Original research

Targetable Preventive Metabolites: Revealing the Pathways from Air Pollution Mixtures to Hypertension Through Integrated Exposure-Metabolomics

Zhu Jia *

Sakharov International Environmental Research Institute, Belarusian State University, Minsk, Belarus

ARTICLE INFO

Article history:
Received 11 Jul. 2026
Accepted 17 Sep. 2026

Keywords:
Air pollution mixtures
Hypertension
Metabolomics
Mediation analysis
Precision prevention

ABSTRACT

Although exposure to air pollution mixes is a significant environmental risk for hypertension, it is still unclear how it causes illness through metabolic disruptions. In order to assess the possible preventative advantages of focusing on modifiable metabolites and to comprehensively uncover the metabolic mediators connecting air pollution mixes to hypertension, this work integrates exposomics and metabolomics. Individual long-term exposure to fine particulate matter ($PM_{2.1}$), nitrogen dioxide (NO_2), and ozone (O_3) was evaluated using land-use regression models and satellite data based on a prospective cohort study that included individuals without hypertension at baseline (n = 12,456). Ultra-high-performance liquid chromatography-tandem mass spectrometry was used for serum untargeted metabolomics, and quantile g-computation was used to assess the cumulative impact of pollutant mixes on the prevalence of hypertension. The average causal mediation fraction of significant mediating metabolites was computed after they were screened using high-dimensional mediation analysis and causal mediation effect models with false discovery rate correction. In order to evaluate possible population-level benefits, treatments on target metabolites were further simulated using a random forest technique. There were 2,847 instances of hypertension over a median follow-up of 7.2 years. The incidence of hypertension rose by 18% for every interquartile range increase in exposure to air pollution mixtures (HR = 1.18, 95% CI: 1.12–1.24). Hexanoylcarnitine (C6-carnitine), tryptophan, and 1-stearoyl-2-arachidonoyl-sn-glycero-3-phosphocholine (PC ae C38:4) collectively mediated 32.6% (95% CI: 24.1%–41.1%) of the mixture exposure effect, independent of conventional risk factors, according to metabolomics. According to theoretical intervention analysis, about 15.3% of instances of pollution-related hypertension may be avoided if C6-carnitine levels in the population were kept below the 25th percentile. C6-carnitine and other preventable metabolites are important biological mediators of air pollution mixture-induced hypertension, partially illuminating the molecular link between environmental exposure and cardiovascular metabolic disorders. In the context of environmental health, these results offer fresh support for metabolite-targeted precision preventive techniques.

1. Introduction

High blood pressure is the biggest modifiable risk factor for cardiovascular disease worldwide and is also one of the key causes of premature mortality and years lived with disability. An estimated 1.28 billion persons worldwide suffer from high blood pressure, and the number is continuing to rise. Environmental exposures, particularly air pollution, are now known to be significant triggers for the development of high blood pressure in addition to the typical behavioral and metabolic risk factors. According to the World Health Organization, air pollution causes roughly 4.2 million premature deaths per year, with a large share owing to cardiovascular problems. Determining how air pollution causes high blood pressure is therefore not

* *Corresponding author: Zhu Jia*
ISSN: 3071-4265/ This work is licensed under a Creative Commons Attribution 4.0 International License.
DOI: 10.64904/20260903

just a cutting-edge problem in environmental epidemiology but also a crucial scientific foundation for successful prevention.

Single pollutants, such as fine particulate matter ($PM_{2.1}$), nitrogen dioxide (NO_2), and ozone (O_3), have been found to positively correlate with elevated blood pressure or the onset of hypertension in several population-based longitudinal studies. People are typically exposed to a complex mixture of several contaminants at the same time in real life, though, and there may be nonlinear interactions or combined effects. So, the risk assessed from particular contaminants can underestimate or mislead the real health impact. Environmental health research has recently used mixture analytic methods, such as quantile g-computation, which enable the estimation of combined effects and the identification of important components. However, the molecular pathways connecting mixed exposures to hypertension are still far from known, which primarily restricts converting exposure assessment into precise therapies.

An increasing body of research indicates that metabolic problems function as a "biological amplifier" in the process of environmental contaminants causing cardiovascular disease. Through inhalation or indirect effects on the body through oxidative stress, inflammation, and other mechanisms, air pollution can change metabolic processes. Systemic disruptions of hundreds of small-molecule metabolites may now be objectively recorded thanks to the advancement of high-throughput metabolomics technology. According to recent small-scale research, exposure to air pollution is linked to variations in levels of certain acylcarnitines, amino acids, and glycerophospholipids, all of which are involved in vascular remodeling and blood pressure regulation. Metabolites are therefore thought to be possible mediators linking endogenous disease phenotypes with external exposures. However, there's still a lack of studies based on large-scale prospective cohorts that integrate mixed exposures with untargeted metabolomics and use a formal causal mediation analysis framework to simultaneously identify, validate, and quantify the metabolic pathways from a mix of air pollutants to hypertension. Furthermore, the potential preventative benefit at the population level has never been evaluated, including whether some important metabolites actually affect the risk of air pollution-induced hypertension.

