age, sex, an indicator of each study waves visit (KORA S4, KORA F4, or KORA FF4), houseman estimated cells, batch, chip, educational attainment, smoking status, alcohol consumption, physical activity, BMI, hypertension, diabetes, HDL, and LDL.

Table S3. Descriptive statistics of participant characteristics by study wave in never smokers (N=1726).

	S4 (N=632)	**F4 (N=673)**	**FF4 (N=421)**	
Variable	**Mean ± SD / N (%)**	**Mean ± SD / N (%)**	**Mean ± SD / N (%)**	***p*–value**
Age (years)	55.0 ± 9.0	61.9 ± 9.0	65.7 ± 8.3	< 0.001
Sex (male)	220 (34.8)	232 (34.5)	146 (34.7)	1
Education				0.48
Primary school	373 (59.0)	392 (58.2)	228 (54.2)	
High school	145 (23.0)	150 (22.3)	100 (23.8)	
College	114 (18.0)	131 (19.5)	93 (22.1)	
BMI (kg/m^2)	27.7 ± 4.4	28.0 ± 4.5	28.0 ± 4.7	0.52
Alcohol consumption (g/day)	12.9 ± 18.0	12.0 ± 17.2	13.2 ± 17.7	0.10
Physical activity				0.001
Low	208 (32.9)	207 (30.8)	110 (26.1)	
Medium	311 (49.2)	294 (43.7)	199 (47.3)	
High	113 (17.9)	172 (25.5)	112 (26.6)	
Hypertension (yes)	262 (41.5)	306 (45.5)	203 (48.2)	0.08
Diabetes (yes)	17 (2.7)	49 (7.3)	47 (11.2)	< 0.001
HDL cholesterol (mmol/l)	1.6 ± 0.4	1.5 ± 0.4	1.8 ± 0.5	< 0.001
LDL cholesterol (mmol/l)	3.7 ± 1.1	3.6 ± 0.9	3.6 ± 0.9	0.24
DNAmHorvathAge*	56.1 ± 8.3 (58.1)	60.0 ± 7.5 (38.2)	65.3 ± 6.2 (48.9)	< 0.001
DNAmHannumAge*	58.9 ± 9.1 (80.1)	71.0 ± 9.6 (96.1)	54.2 ± 7.6 (1.7)	< 0.001
DNAmPhenoAge*	49.8 ± 10.9 (22.6)	54.6 ± 12.4 (19.0)	52.9 ± 9.8 (2.1)	< 0.001
DNAmGrimAge*	53.8 ± 8.1 (39.7)	60.6 ± 8.2 (35.4)	62.7 ± 7.5 (14.7)	< 0.001
DNAmSkinBloodAge*	55.2 ± 9.2 (52.9)	62.8 ± 9.2 (58.1)	60.2 ± 7.7 (6.9)	< 0.001
DNAmTL	7.1 ± 0.3	6.8 ± 0.3	6.9 ± 0.22	< 0.001
CD8 T cells	0.11 ± 0.05	0.1 ± 0.06	0.05 ± 0.04	< 0.001
CD4 T cells	0.22 ± 0.05	0.23 ± 0.05	0.18 ± 0.06	< 0.001
Natural killer cells	0.03 ± 0.03	0.06 ± 0.03	0.07 ± 0.04	< 0.001
B cells	0.06 ± 0.02	0.07 ± 0.02	0.06 ± 0.04	< 0.001
Monocytes	0.1 ± 0.02	0.1 ± 0.02	0.07 ± 0.02	< 0.001

KORA = Cooperative Health Research in the Region of Augsburg, S4 = fourth cross-sectional health survey of the KORA cohort, F4 = first follow-up examination of KORA S4, FF4 = second follow-up examination of KORA S4. BMI = body mass index, HDL = high density lipoprotein, LDL= low density lipoprotein. Physical activity was defined according to the exercise time per week: Low = almost or no sporting activity, Medium = regular/ irregular approx. 1 hour per week, High = regularly more than 2 hours in the week. 347 participants attended two examinations, and 344 attended three examinations (Only one participant attended one visit). * Represents the mean ± SD and the percent of participants with positive age acceleration. *p*-value was based on the Kruskal-Wallis test for continuous variables, and Pearson's Chi-squared test for categorical variables.

Table S4. Absolute/percent changes (95% CIs) in epigenetic aging biomarkers per IQR increase in air pollutants in never smokers.

Epigenetic aging biomarkers	**Exposure**	**Basic**	**Behavioral**	**Clinical**	**Full**
Acceleration of DNAmHorvathAge	PM_{10}	-0.35 (-0.58, -0.11); p = 0.04	-0.36(-0.6, -0.12); p = 0.03	-0.36 (-0.6, -0.12); p = 0.03	-0.36 (-0.6, -0.12); p = 0.03
	PNC	-0.33 (-0.52, -0.14); p = 0.01	-0.34 (-0.53, -0.15); p = 0.007	-0.35 (-0.54, -0.16); p = 0.007	-0.35 (-0.54, -0.16); p = 0.006
	NO_x	-0.33 (-0.53, -0.14); p = 0.01	-0.34 (-0.54, -0.14); p = 0.009	-0.35 (-0.55, -0.16); p = 0.009	-0.35 (-0.55, -0.16); p = 0.007
Acceleration of DNAmSkinBloodAge	$PM_{2.5}$	-0.42 (-0.63, -0.21); p = 0.004	-0.42 (-0.63, -0.21); p = 0.004	-0.43 (-0.64, -0.22); p = 0.002	-0.43 (-0.63, -0.22); p = 0.003

Covariate-adjusted linear mixed-effect regression models were used. The effect estimates of acceleration of DNAmGrimAge and DNAmTL were presented as absolute change (together 95% CI) and percent change, respectively. Basic model: adjusted for age, sex, an indicator of each study wave (KORA S4, KORA F4, or KORA FF4), houseman estimated cells, batch, and chip; Behavioral model: further adjusted for educational attainment, alcohol consumption, physical activity into basic model; Clinical model: additionally included BMI, hypertension, diabetes, HDL, and LDL into the basic models. Full model: all confounders used in other three models. $PM_{2.5}$ = particulate matter with an aerodynamic diameter less than or equal to 2.5 μm; PNC = particle number concentration; NO_x = nitrogen oxide. An IQR increase was 1.5 $\mu g/m^3$ for $PM_{2.5}$, 2.1$\mu g/m^3$ for $PM_{10,}$ $2.0\times10^3/cm^3$ for PNC, and 8.8 $\mu g/m^3$ for NO_x.

Table S5. Descriptive statistics of participant characteristics by study wave in current and former smokers (N=2378).

	S4 (N=848)	**F4 (N=926)**	**FF4 (N=604)**	
Variable	**Mean ± SD / N (%)**	**Mean ± SD / N (%)**	**Mean ± SD / N (%)**	***p*–value**
Age (years)	53.1 ± 8.7	59.9 ± 8.7	64.0 ± 7.8	< 0.001
Sex (male)	516 (60.9)	551 (59.5)	349 (57.8)	0.50
Education				0.79
Primary school	500 (59.0)	527 (56.9)	339 (56.1)	
High school	192 (22.6)	218 (23.5)	140 (23.2)	
College	156 (18.4)	181 (19.6)	125 (20.7)	
BMI (kg/m^2)	27.8 ± 4.6	28.3 ± 5.0	28.4 ± 5.2	0.10
Alcohol consumption (g/day)	20.2 ± 24.2	18.5 ± 22.5	17.3 ± 22.2	0.03
Physical activity				0.003
Low	275 (32.4)	285 (30.8)	169 (28.0)	
Medium	396 (46.7)	403 (43.5)	254 (42.0)	
High	177 (20.9)	238 (25.7)	181 (30.0)	
Hypertension (yes)	358 (42.2)	413 (44.6)	280 (46.4)	0.27
Diabetes (yes)	34 (4.0)	93 (10.0)	83 (13.7)	< 0.001
HDL cholesterol (mmol/l)	1.4 ± 0.4	1.4 ± 0.4	1.7 ± 0.5	< 0.001
LDL cholesterol (mmol/l)	3.7 ± 1.0	3.6 ± 0.9	3.5 ± 0.9	0.027
DNAmHorvathAge*	54.6 ± 8.3 (62.7)	58.4 ± 7.5 (38.0)	64.8 ± 6.0 (57.5)	< 0.001
DNAmAgeHannum*	58.1 ± 9.4 (86.2)	69.9 ± 10.2 (96.8)	53.5 ± 7.7 (0.8)	< 0.001
DNAmPhenoAge*	49.2 ± 11.0 (26.8)	53.6 ± 12.5 (22.7)	52.4 ± 9.6 (4.0)	< 0.001
DNAmGrimAge*	56.5 ± 8.7 (69.8)	62.9 ± 9.1 (65.4)	64.9 ± 8.0 (52.2)	< 0.001
DNAmSkinBloodAge*	53.5 ± 9.3 (53.8)	61.1 ± 9.2 (61.8)	58.9 ± 7.7 (7.8)	< 0.001
DNAmTL	7.0 ± 0.3	6.7 ± 0.3	6.8 ± 0.2	< 0.001
CD8 T cells	0.11 ± 0.05	0.1 ± 0.05	0.05 ± 0.04	< 0.001
CD4 T cells	0.22 ± 0.05	0.24 ± 0.05	0.19 ± 0.06	< 0.001
Natural killer cells	0.03 ± 0.02	0.06 ± 0.02	0.07 ± 0.04	< 0.001
B cells	0.06 ± 0.02	0.07 ± 0.02	0.05 ± 0.03	< 0.001
Monocytes	0.11 ± 0.02	0.1 ± 0.02	0.07 ± 0.02	< 0.001

KORA = Cooperative Health Research in the Region of Augsburg; S4 = fourth cross-sectional health survey of the KORA cohort; F4 = first follow-up examination of KORA S4; FF4 = second follow-up examination of KORA S4. BMI = body mass index, HDL = high density lipoprotein, LDL= low density lipoprotein. Physical activity was defined according to the exercise time per week: Low = almost or no sporting activity, Medium = regular/ irregular approx. 1 hour per week, High = regularly more than 2 hours in the week. 487 participants attended two examinations, and 468 attended three examinations. * Represents the mean ± SD and the percent of participants with positive age acceleration. *p*-value was based on the Kruskal-Wallis test for continuous variables, and Pearson's Chi-squared test for categorical variables.

Table S6. Absolute/percent changes (95% CIs) in epigenetic aging biomarkers per IQR increase in air pollutants in current and former smokers.

Epigenetic aging biomarkers	Exposure	Basic	Behavioral	Clinical	Full
Acceleration of DNAmHorvathAge	$PM_{2.5}$	0.33 (0.13, 0.52); $p = 0.01$	0.31 (0.12,0.51); $p = 0.03$	0.32 (0.12,0.51); $p = 0.03$	0.32 (0.12,0.51); $p = 0.01$
	PM_{10}	0.29 (0.09, 0.49); $p = 0.02$	0.28 (0.08,0.48); $p = 0.03$	0.33 (0.04,0.61); $p = 0.03$	0.26 (0.06,0.46); $p = 0.03$
	PNC	0.24 (0.09, 0.4); $p = 0.01$	0.23 (0.08,0.39); $p = 0.03$	0.23 (0.07,0.39); $p = 0.04$	0.23 (0.07,0.39); $p = 0.02$
	PM_{coarse}	0.31 (0.12, 0.51); $p = 0.01$	0.31 (0.11,0.5); $p = 0.03$	0.32 (0.09,0.52); $p = 0.03$	0.31 (0.12,0.51); $p = 0.01$
	$PM_{2.5abs}$	0.34 (0.12, 0.55); $p = 0.01$	0.31 (0.1,0.52); $p = 0.03$	0.33 (0.09,0.52); $p = 0.04$	0.31 (0.09,0.52); $p = 0.03$
	NO_2	0.36 (0.14, 0.59); $p = 0.01$	0.35 (0.13,0.58); $p = 0.03$	0.37 (0.14,0.59); $p = 0.03$	0.36 (0.14,0.59); $p = 0.01$
	NO_x	0.23 (0.06, 0.39); $p = 0.02$	0.22 (0.05,0.39); $p = 0.04$	0.22 (0.05,0.39); $p = 0.03$	0.22 (0.05,0.39); $p = 0.03$
Acceleration of DNAmPhneoAge	$PM_{2.5}$	0.51 (0.13, 0.89); $p = 0.02$	0.5 (0.12, 0.88); $p = 0.03$	0.51 (0.13, 0.89); $p = 0.02$	0.5 (0.12, 0.88); $p = 0.03$
	PM_{coarse}	0.55 (0.17, 0.94); $p = 0.02$	0.55 (0.17, 0.96); $p = 0.03$	0.55 (0.17, 0.94); $p = 0.02$	0.58 (0.19, 0.96); $p = 0.02$
	$PM_{2.5abs}$	0.56 (0.13, 0.96); $p = 0.02$	0.55 (0.14, 0.97); $p = 0.03$	0.6 (0.19, 1.02); $p = 0.02$	---
	NO_2	0.65 (0.21, 1.09); $p = 0.02$	0.64 (0.19, 1.08); $p = 0.03$	0.65 (0.21, 1.09); $p = 0.02$	0.61 (0.17, 1.06); $p = 0.03$
DNAmTL	$PM_{2.5}$	-0.3 (-0.4, -0.2); $p = 7.3\times10^{-7}$	-0.3 (-0.4, -0.2); $p = 1.3\times10^{-6}$	-0.3 (-0.4, -0.2); $p = 7.4\times10^{-7}$	-0.3 (-0.4, -0.2); $p = 1.6\times10^{-5}$
	PM_{10}	-0.2 (-0.3, -0.1); $p = 0.0003$	-0.2 (-0.3, -0.1); $p = 0.0008$	-0.2 (-0.3, -0.1); $p = 0.0003$	-0.2 (-0.3, -0.1); $p = 0.005$
	PNC	-0.1 (-0.2, 0); $p = 0.02$	-0.1 (-0.2, 0); $p = 0.04$	-0.1 (-0.2, 0); $p = 0.02$	---
	PM_{coarse}	-0.3 (-0.4, -0.1); $p = 2.3\times10^{-5}$	-0.3 (-0.4, -0.2); $p = 1.8\times10^{-5}$	-0.3 (-0.4, -0.1); $p = 1.6\times10^{-5}$	-0.3 (-0.4, -0.1); $p = 0.0001$
	$PM_{2.5abs}$	-0.3 (-0.4, -0.2); $p = 5.9\times10^{-6}$	-0.3 (-0.4, -0.2); $p = 1.8\times10^{-6}$	-0.3 (-0.4, -0.2); $p = 8.9\times10^{-6}$	-0.3 (-0.4, -0.2); $p = 0.0001$
	NO_2	-0.3 (-0.4, -0.2); $p = 4.2\times10^{-5}$	-0.3 (-0.4, -0.1); $p = 0.0001$	-0.3 (-0.4, -0.2); $p = 4.8\times10^{-5}$	-0.3 (-0.4, -0.1); $p = 0.001$

Covariate-adjusted linear mixed-effect regression models were used. The effect estimates of acceleration of DNAmGrimAge and DNAmTL were presented as absolute change (together 95% CI) and percent change, respectively. Basic model: adjusted for age, sex, an indicator of each study wave (KORA S4, KORA F4, or KORA FF4), houseman estimated cells, batch, and chip; Behavioral model: further adjusted for educational attainment, alcohol consumption, physical activity into basic model; Clinical model: additionally included BMI, hypertension, diabetes, HDL, and LDL into the basic models. Full model: all confounders used in other three models. $PM_{2.5}$ = particulate matter with an aerodynamic

diameter less than or equal to 2.5 μm; PM_{10} = particulate matter with an aerodynamic diameter less than or equal to 10 μm; PNC = particle number concentration; PM_{coarse} = particulate matter with an aerodynamic diameter of 2.5-10 μm; $PM_{2.5abs}$ = $PM_{2.5}$ absorbance; NO_2 = nitrogen dioxide. An IQR increase was 1.3 $\mu g/m^3$ for $PM_{2.5}$, 2.1 $\mu g/m^3$ for PM_{10}, $2.0\times10^3/cm^3$ for PNC, 1.3 $\mu g/m^3$ for PM_{coarse}, $0.3\times10^{-5}/m$ for $PM_{2.5abs}$, and 6.9 $\mu g/m^3$ for NO_2.

Table S7. Pathways identified by exposure-related CpG sites via Ingenuity Pathway Analysis (IPA)**.**

Canonical pathways		Top diseases		Top bio functions	
Name	*p*-value	Name	*p*-value range	Name	*p*-value range
Triacylglycerol biosynthesis	0.01	Cardiovascular Disease	$0.02 - 2\times10^{-4}$	Lipid Metabolism	$0.048 - 9\times10^{-6}$
Stearate biosynthesis I	0.01	Hematological Disease	$0.002 - 2\times10^{-4}$	Small Molecule Biochemistry	$0.048 - 9\times10^{-6}$
D-myo-inositol (1,4,5,6)-tetrakisphosphate biosynthesis	0.036	Hereditary Disorder	$0.002 - 2\times10^{-4}$	Molecular Transport	$0.035 - 4\times10^{-4}$
D-myo-inositol (3,4,5,6)-tetrakisphosphate biosynthesis	0.036	Metabolic Disease	$0.002 - 2\times10^{-4}$	Cell-To-Cell Signaling and Interaction	$0.02 - 6\times10^{-4}$
3-phosphoinositide degradation	0.038	Ophthalmic Disease	$0.001 - 2\times10^{-4}$	Energy Production	$0.01 - 6\times10^{-4}$