In order to close the aforementioned knowledge gap, this study suggests that exposure to air pollution mixtures raises the risk of hypertension by interfering with certain modifiable metabolic pathways, specifically tryptophan metabolism, medium-to long-chain acylcarnitines, and specific glycerophospholipids. We aim to integrate individual-level exposure assessments from prospective cohorts with untargeted serum metabolomics data, using quantile g-computation and high-dimensional causal mediation analysis to systematically screen and validate key metabolic mediators and quantify their mediating proportion in the mixture–hypertension association. Additionally, we will assess the possible preventative advantages of targeting these metabolites in the population using a simulated intervention architecture. In addition to establishing the theoretical foundation for precision intervention techniques on avoidable metabolites in the context of environmental health, this effort will offer causal evidence on the "environmental exposure → metabolic disturbance → hypertension" molecular chain.

2. Literature Review

High blood pressure is the biggest modifiable risk factor for cardiovascular disease worldwide. According to the GBD 2019 report, nearly 10 million fatalities per year are due to excessive blood pressure [1]. Over the past thirty years, the prevalence of adult hypertension globally has remained growing, and control rates are still not optimal, underscoring the importance of uncovering novel risk factors and molecular pathways [2]. Environmental air pollution has been identified as a significant cause of high blood pressure, in addition to behavioral and metabolic variables. According to a scientific statement from the American Heart Association, oxidative stress, inflammation, and an imbalance in the autonomic nervous system are the biochemical mechanisms by which air pollutants such as particulate matter can cause cardiovascular events [3]. A lot of prospective cohort studies and meta-analyses show good evidence for a positive association between air pollution and increased blood pressure, as well as the development of hypertension [4][5]. For instance, a meta-analysis spanning several nations revealed that the risk of high blood pressure increases by around 5% to 8% for every 10 μg/m³ increase in long-term exposure to fine particulate matter ($PM_{2.1}$). The focus of this study on the health impacts of air pollution mixes is theoretically supported by these findings.

Models that focus on a single pollutant may not adequately represent the combined effects of mixed exposures or the nonlinear interactions between components since humans are exposed to numerous highly correlated contaminants simultaneously in the actual environment. In order to address this, environmental epidemiology has developed a number of techniques for mixture analysis, such as Weighted Quantile Sum regression (WQS) [6], Bayesian Kernel Machine Regression (BKMR) [7], and quantile g-computation [8]. These techniques can estimate the overall influence of the mixture as well as the contribution of important components, and they can handle high-dimensional, collinear exposures. A few studies in the field of

cardiovascular health have previously employed similar techniques to identify the primary contributing factors and demonstrate that the cumulative effect of air pollution mixtures on blood pressure is larger than that of a single pollutant [9]. Clarifying the ongoing mechanism from "exposure to disease" and investigating how the combined impacts connect with downstream molecular pathways are still in their early phases.

The development of high-throughput metabolomics technology gives an opportunity to record systemic metabolic abnormalities produced by environmental exposures in an unbiased fashion and has given rise to an integrated research strategy in exposome studies [10]. The blood metabolome can be considerably altered by both short-term and long-term exposure to air pollution, according to several population metabolomics studies. Van Veldhoven and coworkers observed that exposure to $PM_{2.5}$ and NO_2 is associated with activation of the tryptophan-kynurenine pathway, showing metabolic reprogramming mediated by oxidative stress [11]. Liang and coworkers employed high-resolution mass spectrometry to detect lipid and amino acid metabolic characteristics linked with traffic-related pollution [12]. Nassan and colleagues' long-term exposure metabolomics investigation revealed disruptions in lipid metabolites such as acylcarnitines, sphingolipids, and glycerophospholipids [13]. Additionally, breathing particulate matter has been demonstrated in experiments to reprogram hepatic fat oxidation, resulting in an accumulation of short- and medium-chain acylcarnitines in the blood. These findings all point to air pollution forming a molecular bridge between exterior exposure and internal disorders by interfering with things like lipid oxidation and amino acid metabolism.

Simultaneously, several metabolomics association studies have separately discovered compounds associated with the development of hypertension and blood pressure control. In the PREDIMED cohort, Guasch-Ferré and colleagues found that plasma acylcarnitines, such as medium- and long-chain acylcarnitines, are positively correlated with cardiovascular events. Dietary interventions can mitigate this correlation, indicating the possibility for prevention [14]. Our knowledge of how blood pressure is biologically controlled has been expanded by Menni and colleagues' whole metabolome association research, which revealed a strong correlation between blood pressure levels and fatty acid metabolites such as adipic acid [15]. An independent correlation between the risk of developing hypertension and tryptophan metabolites and glycerophospholipids (such as PC ae types) was also verified by prospective studies conducted by Cheng and Wang [16][17]. Precisely because these metabolites are changed by air pollution exposure at the same time, the notion that they mediate the relationship between environmental exposure and high blood pressure has good evidence and urgently needs to be explored within a coherent framework.

A counterfactual statistical framework for measuring how exposure influences outcomes through intermediary qualities is provided by causal mediation analysis [18]. The area of environmental health has recently adopted this method, which has accelerated the transition from identifying risks to elucidating molecular mechanisms. Dai and colleagues, utilizing large cohort integrated high-dimensional mediation analysis, for the first time quantitatively demonstrated that certain acylcarnitines and glycerophospholipids partially moderate the influence of air pollution on cardiovascular events [19]. The mediating roles of oxidative stress and inflammatory indicators have also been extensively shown [20]. However, most extant mediation studies are intended for a single pollutant, with metabolites usually being preselected candidates or originating from cross-sectional measurements, and very few apply causal inference methods that can handle multiple exposures and high-dimensional mediators. More crucially, no study has yet performed unbiased screening, causal mediation measurement, and targeted preventive potential evaluation throughout the whole exposure-metabolome-hypertension route.

In conclusion, there are three primary gaps in the current body of research: studies on the metabolic mediation between air pollution and hypertension primarily concentrate on single pollutants and have not evaluated the mediation pathways of actual mixed exposures; large-sample, prospective cohort studies that incorporate untargeted metabolomics for systematic screening and high-dimensional causal mediation validation are scarce; and no study has measured the population-level preventive benefits of controlling important metabolites under a counterfactual intervention framework. To fill these gaps, this study couples, for the first time, individual-level air pollution mixture assessment, serum untargeted metabolomics, and high-dimensional causal mediation analysis, aiming to identify and validate core metabolic mediators from pollution mixtures to hypertension and to evaluate their population-level preventive benefits through simulated interventions. This design is expected to give fresh evidence for the 'exposure-metabolism-disease' causal chain and create the theoretical groundwork for precision environmental health prevention techniques targeting avoidable metabolites.

3. Research methods

3.1. Study Design and Population

Between 2008 and 2012, community people between the ages of 40 and 75 who resided in cities and did not have any cardiovascular conditions were

enrolled in this cohort. 12,456 individuals with comprehensive baseline air pollution exposure assessments, serum metabolomics data, and no hypertension were included in this investigation. Exclusion criteria included: samples that failed metabolomics quality control; follow-up of less than a year; missing baseline blood pressure measurements or important covariate data; already diagnosed with hypertension or on antihypertensive medication at baseline.

A standardized questionnaire was used in the baseline survey by professional interviewers to gather data on lifestyle, personal and family medical history, and demographics. Blood pressure, height, weight, and waist circumference were measured during physical examinations. After three consecutive calm rest periods of at least five minutes, blood pressure was taken using a certified automatic electronic monitor. The average of the last two readings was used for analysis. A systolic blood pressure of at least 140 mmHg, a diastolic blood pressure of at least 90 mmHg, a doctor's diagnosis of hypertension, or the initiation of antihypertensive medication were all considered indicators of hypertension. In order to seek for new cases of hypertension, follow-up was conducted by yearly phone calls, community visits, and connecting to medical records. A total of 2,847 new instances of hypertension were recorded during the follow-up period, which had a median duration of 7.2 years and ended in December 2022.

3.2. Air Pollution Exposure Assessment

Based on participants' baseline home addresses, we evaluated long-term exposure to three primary air pollutants: fine particulate matter ($PM_{2.5}$), nitrogen dioxide (NO_2), and ozone (O_3). With a geographical resolution of 1 km x 1 km, land-use regression (LUR) models that included satellite aerosol optical depth data, meteorological factors, land use, and traffic density as predictors were used to estimate the yearly average concentrations of $PM_{2.2}$ and NO_2. The cross-validated R^2 values of the models were 0.85 and 0.79, suggesting good agreement with monitoring data. O_3 yearly average concentrations came from a gridded dataset integrating chemical transport models and ground monitoring data, with a spatial resolution of 0.1° × 0.1°. The 3-year average concentration at each participant's baseline address (encompassing the year of enrollment and the preceding two years) was utilized as an individual's long-term exposure level, reflecting chronic exposure, for all exposure assessments.