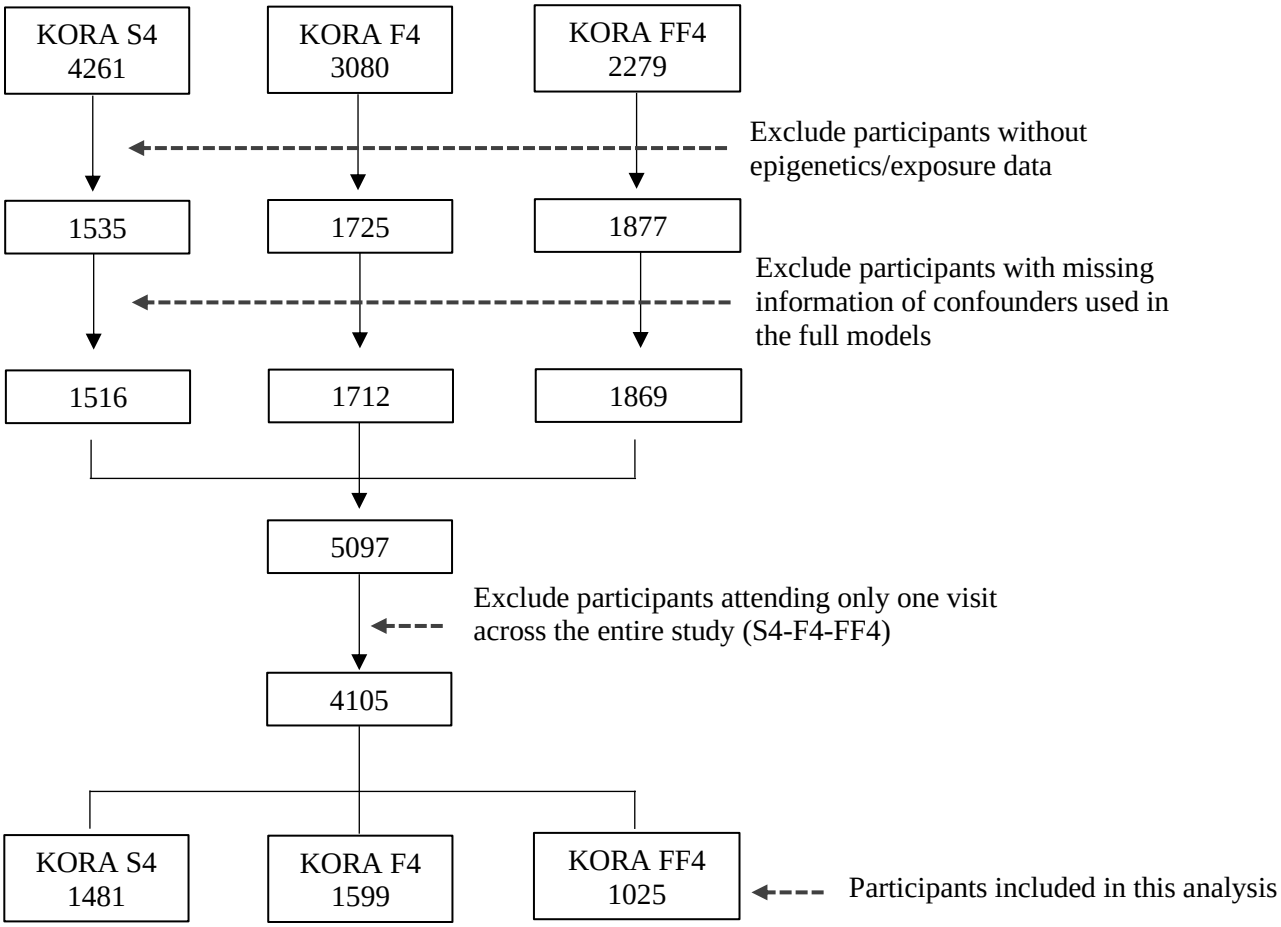


Figure S1. Flow chart of participants exclusion process in this study

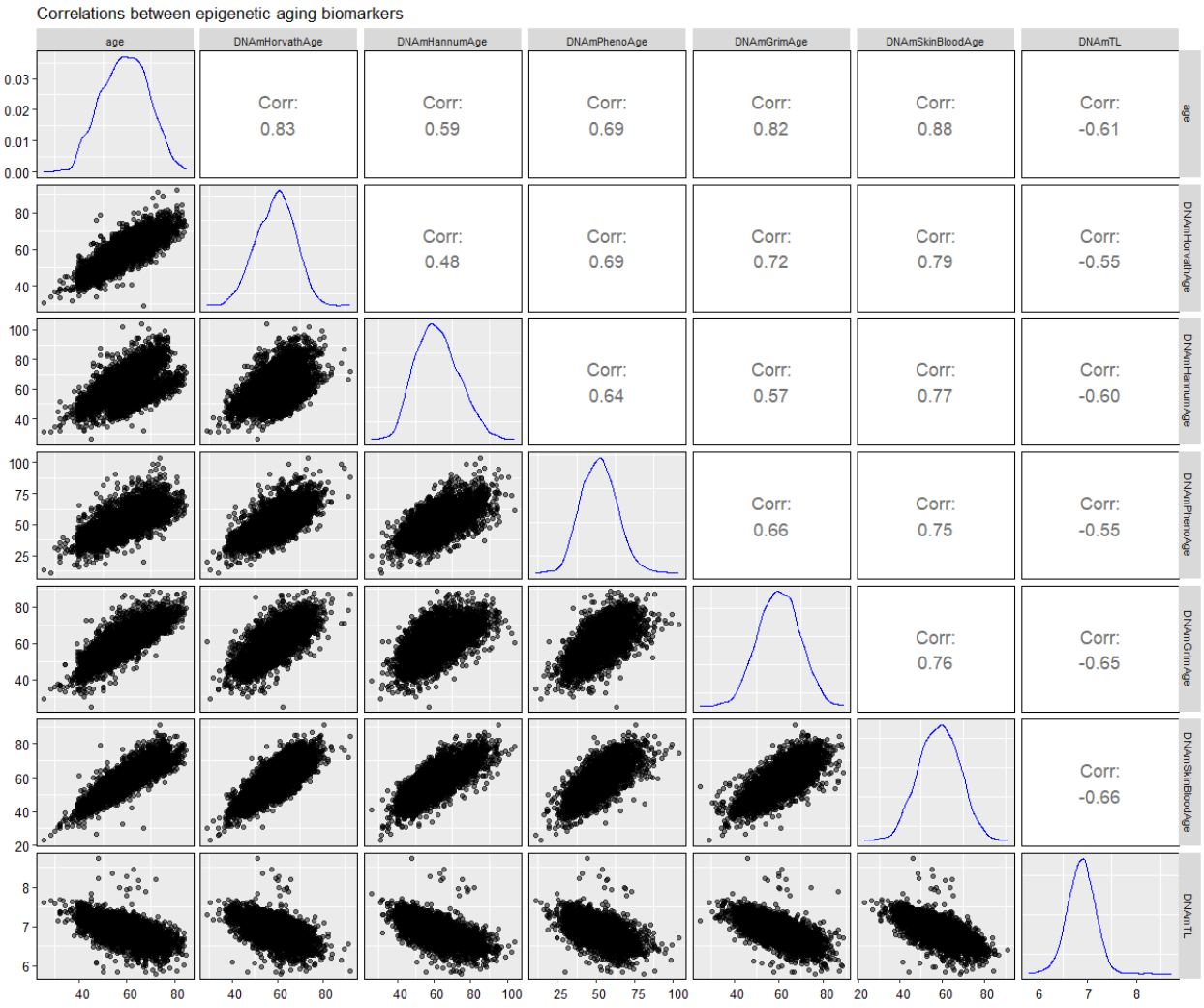


Figure S2. Pearson correlations between Age, DNAmHorvathAge, DNAmHannumAge, DNAmPhenoAge, DNAmGrimAge, DNAmSkinBloodAge, and DNAmTL. Pearson's correlation coefficients were calculated to determine the correlations between chronological age and epigenetic aging biomarkers.

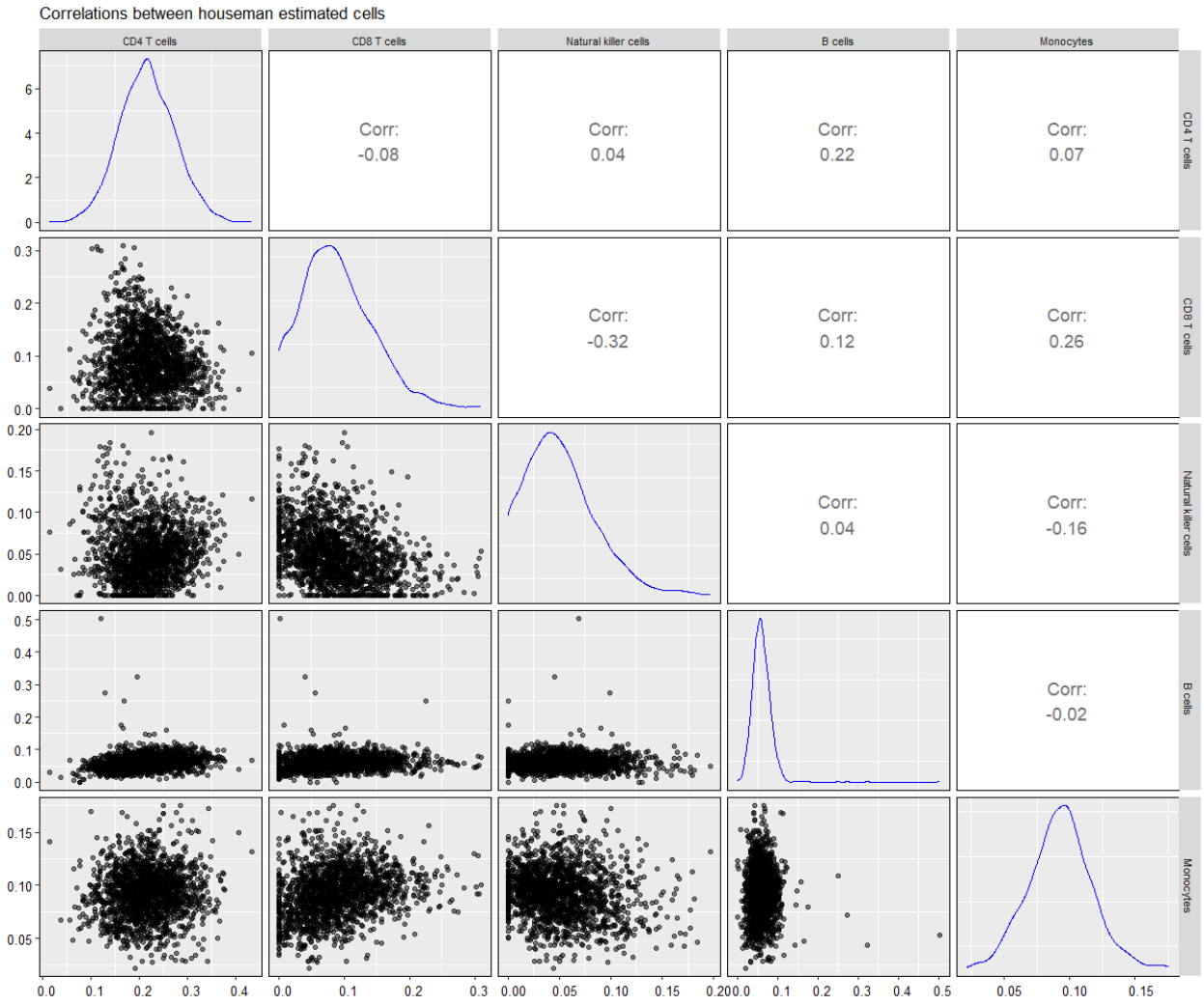


Figure S3. Pearson correlation between houseman estimated cells.

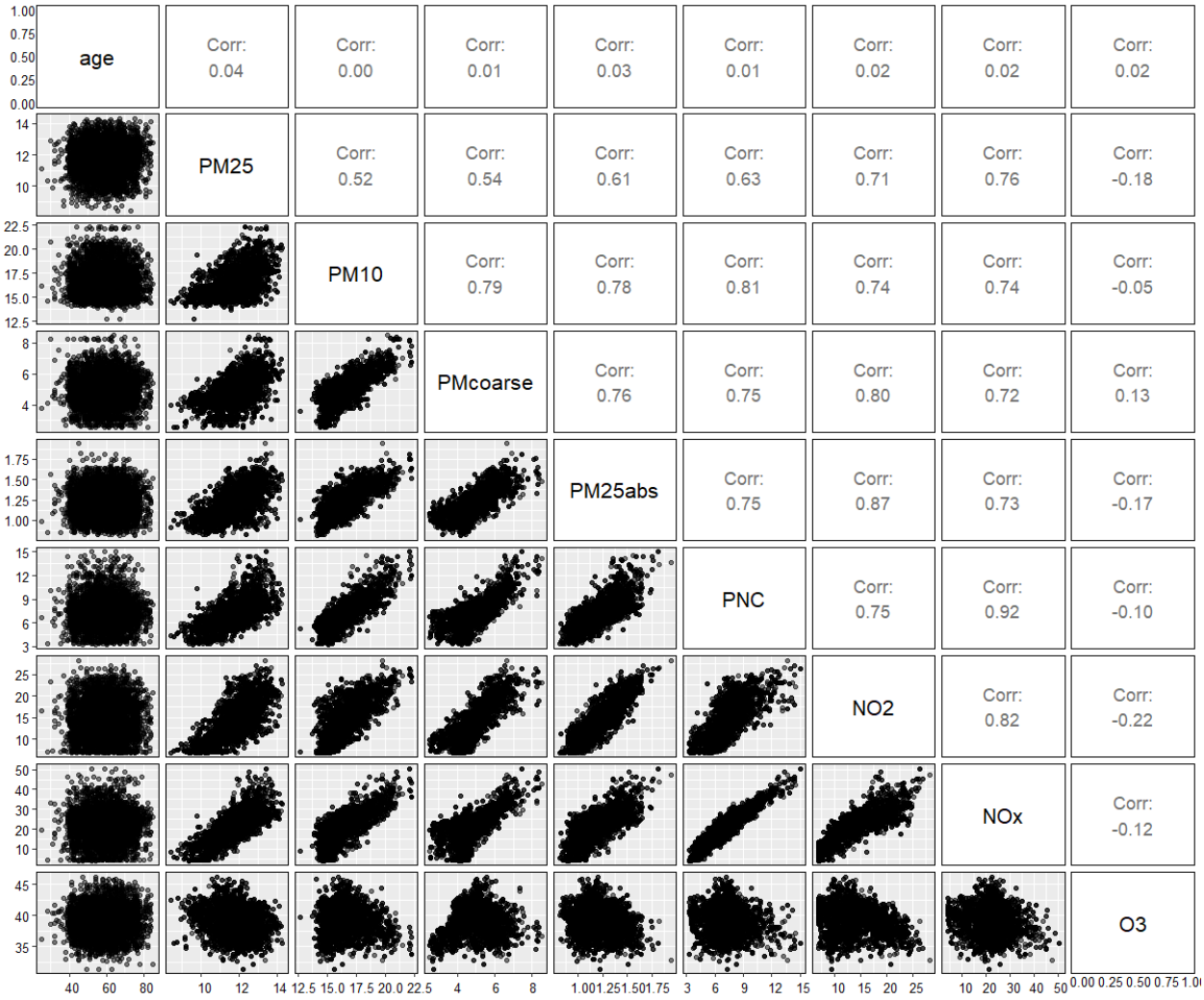


Figure S4. Pearson correlations between chronological age, $PM_{2.5}$, PM_{10}, PM_{coarse}, $PM_{2.5abs}$, PNC, NO_2, NO_x and O_3.