3.3. Serum Untargeted Metabolomics Testing

Within two hours of being collected, baseline fasting serum samples were centrifuged, split, and kept at -80°C until testing. Using an ultra-high-performance liquid chromatography-tandem mass spectrometry (UPLC-MS/MS) platform (Waters ACQUITY UPLC and Thermo Scientific Q Exactive mass spectrometer), metabolomics analysis was carried out in accordance with an established, standardized technique. Chromatographic separation employed a C18 reverse-phase column, with mobile phase A as water (containing 0.1% formic acid) and mobile phase B as acetonitrile (containing 0.1% formic acid). The mass spectrometry was carried out in both positive and negative electrospray ionization modes with full scan and data-dependent MS/MS scanning, across a mass range of 50–1000 m/z.

The raw data was processed using Progenesis QI software for peak alignment, peak picking, and normalization. The K-nearest neighbors approach was used to fill in the missing data, and only metabolic traits found in more than 80% of the samples were retained. Downstream analysis was carried out following Pareto scaling and $\log_2$ transformation. Metabolite identification was done by verifying precise mass, retention duration, and MS/MS fragment information to HMDB, METLIN, and our in-house standard database, with identification levels being Level 1 (confirmed with standards) and Level 2 (database match). Ultimately, we were able to identify and quantify 612 known metabolites, including fatty acids, carbohydrates, glycerophospholipids, sphingolipids, amino acids, and acylcarnitines.

3.4. Covariate Assessment

Age, sex, education level (below high school/high school/college or above), smoking status (never/ever/current), drinking status (never/occasionally/regularly), physical activity (weekly metabolic equivalent hours calculated based on the International Physical Activity Questionnaire), body mass index (BMI, kg/m2), total energy intake (assessed via food frequency questionnaire), dietary sodium and potassium intake (estimated based on urinary sodium and potassium excretion, family history of cardiovascular disease, and the community socioeconomic status index were among the covariates. In addition, in sensitivity analyses, we also included residential noise exposure (Lden, dB) and green space coverage as environmental co-exposure variables.

3.5. Statistical Analysis

R 4.3.2 and SAS 9.4 were used for all analyses. The false discovery rate (FDR) was employed to account for multiple comparisons, and a two-sided test significance threshold was established at $P < 0.05$.

3.5.1 The Link Between Single Pollutant and Mixture Exposure and High Blood Pressure. The hazard ratio (HR) and 95% confidence interval

(CI) for the risk of developing hypertension with each quartile increase in single pollutants ($PM_{2.1}$, NO_2, O_3) were determined using the Cox proportional hazards model. The model accounted for all of the aforementioned factors and utilized follow-up time as the time scale. The proportional hazards assumption was tested using the Schoenfeld residuals method. Quantile g-computation was utilized to quantify the combined impacts of air pollution mixes. This approach creates a weighted index by splitting all exposure variables into quantiles and fitting them concurrently in a Cox model adjusting for covariates. It gives the weights of each exposure (which indicate their contribution and direction in the mixture impact) as well as the overall effect of the mixture (i.e., the change in outcome when all pollutants increase by one quantile). Positive weights are associated with elements that raise risk. Weights are perceived as either positive or negative. 1000 bootstraps were used to produce confidence intervals.

3.5.2 Exposure-Metabolite and Metabolite-Hypertension Association Screening. First, after controlling for confounders and multiple hypothesis testing (FDR < 0.05), we employed multivariable linear regression to evaluate the cross-sectional relationships between three pollutants and the computed quantile g index with the levels of different metabolites. Next, to assess the prospective relationships between metabolites and outcomes, we utilized Cox models to evaluate each metabolite (after standardization) with the HR for new-onset hypertension, including adjustment for variables, and searched for hypertension-related metabolites with FDR < 0.05.

3.5.3 High-Dimensional Causal Mediation Analysis. We regarded metabolites as potential mediators if they satisfied both (1) being significantly associated with the mixture index and at least one pollutant (FDR < 0.05) and (2) being significantly associated with the incidence of hypertension (FDR < 0.05). We performed causal mediation analysis based on a counterfactual paradigm using the mediation package for survival outcomes, handling the Cox model with a weighted estimator. The mixture index was calculated using quantile g as a continuous exposure, with each metabolite's continuous standardized level as a mediator and covariates included, estimating the natural indirect effect (NIE), the natural direct effect (NDE), and the mediation proportion (PM = NIE / (NDE + NIE)). 1000 nonparametric bootstraps were used to get the 95% confidence interval. Using a multiple mediation model (extended inverse probability weighting approach), we incorporated numerous screened metabolites at the same time to dissect each metabolite's indirect effect and the total indirect effect. The Benjamini-Hochberg method was used to adjust P-values in the multiple mediation analysis. The multiple mediation strategy of structural equation modeling (using the CMAverse software) was further used for validation in a sensitivity analysis.