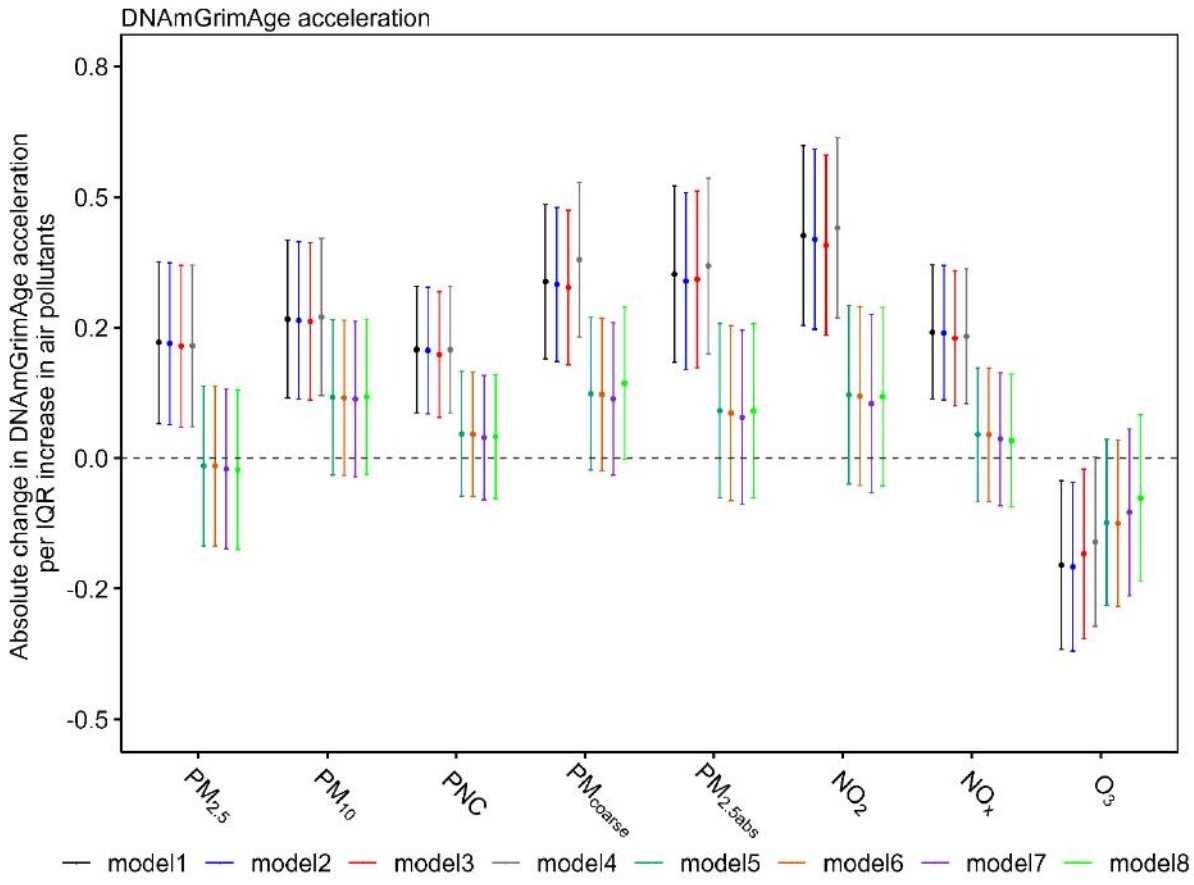


Figure S5. Results of analysis on DNAmGrimAge with an inclusion of confounders step by step.

Covariate-adjusted linear mixed-effect regression models were used.

Random effect: participant ID, batch, chip

model1: age + sex + houseman cells + indicator of study + BMI + hypertension + diabetes + HDL + LDL

model2: model1 + alcohol consumption

model3: model1 + physical activity

model4: model1 + education

model5: model1 + smoking

model6: model1 + smoking + alcohol consumption

model7: model1 + smoking + alcohol consumption + physical activity

model8: model1 + smoking + alcohol consumption + physical activity + education

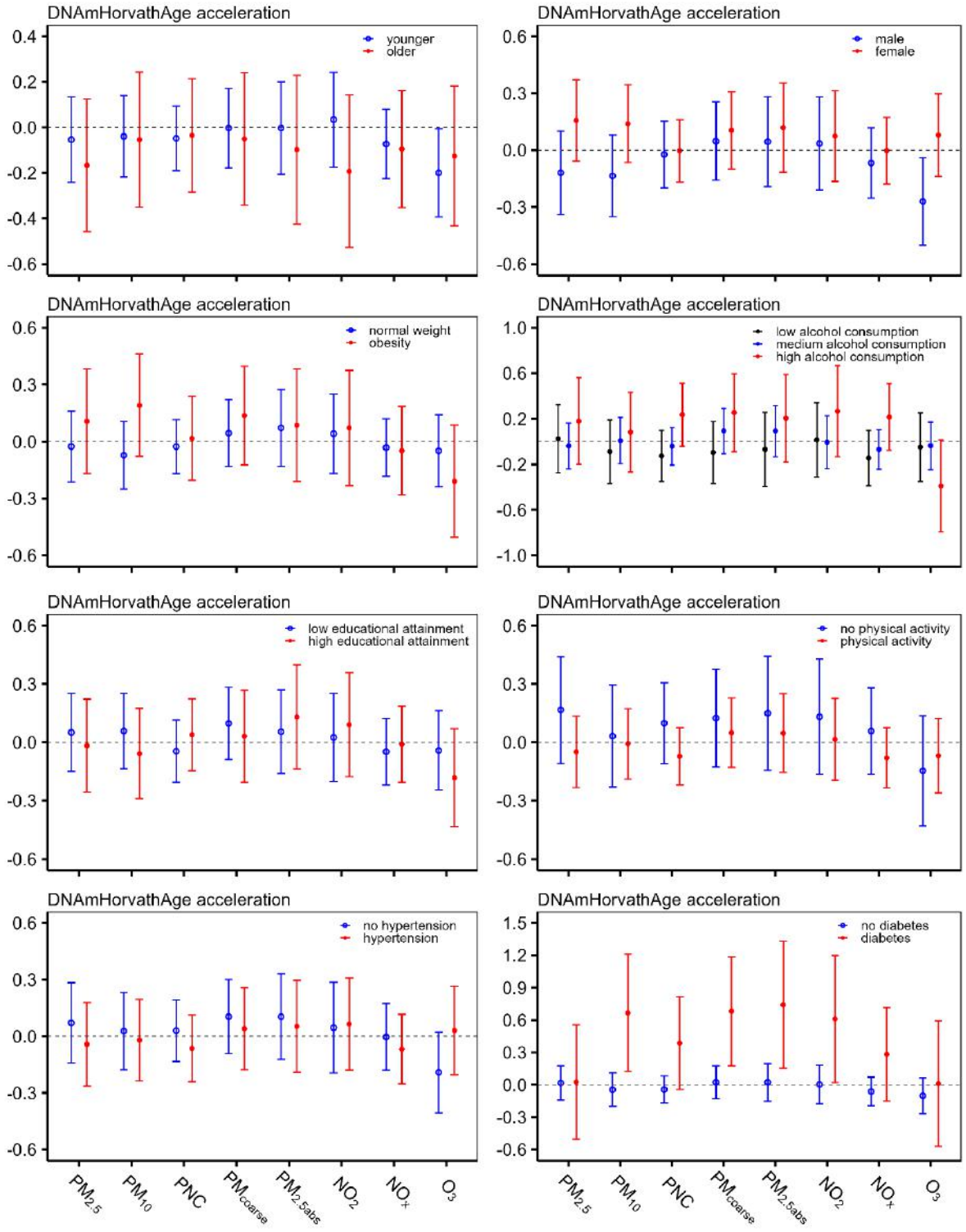


Figure S6. Absolute changes (95% CI, years) in DNAmHorvathAge acceleration per IQR increase in air pollutant concentrations stratified by categorized age group (≤ 65 vs > 65), sex, categorized BMI group, categorized group of alcohol consumption, educational attainment, physical activity, and history of diseases (hypertension, diabetes).

Covariate-adjusted linear mixed-effect regression models were used. Results were from our full model adjusted by age, sex, an indicator of each study waves visit (KORA S4, KORA F4, or KORA FF4), houseman estimated cells, batch, chip, educational attainment, smoking status, alcohol consumption, physical activity, BMI, hypertension, diabetes, HDL, and LDL.

$PM_{2.5}$ = particulate matter with an aerodynamic diameter less than or equal to 2.5 μm; PM_{10} = particulate matter with an aerodynamic diameter less than or equal to 10 μm; PNC = particle number concentration; PM_{coarse} = particulate matter with an aerodynamic diameter of 2.5-10 μm; $PM_{2.5abs}$ = $PM_{2.5}$ absorbance; NO_2 = nitrogen dioxide; NO_x = nitrogen oxide; O_3 = ozone. An IQR increase was 1.4 $\mu g/m^3$ for $PM_{2.5}$, 2.1 $\mu g/m^3$ for PM_{10}, $2.0\times10^3/cm^3$ for PNC, 1.4 $\mu g/m^3$ for PM_{coarse}, $0.3\times10^{-5}/m$ for $PM_{2.5abs}$, 7.2 $\mu g/m^3$ for NO_2, 8.8 $\mu g/m^3$ for NO_x, and 3.4 $\mu g/m^3$ for O_3.

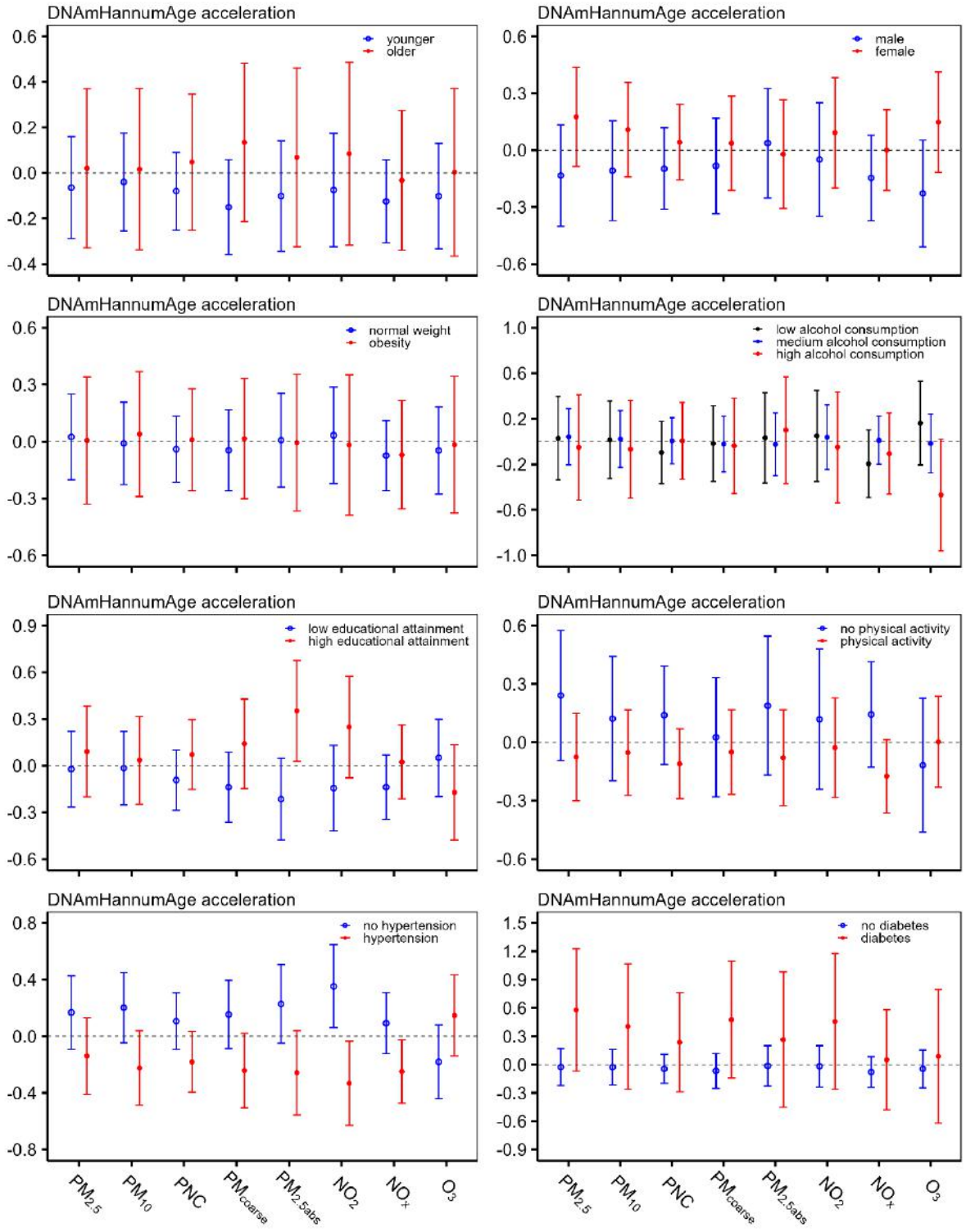


Figure S7. Absolute changes (95% CI, years) in DNAmHannumAge acceleration per IQR increase in air pollutant concentrations stratified by categorized age group (≤ 65 vs > 65), sex, categorized BMI group, categorized group of alcohol consumption, educational attainment, physical activity, and history of diseases (hypertension, diabetes).

Covariate-adjusted linear mixed-effect regression models were used. Results were from our full model adjusted by age, sex, an indicator of each study waves visit (KORA S4, KORA F4, or KORA FF4), houseman estimated cells, batch, chip, educational attainment, smoking status, alcohol consumption, physical activity, BMI, hypertension, diabetes, HDL, and LDL.

$PM_{2.5}$ = particulate matter with an aerodynamic diameter less than or equal to 2.5 μm; PM_{10} = particulate matter with an aerodynamic diameter less than or equal to 10 μm; PNC = particle number concentration; PM_{coarse} = particulate matter with an aerodynamic diameter of 2.5-10 μm; $PM_{2.5abs}$ = $PM_{2.5}$ absorbance; NO_2 = nitrogen dioxide; NO_x = nitrogen oxide; O_3 = ozone. An IQR increase was 1.4 $\mu g/m^3$ for $PM_{2.5}$, 2.1 $\mu g/m^3$ for PM_{10}, $2.0\times10^3/cm^3$ for PNC, 1.4 $\mu g/m^3$ for PM_{coarse}, $0.3\times10^{-5}/m$ for $PM_{2.5abs}$, 7.2 $\mu g/m^3$ for NO_2, 8.8 $\mu g/m^3$ for NO_x, and 3.4 $\mu g/m^3$ for O_3.

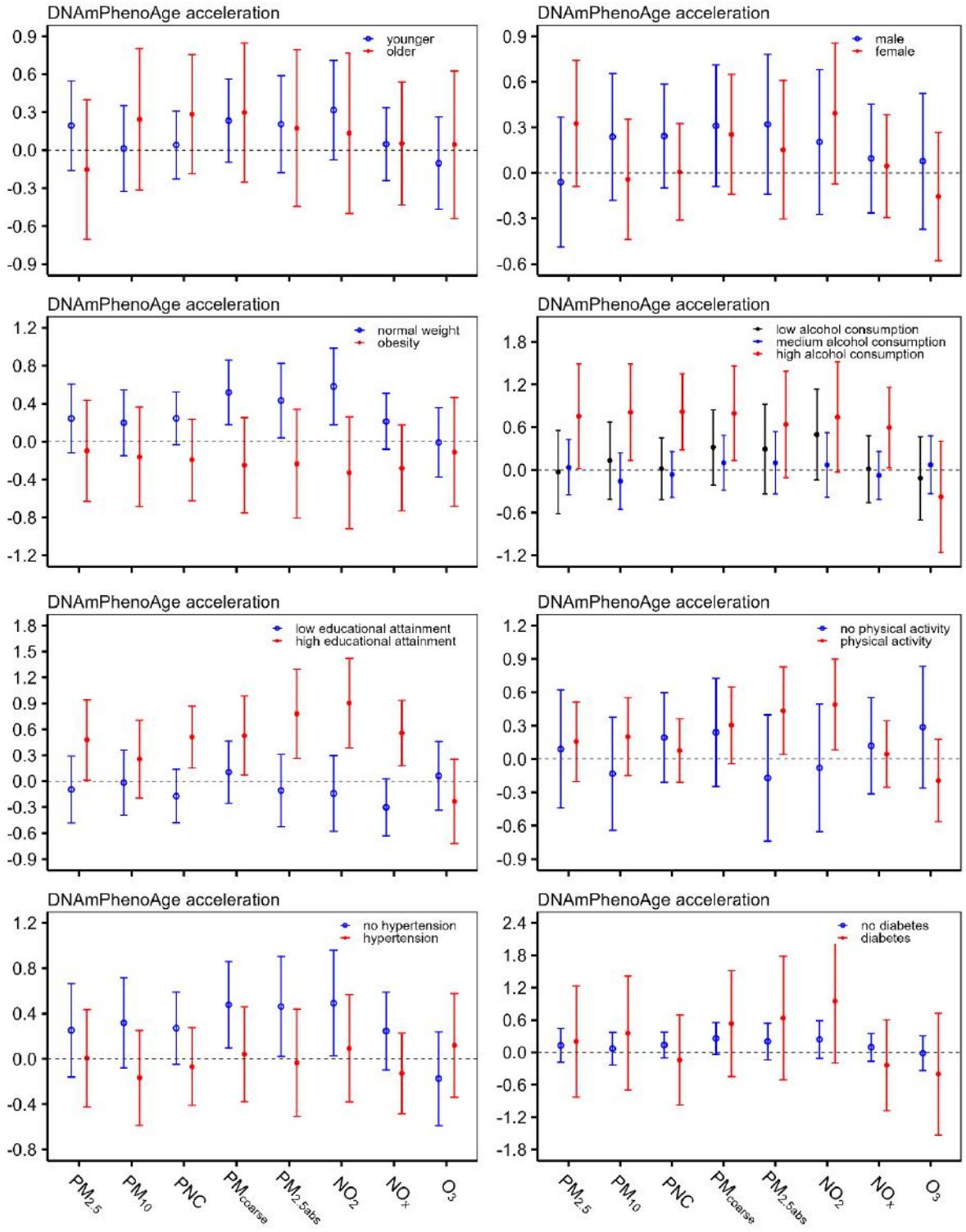


Figure S8. Absolute changes (95% CI, years) in DNAmPhenoAge acceleration per IQR increase in air pollutant concentrations stratified by categorized age group (≤ 65 vs > 65), sex, categorized BMI group, categorized group of alcohol consumption, educational attainment, physical activity, and history of diseases (hypertension, diabetes).

Covariate-adjusted linear mixed-effect regression models were used. Results were from our full model adjusted by age, sex, an indicator of each study waves visit (KORA S4, KORA F4, or KORA FF4), houseman estimated cells, batch, chip, educational attainment, smoking status, alcohol consumption, physical activity, BMI, hypertension, diabetes, HDL, and LDL.

$PM_{2.5}$ = particulate matter with an aerodynamic diameter less than or equal to 2.5 μm; PM_{10} = particulate matter with an aerodynamic diameter less than or equal to 10 μm; PNC = particle number concentration; PM_{coarse} = particulate matter with an aerodynamic diameter of 2.5-10 μm; $PM_{2.5abs}$ = $PM_{2.5}$ absorbance; NO_2 = nitrogen dioxide; NO_x = nitrogen oxide; O_3 = ozone. An IQR increase was 1.4 $\mu g/m^3$ for $PM_{2.5}$, 2.1 $\mu g/m^3$ for PM_{10}, $2.0\times10^3/cm^3$ for PNC, 1.4 $\mu g/m^3$ for PM_{coarse}, $0.3\times10^{-5}/m$ for $PM_{2.5abs}$, 7.2 $\mu g/m^3$ for NO_2, 8.8 $\mu g/m^3$ for NO_x, and 3.4 $\mu g/m^3$ for O_3.

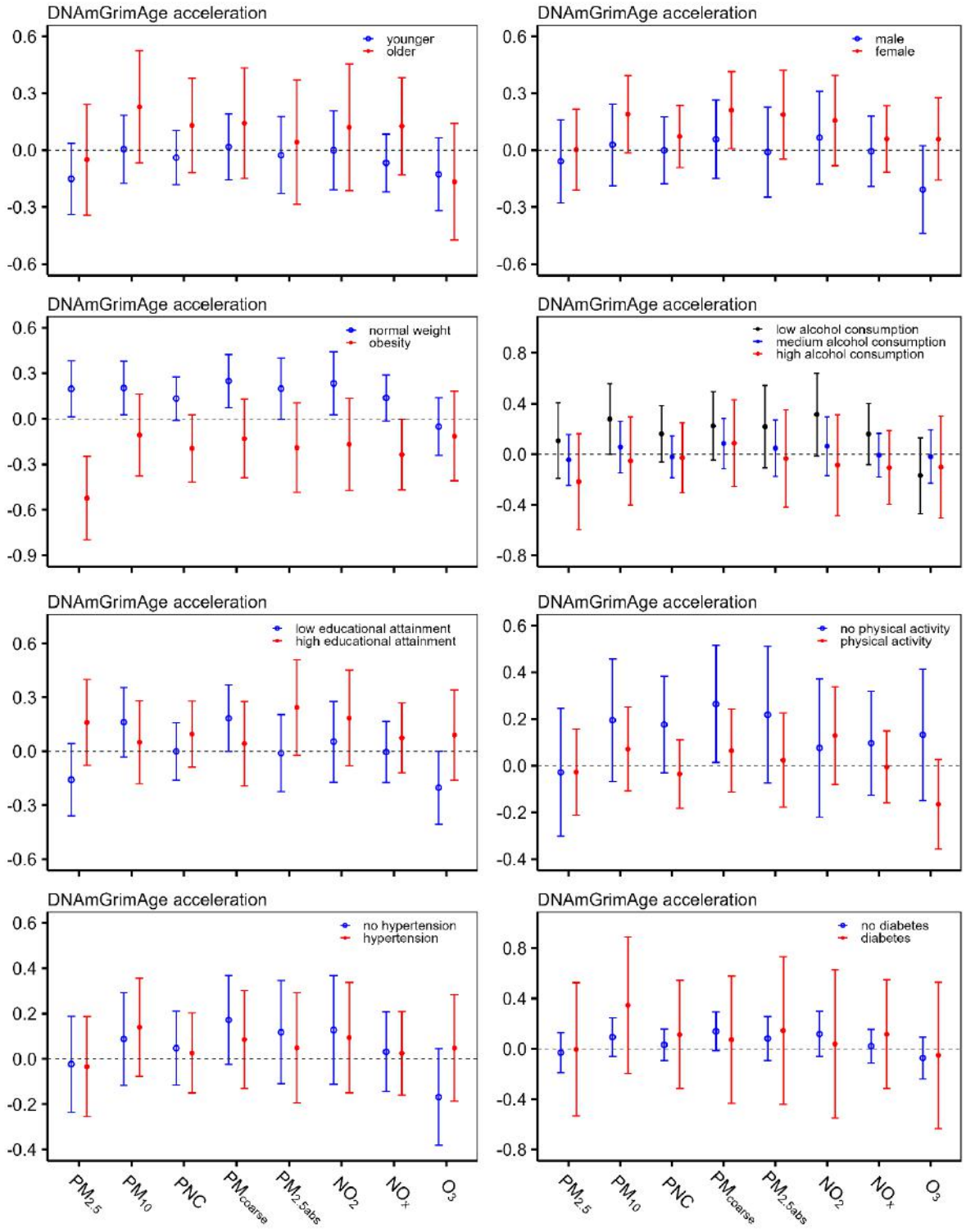


Figure S9. Absolute changes (95% CI, years) in DNAmGrimAge acceleration per IQR increase in air pollutant concentrations stratified by categorized age group (≤ 65 vs > 65), sex, categorized BMI group, categorized group of alcohol consumption, educational attainment, physical activity, and history of diseases (hypertension, diabetes).