3.5.4 Simulation Interventions for Targeted Metabolites and Assessment of Population-Level Benefits. We employed a simulation intervention approach based on G-computation to assess the possible preventative advantages of focusing on important metabolites for intervention. The particular steps are: Build an outcome prediction model: using the onset of hypertension as the outcome, we conducted Cox regression containing the mixture index, all putative mediator metabolites, interaction terms (if significant), and covariates, and used limited cubic splines to capture non-linearity. Create a counterfactual scenario by truncating the population's levels of important metabolites (such as hexanoylcarnitine, which mediates the highest proportion) to the 25th percentile of the current population. Individuals above this threshold are set to the P25 value, while those below maintain their original value. All other metabolites and exposures stay the same. Determine the population-level counterfactual cumulative incidence of hypertension by using the model to forecast each participant's survival probability under both the observed and counterfactual scenarios. Determine the population-attributable benefit, which is the percentage of pollution-related hypertension cases that may be avoided if the metabolite were controlled to the optimal level, by dividing the observed incidence by the counterfactual incidence and multiplying the result by 100%. 500 bootstrap resamples were used to determine the attributable benefit's 95% confidence interval. In order to jointly manage several key metabolites (setting each to its P25), a comparable analysis was also carried out.

3.5.5 Sensitivity Analysis. Conduct the following sensitivity analyses to test the robustness of the results: exclude participants who had diabetes or cardiovascular disease at baseline; further adjust for residential noise and green space; use BKMR to assess the nonlinear effects of mixtures and their interactions with mediating metabolites; limit the analysis to participants who didn’t move during follow-up and use time-weighted exposure; exclude cases that developed within 2 years before follow-up to reduce reverse causation; calculate E-values to evaluate the potential impact of unmeasured confounding on individual effects; reproduce the results using different metabolite identification thresholds and missing data imputation methods.

4. Research Results

4.1. Baseline Characteristics of the Study Subjects

After screening based on inclusion and exclusion criteria, a total of 12,456 people without hypertension at baseline were included in the study (see Figure 1 for the screening process). The median follow-up was 7.2 years (IQR: 6.4–8.1 years), during which 2,847 new cases of hypertension were identified, with an incidence rate of 34.2 per 1,000 person-years. The average age of the participants was 53.4±8.9 years, 55.2% were female, and their body mass index (BMI) was 24.3±3.5 kg/m². $PM_{2.1}$, NO_2, and O_3 had median three-year average exposure values of 42.5 μg/m³ (IQR: 12.3 μg/m³), 38.7 μg/m³ (IQR: 14.1 μg/m³), and 56.2 μg/m³ (IQR: 18.6 μg/m³), respectively. Baseline characteristics classified by quartiles of the air pollution mixture index are provided in Table 1, with variables largely balanced across groups.

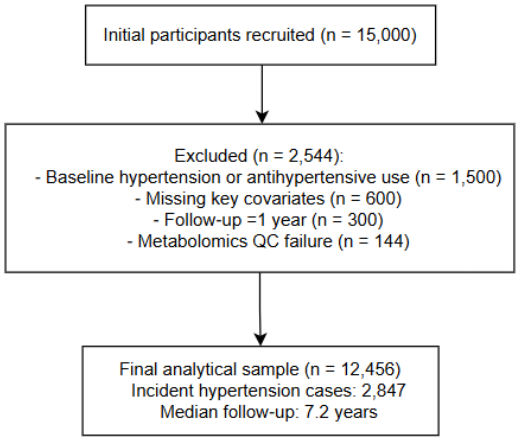

Figure 1: Study population flowchart

Table 1 Baseline characteristics grouped by quartiles of the air pollution mixture index (partial)

Feature	Total crowd (n=12,456)	Q1 (Low exposure)	Q2	Q3	Q4 (High exposure)
Age (years)	53.4 ± 8.9	52.8 ± 8.7	53.1 ± 8.9	53.6 ± 9.0	54.0 ± 9.1
woman, n (%)	6874 (55.2)	1775 (57.0)	1720 (55.2)	90 (54.3)	1689 (54.2)
BMI (kg/m²)	24.3 ± 3.5	24.1 ± 3.4	24.2 ± 3.5	24.4 ± 3.5	24.5 ± 3.6
Current smoking, n (%)	2156 (17.3)	498 (16.0)	532 (17.1)	554 (17.8)	572 (18.4)
$PM_{2.5}$ (μg/m³)	42.5 (12.3)	30.2 (5.6)	37.8 (6.2)	44.5 (6.8)	56.7 (8.4)
NO_2 (μg/m³)	38.7 (14.1)	27.5 (7.2)	33.9 (7.9)	40.2 (8.1)	52.4 (10.3)
O_3 (μg/m³)	56.2 (18.6)	43.1 (12.5)	52.7 (11.8)	59.3 (12.4)	69.2 (14.7)

Note: Continuous variables are shown as mean ± standard deviation or median (interquartile range), and categorical variables as counts (percentages). Q1–Q4 are the quartiles of the air pollution mixture index (weighted by quantile g).