Covariate-adjusted linear mixed-effect regression models were used. Results were from our full model adjusted by age, sex, an indicator of each study waves visit (KORA S4, KORA F4, or KORA FF4), houseman estimated cells, batch, chip, educational attainment, smoking status, alcohol consumption, physical activity, BMI, hypertension, diabetes, HDL, and LDL.

$PM_{2.5}$ = particulate matter with an aerodynamic diameter less than or equal to 2.5 μm; PM_{10} = particulate matter with an aerodynamic diameter less than or equal to 10 μm; PNC = particle number concentration; PM_{coarse} = particulate matter with an aerodynamic diameter of 2.5-10 μm; $PM_{2.5abs}$ = $PM_{2.5}$ absorbance; NO_2 = nitrogen dioxide; NO_x = nitrogen oxide; O_3 = ozone. An IQR increase was 1.4 $\mu g/m^3$ for $PM_{2.5}$, 2.1 $\mu g/m^3$ for PM_{10}, $2.0\times10^3/cm^3$ for PNC, 1.4 $\mu g/m^3$ for PM_{coarse}, $0.3\times10^{-5}/m$ for $PM_{2.5abs}$, 7.2 $\mu g/m^3$ for NO_2, 8.8 $\mu g/m^3$ for NO_x, and 3.4 $\mu g/m^3$ for O_3.

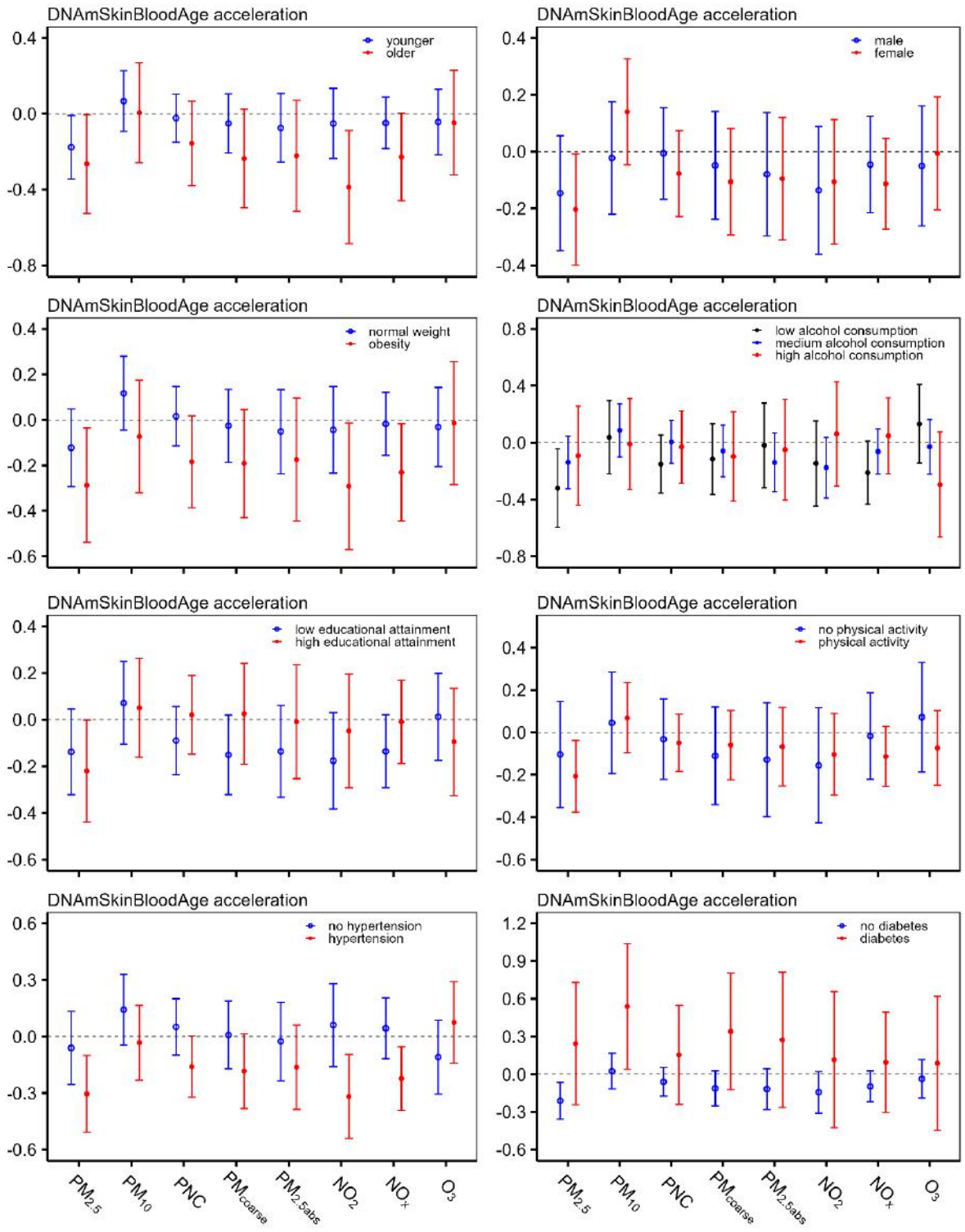


Figure S10. Absolute changes (95% CI, years) in DNAmSkinBloodAge acceleration per IQR increase in air pollutant concentrations stratified by categorized age group (≤ 65 vs > 65), sex, categorized BMI group, categorized group of alcohol consumption, educational attainment, physical activity, and history of diseases (hypertension, diabetes).

Covariate-adjusted linear mixed-effect regression models were used. Results were from our full model adjusted by age, sex, an indicator of each study waves visit (KORA S4, KORA F4, or KORA FF4), houseman estimated cells, batch, chip, educational attainment, smoking status, alcohol consumption, physical activity, BMI, hypertension, diabetes, HDL, and LDL.

$PM_{2.5}$ = particulate matter with an aerodynamic diameter less than or equal to 2.5 μm; PM_{10} = particulate matter with an aerodynamic diameter less than or equal to 10 μm; PNC = particle number concentration; PM_{coarse} = particulate matter with an aerodynamic diameter of 2.5-10 μm; $PM_{2.5abs}$ = $PM_{2.5}$ absorbance; NO_2 = nitrogen dioxide; NO_x = nitrogen oxide; O_3 = ozone. An IQR increase was 1.4 $\mu g/m^3$ for $PM_{2.5}$, 2.1 $\mu g/m^3$ for PM_{10}, $2.0\times10^3/cm^3$ for PNC, 1.4 $\mu g/m^3$ for PM_{coarse}, $0.3\times10^{-5}/m$ for $PM_{2.5abs}$, 7.2 $\mu g/m^3$ for NO_2, 8.8 $\mu g/m^3$ for NO_x, and 3.4 $\mu g/m^3$ for O_3.

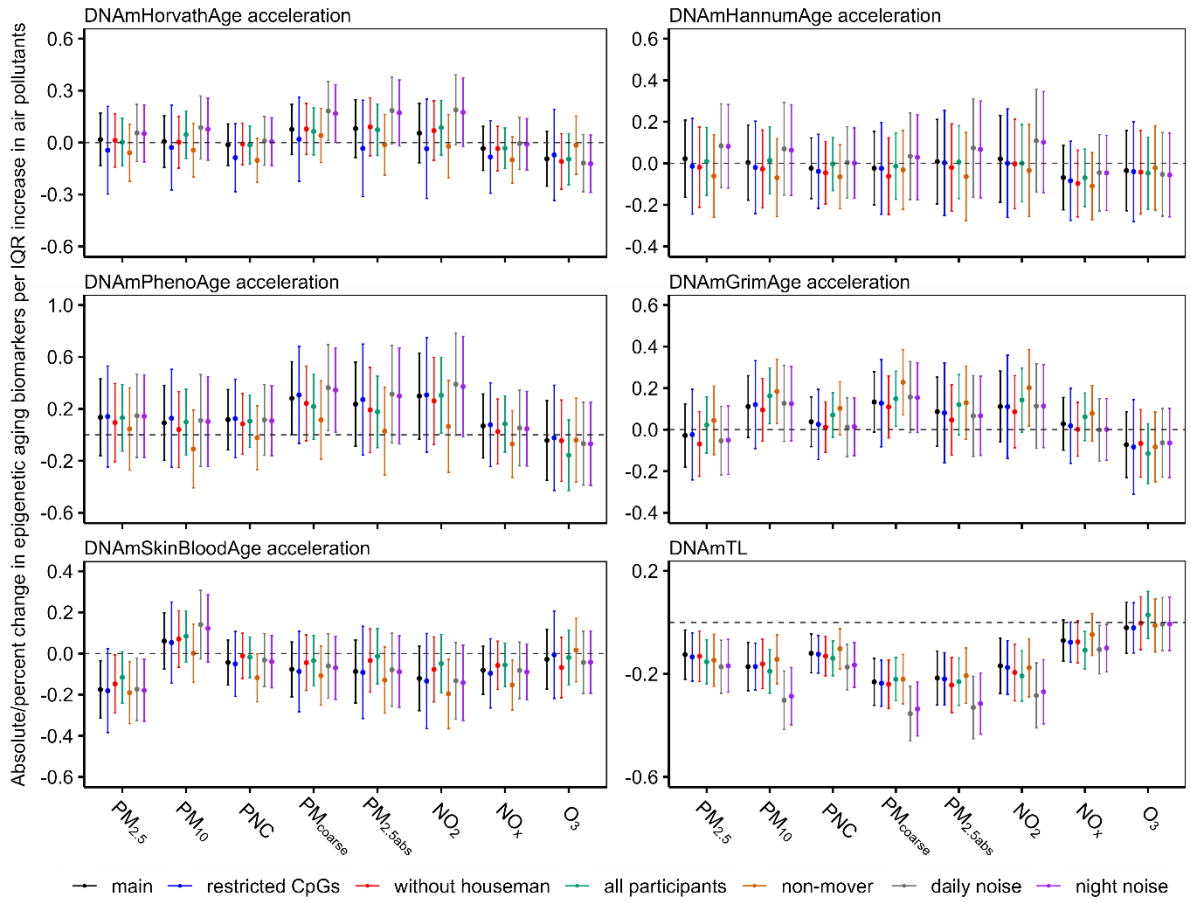


Figure S11. Absolute change (95% CI, years) of epigenetic age acceleration per IQR increase in air pollutant concentrations with basic, behavioral, clinical, and full covariate adjustment. For DNAmTL, the percent change (95% CI) is displayed.

Covariate-adjusted linear mixed-effect regression models were used. The confounders used in different sensitivity analyses were the same as used in our full model including by age, sex, an indicator of each study waves visit (KORA S4, KORA F4, or KORA FF4), houseman estimated cells, batch, chip, educational attainment, smoking status, alcohol consumption, physical activity, BMI, hypertension, diabetes, HDL, and LDL.

Main: results from our full model; Restricted_CpG: Epigenetic clocks calculation after exclusion of missing CpG sites in MethylationEPIC (KORA FF4) from KORA S4 and F4 methylation dataset; Without houseman: an exclusion of houseman estimated cells from the confounders; All participants: An inclusion of all participants whoever attended KORA S4, F4 or FF4 once or more times and meet the data recruitments. Non-mover: participants without changing their residential address throughout the entire study period; Daily/night noise: an inclusion of daily or night noise correspondingly as the second environmental exposure.

$PM_{2.5}$ = particulate matter with an aerodynamic diameter less than or equal to 2.5 µm; PM_{10} = particulate matter with an aerodynamic diameter less than or equal to 10 µm; PNC = particle number concentration; PM_{coarse} = particulate matter with an aerodynamic diameter of 2.5-10 µm; $PM_{2.5abs}$ = $PM_{2.5}$ absorbance; NO_2 = nitrogen dioxide; NO_x = nitrogen oxide; O_3 = ozone. An IQR increase was 1.4 µg/m^3 for $PM_{2.5}$, 2.1 µg/m^3 for PM_{10}, 2.0×10^3/cm^3 for PNC, 1.4 µg/m^3 for PM_{coarse}, 0.3×10^{-5}/m for $PM_{2.5abs}$, 7.2 µg/m^3 for NO_2, 8.8 µg/m^3 for NO_x, and 3.4 µg/m^3 for O_3.

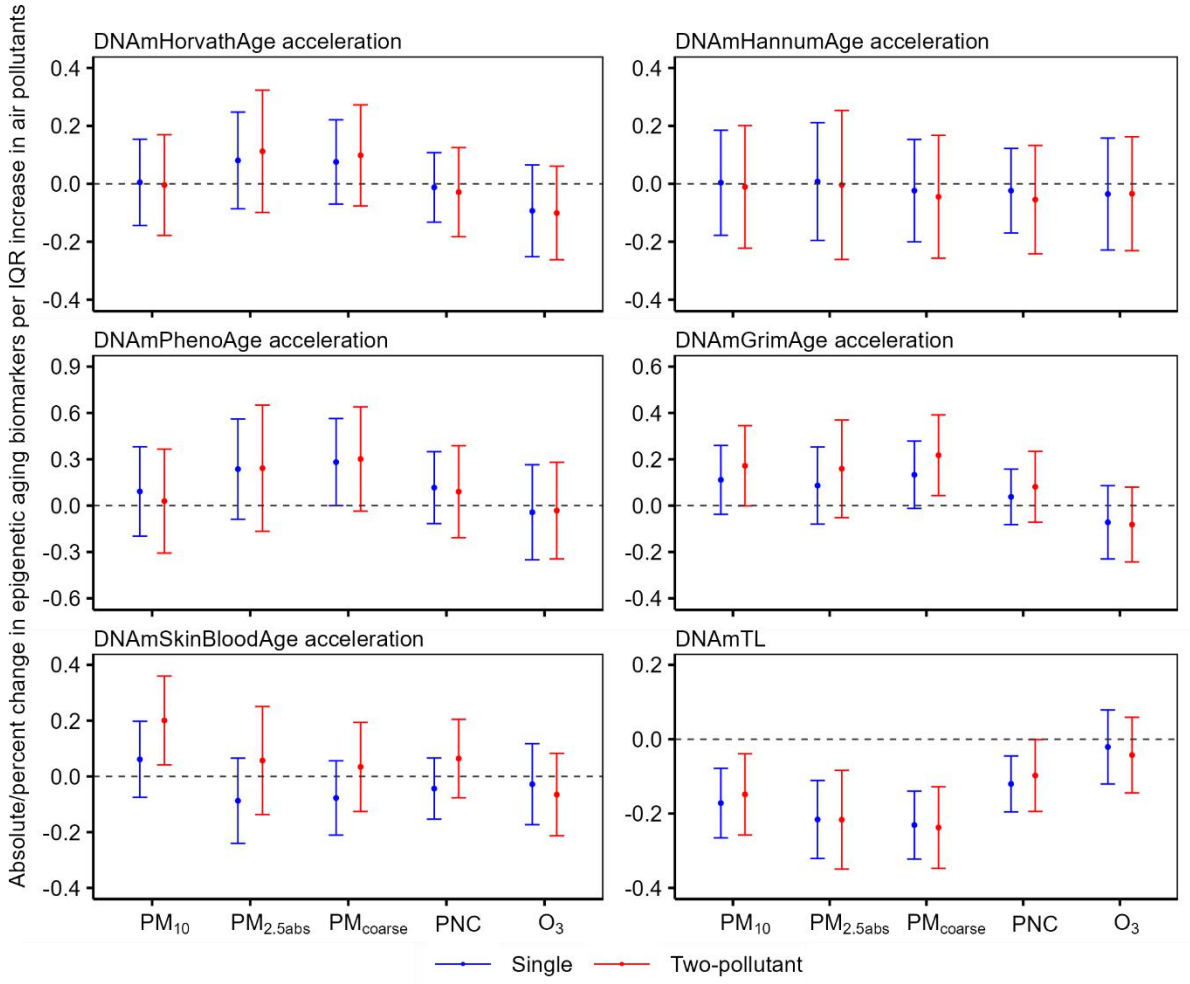


Figure S12. Comparison between single and two-pollutant models after additional inclusion of $PM_{2.5}$ (percent change only for DNAmTL).

Covariate-adjusted linear mixed-effect regression models were used. Results were from our full model adjusted by age, sex, an indicator of each study waves visit (KORA S4, KORA F4, or KORA FF4), houseman estimated cells, batch, chip, educational attainment, smoking status, alcohol consumption, physical activity, BMI, hypertension, diabetes, HDL, and LDL.

$PM_{2.5}$ = particulate matter with an aerodynamic diameter less than or equal to 2.5 μm; PM_{10} = particulate matter with an aerodynamic diameter less than or equal to 10 μm; PNC = particle number concentration; PM_{coarse} = particulate matter with an aerodynamic diameter of 2.5-10 μm; $PM_{2.5abs}$ = $PM_{2.5}$ absorbance; NO_2 = nitrogen dioxide; NO_x = nitrogen oxide; O_3 = ozone. An IQR increase was 1.4 $\mu g/m^3$ for $PM_{2.5}$, 2.1 $\mu g/m^3$ for PM_{10}, 1.4 $\mu g/m^3$ for PM_{coarse}, $2.0\times10^3/cm^3$ for PNC, $0.3\times10^{-5}/m$ for $PM_{2.5abs}$, 7.2 $\mu g/m^3$ for NO_2, 8.8 $\mu g/m^3$ for NO_x, and 3.4 $\mu g/m^3$ for O_3.

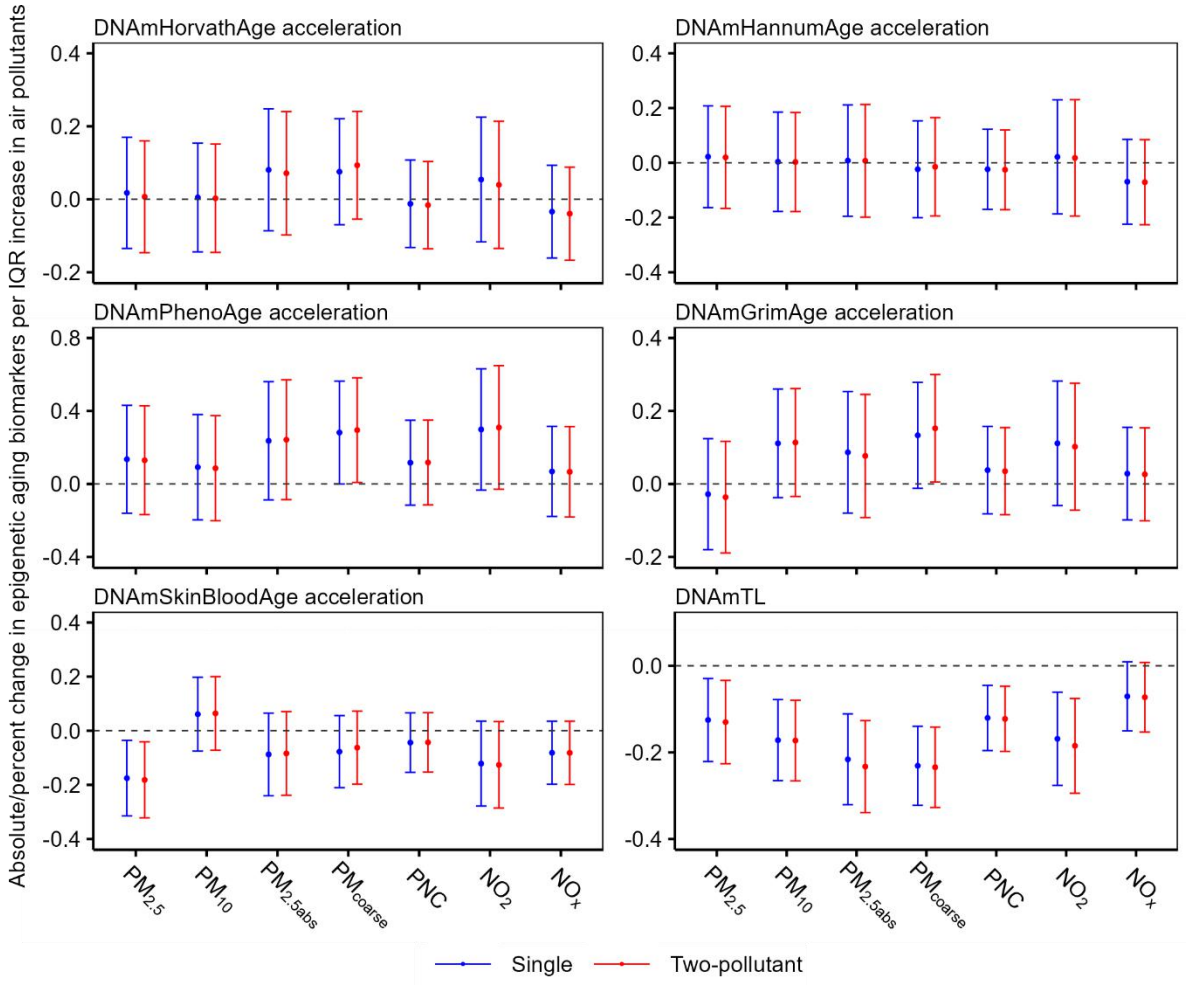


Figure S13. Comparison between single and two-pollutant models after additional inclusion of O_3 (percent change only for DNAmTL).

Covariate-adjusted linear mixed-effect regression models were used. Results were from our full model adjusted by age, sex, an indicator of each study waves visit (KORA S4, KORA F4, or KORA FF4), houseman estimated cells, batch, chip, educational attainment, smoking status, alcohol consumption, physical activity, BMI, hypertension, diabetes, HDL, and LDL.

$PM_{2.5}$ = particulate matter with an aerodynamic diameter less than or equal to 2.5 μm; PM_{10} = particulate matter with an aerodynamic diameter less than or equal to 10 μm; PNC = particle number concentration; PM_{coarse} = particulate matter with an aerodynamic diameter of 2.5-10 μm; $PM_{2.5abs}$ = PM2.5 absorbance; NO_2 = nitrogen dioxide; NO_x = nitrogen oxide; O_3 = ozone. An IQR increase was 1.4 μg/m^3 for $PM_{2.5}$, 2.1 μg/m^3 for PM_{10}, 1.4 μg/m^3 for PM_{coarse}, 2.0×10^3/cm^3 for PNC, 0.3×10^{-5}/m for $PM_{2.5abs}$, 7.2 μg/m^3 for NO_2, 8.8 μg/m3 for NO_x, and 3.4 μg/m^3 for O_3.