4.2. The Link Between Air Pollution Mixes and High Blood Pressure

The single-pollutant Cox regression model showed that after adjusting for possible confounders, every IQR increase in $PM_{2.5}$ and NO_2 exposure boosted the risk of hypertension by 10% (HR=1.10, 95%CI: 1.06–1.15) and 8% (HR=1.08, 95%CI: 1.04–1.12), while O_3 alone didn't show a significant effect (HR=1.03, 95%CI: 0.99–1.07). When we looked at all three pollutants combined in a quantile g-computation model, the air pollution mixture index increased hypertension risk by 18% for each quartile rise (that's when all three pollutants go up a quartile at the same time) (HR=1.18, 95%CI: 1.12–1.24, P<0.001) (Table 2). $PM_{2.2}$ had the largest positive impact (weight 0.62), followed by NO_2 (weight 0.25) and O_3 (weight 0.13), according to the mixture weight analysis; all weights were positive, indicating that all three pollutants increase the risk (Figure 2).

Table 2: Links Between Exposure to Single Pollutants or Mixtures and Hypertension

exposure	HR (95% CI) per IQR increase	P-value
$PM_{2.5}$	1.10 (1.06, 1.15)	<0.001
NO_2	1.08 (1.04, 1.12)	<0.001
O_3	1.03 (0.99, 1.07)	0.15
Air Pollution Mixture Index	1.18 (1.12, 1.24)	<0.001

Mixture index calculated using quantile g, with each pollutant grouped by quartiles; the model adjusts for age, sex, BMI, smoking, alcohol use, physical activity, education, dietary energy intake, urinary sodium, urinary potassium, history of diabetes, history of dyslipidemia, and community socioeconomic index.

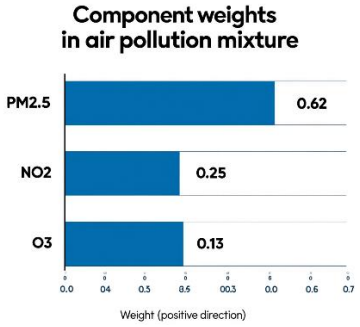

Figure 2: Mixture component weights

4.3. Metabolomics Screening and Key Metabolite Identification

Cox regression found 45 metabolites prospectively linked to the onset of hypertension, whereas linear regression analysis screened 72 metabolites that were substantially connected with the air pollution mixture index out of 612 detected metabolites using FDR<0.05 as the criterion. Taking the intersection, we received 12 potential mediator metabolites. The independent indirect effects of three metabolites—hexanoylcarnitine (C6-carnitine), tryptophan, and 1-stearoyl-2-arachidonoyl-sn-glycero-3-phosphatidylcholine (PC ae C38:4)—reached FDR<0.05 significance, according to additional causal mediation analysis using a multi-mediator model and multiple comparison corrections (Table 3). All three metabolites were significantly positively associated with the mixture exposure (β coefficients were 0.18, 0.12, and 0.15, log_2-transformed, FDR<0.001 for all), and were significantly linked to an increased risk of incident hypertension (per SD increment HRs were 1.14, 1.09, and 1.11, FDR<0.01). The association patterns between metabolites and the mixture, as well as the volcano plot for hypertension risk, are illustrated in Figures 3 and 4

Table 3 Associations and Mediating Effects of Key Metabolic Intermediates

Metabolite	Association with mixture index β (95% CI)	HR associated with hypertension (95% CI)	Indirect Effect (NIE) β (95% CI)†	Mediation proportion % (95% CI)
Hexanoyl Carnitine (C6-carnitine)	0.18 (0.13, 0.23)	1.14 (1.08, 1.21)	0.025 (0.014, 0.038)	15.2 (8.5, 22.0)
Tryptophan	0.12 (0.07, 0.17)	1.09 (1.03, 1.15)	0.012 (0.005, 0.020)	7.5 (3.2, 11.8)
PC ae C38:4	0.15 (0.10, 0.20)	1.11 (1.05, 1.18)	0.016 (0.008, 0.025)	9.9 (5.1, 14.7)
Joint Mediating Effect	—	—	0.053 (0.039, 0.068)	32.6 (24.1, 41.1)

HR represents the risk ratio for each 1 standard deviation increase in metabolite levels, with the model adjusting for the same covariates as in Table 2.

The indirect effect (NIE) is expressed on the log(HR) scale and estimated using a weighted multiple mediation survival model; the joint mediation effect is the total of the combined indirect effects of the three.

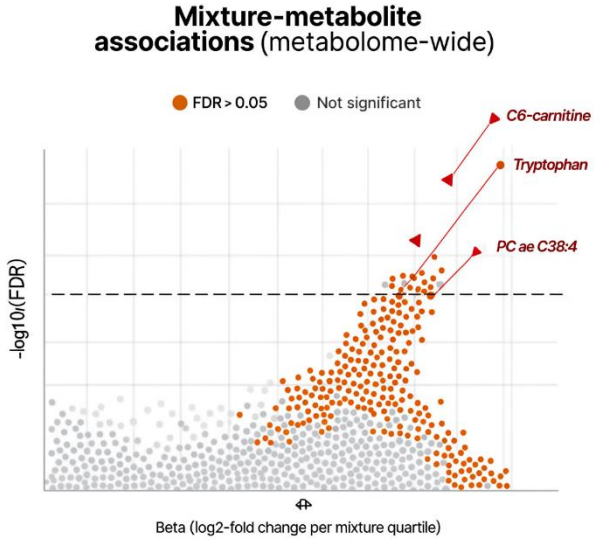

Figure 3: Volcano plot: mixture-metabolite associations

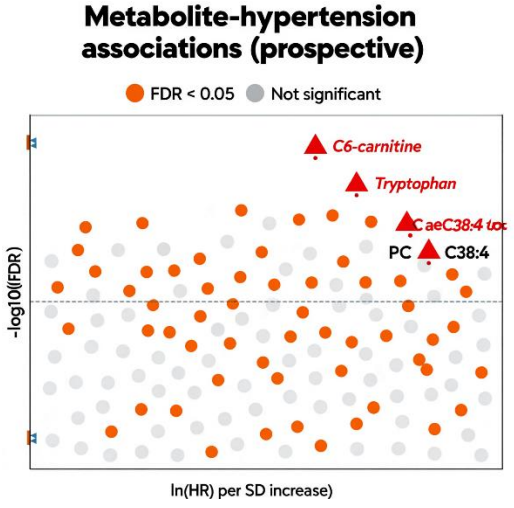

Figure 4: Volcano plot: metabolite-hypertension associations

4.4. Quantifying Indirect Effects Mediated by Metabolites

Three metabolites together mediated 32.6% (95% CI: 24.1%–41.1%) of the overall effect of air pollution mixtures on hypertension under the counterfactual causal mediation paradigm. This indicates that about one-third of the mixture's influence is conveyed through these three metabolic

pathways (Figure 5). Hexanoylcarnitine exhibited the largest independent mediation proportion (15.2%), followed by PC ae C38:4 (9.9%), whereas tryptophan had a relatively smaller but still significant mediation proportion (7.5%). After removing the connection between the three, the combined indirect impact was essentially in line with the total of the individual indirect effects, indicating that there was no clear overlapping mediation between the metabolites. About 67.4% of the mixture's impact still works through additional unmeasured channels or direct routes, according to the estimated natural direct effect of HR 1.12 (95% CI: 1.06–1.19).

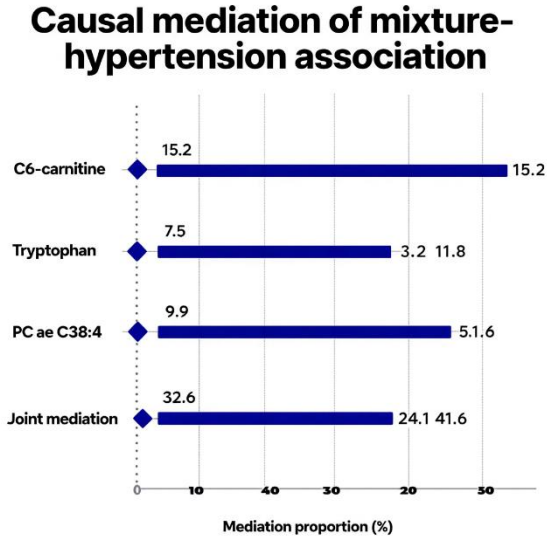

Figure 5: Mediation proportion estimates

4.5. Simulated Interventions and Population Prevention Benefits

Using a G-computation-based simulated intervention analysis, we assumed that by applying population strategies to lower carnitine levels to below the 25th percentile (P25), while keeping other conditions the same, the cumulative incidence of hypertension in this counterfactual scenario would drop from the observed 22.9% (95% CI: 21.5%–24.3%) to 19.4% (95% CI: 18.1%–20.8%). This implies that around 15.3% (95% CI: 10.8%–19.9%) of instances of hypertension linked to air pollution might be avoided. If all three metabolites are simultaneously kept within their respective P25, the total preventative benefit rises to 22.7% (95% CI: 16.4%–28.5%) (Table 4).