Further projects

I was responsible for a research project investigating the associations between genome-wide DNA methylation patterns and short-, medium-, and long-term exposure to air pollutants. This project utilized longitudinal data from the KORA cohort study conducted in Augsburg, Germany, and DNA methylation was quantified using Infinium methylation arrays. My responsibilities encompassed drafting the statistical analysis plan, managing data processing and quality control, conducting statistical analyses, summarizing results, interpreting findings, and preparing visualizations. The analyses have been completed, and I am currently working on the results for manuscripts.

Additionally, I contributed to the CHARGE consortium's pollution-wide Epigenome-Wide Association Study (EWAS) project, specifically utilizing the KORA F4 DNA methylation dataset. Our objective was to replicate findings related to long-term exposure to near-roadway air pollution and its associations with DNA methylation patterns. My role included data application, data quality control procedures, data processing, statistical analyses, and results summarization and visualization. This project was successfully completed and is currently awaiting publication.

I was also involved in a collaborative project assessing the predictive capability of DNA methylation-based biomarkers for major health outcomes, including all-cause mortality, myocardial infarction, stroke, and cancer. KORA F4 served as a replication cohort, and my contributions involved data processing, conducting sensitivity analyses, and actively participating in manuscript reviewing, editing, and revision. This project was successfully completed and resulted in a peer-reviewed publication titled "DNA methylation-based biomarkers of age acceleration and all-cause death, myocardial infarction, stroke, and cancer in two cohorts: the NAS and KORA F4."

Finally, I participated in a study examining the relationship between epigenetic aging and allergic diseases. Specifically, I was responsible for calculating epigenetic clock measures using DNA methylation data from KORA F4. Additionally, I contributed to manuscript editing and revision processes. This study has successfully concluded, resulting in a published article titled "Childhood Asthma and Allergy Are Related to Accelerated Epigenetic Aging."

Acknowledgements

Standing at the end of my PhD journey, I look back on this path—one intertwined with both thorns and starlight. It has deepened my understanding of academic exploration and the resilience of life. The pandemic, cyberattacks, and unexpected accidents—events that seemed more fitting for books and movies—dramatically became part of my academic reality. However, I never retreated, and now I have finally reached this moment. With a heart full of gratitude, I wish to extend my sincerest thanks to everyone who has supported me throughout my PhD.

First and foremost, I would like to express my sincere gratitude to my supervisor Prof. Annette Peters and my reporting supervisor Dr. Alexandra Schneider for giving me the opportunity to pursue my PhD at LMU and HMGU. I still remember that during the interview, I was so nervous that I spoke English haltingly, fearing that I would lose this opportunity. But Dr. Alexandra Schneider saw potential in me, and helped me complete all the necessary documents for my scholarship application within only two weeks. Throughout my PhD, she provided meticulous guidance in every aspect from developing my research plan to structuring each paper's analysis, discussing the results, and refining manuscripts.

I am deeply grateful to Prof. Annette Peters. The questions, discussions, and insights she provided during our Thesis Advisory Committee meetings, courses, conferences, presentations, and paper revisions have been invaluable. I have always been amazed by her vast knowledge and profound perspectives, and I feel incredibly lucky to have been able to learn from her. I also extend my sincere thanks to Prof. Dennis Nowak and Dr. Cavin Ward-Caviness, members of my Thesis Advisory Committee, for their professional and insightful suggestions on my analyses.

I would like to thank Dr. Susanne Breitner and Dr. Kathrin Wolf for their support with my projects. Their detailed feedback on my statistical analysis plans, results, presentations, as well as their thorough revisions of my papers, were instrumental. Thanks to Dr. Siqi Zhang for patiently explaining coding and statistical methods and guiding me through my research. Thanks to Dr. Rui Wang-Sattler for providing metabolomics data and assisting me in data normalization. Thanks to Dr. Melanie Waldenberger, whose expertise helped me navigate my initial encounters with DNA methylation data analysis. Special thanks to Dr. Marco Dallavalle, who patiently guided me when I struggled with the data analysis on the remote server.

I had originally planned to complete and submit my PhD thesis by the end of 2023. However, an unexpected accident in August abruptly put my life on pause. At first, unaware of the severity of my injuries, I assumed I would recover quickly. But when I had to relearn even the simplest actions, I realized the true complexity of the situation. The physical pain and psychological distress pushed me to the edge more than once. However, I was never alone. I am immensely grateful to all my supervisors, colleagues, and friends who provided unwavering support after the accident. Their care allowed me to focus entirely

on my recovery in those crucial early days. Special thanks to Dr. Alexandra Schneider, who fought to secure an additional year for my PhD contract, and to Prof. Annette Peters, for offering me the opportunity to continue as a postdoctoral researcher in the future. Thank you, Dr. Kathrin Wolf, for your endless kindness and for taking care of so many administrative matters despite your busy schedule. Your support has been truly heart-warming.

Setting aside the accident, my five years at Helmholtz have been filled with joy and unforgettable experiences. I am deeply grateful for the friends I have met here, for your unwavering support and warmth. There are so many names that I will cherish forever. Thank you, Dr. Mahnaz Badpa and Dr. Masna Rai, for helping me navigate my initial anxieties in a foreign country and for bringing laughter and surprises into my life. Thank you, Dr. Megi Vogli, for always bringing me to exciting new activities, enriching my time outside of academia, and standing by me throughout my hospitalization and recovery. Thanks to Dr. Maximilian Schwarz, Lisa Maier, Fiona Niedermayer, and Dr. Nikolaos Nikolaou, your companionship made my PhD journey much more vibrant and also made post-accident life much easier.

Thank you, Dr. Hong Luo, Minqi Liao, Feiling Ai, Dr. Jiesheng Lin, Yujiao Li, and Dr. Mingming Wang, all journeys and the sunsets we watched together remain some of my most cherished memories. Thank you for taking turns visiting me in the hospital every week, ensuring I never felt alone during my recovery. Thank you, Bing Huo, Wenlu Sun, and Xiaoxue Bao, for braving the storm to bring me food I could eat, helping me get through two consecutive surgeries. A special thanks to Hong, most of my memories from the intensive care unit are pieced together through our chat records. No matter how exhausted you were, you always found time to visit me in the hospital. When I was at my lowest, your words pulled me back from despair. And thank you for always taking care of my family. To my dearest friends back home, Hongwei Yan and Mingxing Zhao, thank you for encouraging me at every stage of my recovery and for your patience in understanding all my emotions. Thank you, Changying Jing, for taking me on trips to help me find peace and joy again.

The deepest gratitude belongs to my parents and two younger brothers. Thank you for always supporting my decisions. I know there were moments you questioned whether this accident could have been avoided if I had not pursued a PhD abroad. Yet, despite those doubts, you have provided me endless encouragement, strengthening my resolve to face any challenge. Thank you to my extended family, even though we are thousands of miles apart, your unconditional love and care remain steadfast.

Furthermore, I would like to express my heartfelt gratitude to the China Scholarship Council for providing me with a four-year scholarship. I would also like to extend my heartfelt thanks to the PhD-EPH community at IBE. A special note of appreciation goes to Dr. Annette Hartmann and Ms. Monika Darchinger from the PhD office. Their unwavering support, dedication, and behind-the-scenes efforts have been the backbone of this program, making a profound difference in our academic journey. There are so many people I need to thank, and while I may not have listed every name here, I will never

forget each act of kindness and support I received. Life can be harsh and unforgiving at times, but the warmth you have given me has been enough to dispel the darkness and light up the nights when I felt like giving up.

Finally, I want to thank myself. I always believed that hard work could accomplish anything. But this journey has shown me that, no matter how much effort I put in, some things are simply beyond my control. There were frustrating moments, but no matter how painful or difficult, I would persevere and eventually reclaim my freedom.

I have always loved this quote from *To Kill a Mockingbird*:

“I wanted you to see what real courage is, instead of getting the idea that courage is a man with a gun in his hand. It’s when you know you’re licked before you begin but you begin anyway and you see it through no matter what. You rarely win, but sometimes you do.”

An ending is also a new beginning. I will carry all the warmth I have received along this journey and step forward with courage. “*Shoot for the moon, even if you miss, you will land among the stars*.”

List of all scientific publications to date

Yao Y, Schneider A, Wolf K, Zhang S, Wang-Sattler R, Prehn C, Adamski J, Peters A, Breitner S. Longitudinal associations between metabolites and immediate, short- and medium-term exposure to ambient air pollution: Results from the KORA cohort study. Sci Total Environ. 2023; 900:165780. doi: 10.1016/j.scitotenv.2023.165780.

Yao Y, Schneider A, Wolf K, Zhang S, Wang-Sattler R, Peters A, Breitner S. Longitudinal associations between metabolites and long-term exposure to ambient air pollution: Results from the KORA cohort study. Environ Int. 2022; 170: 107632. doi: 10.1016/j.envint.2022.107632.

Yao Y, Wang D, Ma H, Li C, Chang X, Low P, Hammond SK, Turyk ME, Wang J, Liu S. The impact on T-regulatory cell related immune responses in rural women exposed to polycyclic aromatic hydrocarbons (PAHs) in household air pollution in Gansu, China: A pilot investigation. Environ Res. 2019; 173:306-317. doi: 10.1016/j.envres.2019.03.053.

Yao Y, Chang X, Wang D, Ma H, Wang H, Zhang H, Li C, Wang J. Roles of ERK1/2 and PI3K/AKT signaling pathways in mitochondria-mediated apoptosis in testes of hypothyroid rats. Toxicol Res (Camb). 2018; 27;7(6):1214-1224. doi: 10.1039/c8tx00122g.

Chang XR*, **Yao YL***, Wang D, Ma HT, Gou PH, Li CY, Wang JL. Influence of hypothyroidism on testicular mitochondrial oxidative stress by activating the p38 mitogen-activated protein kinase and c-Jun NH_2-terminal kinase signaling pathways in rats. Hum Exp Toxicol. 2019; 38(1):95-105. doi: 10.1177/0960327118781927. (**co-first author*)

Wang C, Ni W, **Yao Y,** Just A, Heiss J, Wei Y, Gao X, Coull BA, Kosheleva A, Baccarelli AA, Peters A, Schwartz JD. DNA methylation-based biomarkers of age acceleration and all-cause death, myocardial infarction, stroke, and cancer in two cohorts: The NAS, and KORA F4. EBioMedicine. 2021; 63:103151. doi: 10.1016/j.ebiom.2020.103151.

Leskien M, Thiering E, Yu Z, Huels A, **Yao Y**, Merid SK, Gruzieva O, Weidinger S, Waldenberger M, Peters A, Melén E, Standl M. Childhood Asthma and Allergy Are Related to Accelerated Epigenetic Aging. Allergy. 2025; 0:0-11. doi: 10.1111/all.16583

Gou P, Chang X, Ye Z, **Yao Y,** Nguyen PK, Hammond SK, Wang J, Liu S. A pilot study comparing T-regulatory cell function among healthy children in different areas of Gansu, China. Environ Sci Pollut Res Int. 2017; 24(28):22579-22586. doi: 10.1007/s11356-017-9907-3.