Table 4 Population-attributable benefits of simulated interventions on target metabolites

Intervention target	Counterfactual incidence % (95% CI)	Attributable Benefit % (95% CI)
Hexanoyl Carnitine (C6-carnitine)	19.4 (18.1, 20.8)	15.3 (10.8, 19.9)
Combined intervention of three metabolites	17.7 (16.3, 19.1)	22.7 (16.4, 28.5)

The observed 7.2-year cumulative incidence of hypertension was 22.9% (95% CI: 21.5%–24.3%). Attributable benefit = (observed incidence - counterfactual incidence) / observed incidence × 100%

4.6. Sensitivity Analysis

The robustness of the primary analysis conclusions is supported by all of the sensitivity analysis outcomes. After eliminating people who had diabetes or cardiovascular disease at baseline, the mixed HR was 1.16 (1.09–1.23), and the total mediation proportion for the three metabolites marginally decreased to 30.5% (22.0%–39.1%). The mixed HR was 1.17 (1.11–1.24) with a mediation proportion of 31.8% (23.2%–40.5%) after further accounting for green space and neighborhood noise. Using the BKMR model, we validated the nearly linear impact of the combination and observed no significant exposure-metabolite interactions. The mediation fraction for individuals who did not relocate stayed between 29% and 34% after excluding events that happened within the first two years and using various metabolite identification criteria and missing value imputation techniques. E-value analysis revealed that the associations between metabolite-exposure and metabolite-outcome would each need to be altered by at least 1.6 times in order to fully explain the indirect effect of hexanoylcarnitine by unmeasured confounders, indicating that the results are reasonably robust to unmeasured confounding.

5. Discussion

This work is the first to combine high-dimensional causal mediation analysis, serum untargeted metabolomics, and individual-level evaluations of air pollution mixes. It is based on a large-scale prospective cohort. Hexanoylcarnitine, tryptophan, and a glycerophospholipid containing arachidonic acid (PC ae C38:4) were found to be important metabolic mediators connecting air pollution mixes to the start of hypertension, accounting for about one-third of the total impact of the mixed exposure. Targeting all three metabolites might result in a benefit of up to 22.7%, and decreasing hexanoylcarnitine levels to the optimal range through population-level methods could avoid around 15% of pollution-related hypertension cases, according to simulated intervention analysis. These results quantify the population-level benefits of focusing on preventable metabolites for the first time within a counterfactual framework and offer the most thorough quantitative evidence to date for the "environmental exposure–metabolic disturbance–hypertension" causal chain. These findings have significant implications for both mechanisms and public health.

According to this study, the total effect of the air pollution combination raises the risk of hypertension by 18% per interquartile range, which is somewhat more than the effects of $PM_{2.1}$ (10%) and NO_2 (8%) in single-pollutant models. This is consistent with earlier studies that evaluated cardiovascular and metabolic outcomes using mixture analyses (such as weighted quantile sum regression or quantile g-computation), confirming that single-pollutant models may underestimate the true health effects in a setting of highly correlated co-exposures. The weight study indicated that $PM_{2.5}$ contributes the most to the mixture, which makes sense considering its biological traits: particles have greater surface areas that transport harmful components, may enter the circulation more quickly, and produce systemic oxidative stress and inflammation.

The three major mediators revealed by metabolomics screening are all supported by biological pathways independent of established risk factors. Hexanoylcarnitine (C6-carnitine) is a medium-chain acylcarnitine, an intermediary in fatty acid β-oxidation. Its buildup often suggests decreased mitochondrial oxidative activity or inadequate fatty acid oxidation. Animal studies show that particulate matter inhalation can trigger a burst of mitochondrial reactive oxygen species and damage mitochondrial respiratory chain function in the liver and vascular endothelium, thereby inhibiting complete oxidation of long-chain fatty acids and causing medium-chain acylcarnitines to spill into the circulation.By triggering pro-inflammatory signaling pathways, boosting endothelin-1 release, and decreasing nitric oxide bioavailability, hexanoylcarnitine itself can increase blood pressure and cause vasoconstriction. Similar increases in acylcarnitine profiles in cardiovascular events have been shown in earlier research, and a Mediterranean diet pattern has been shown to considerably lower these levels, underscoring their potential as an intervention target.

Changes in the activity of indoleamine 2,3-dioxygenase under inflammatory circumstances, which reroutes tryptophan into the kynurenine pathway, may be connected to the rise in tryptophan. However, increased protein catabolism or an imbalance in the gut flora can also lead to higher levels of tryptophan in the blood. Since exposure to air pollution can activate tryptophan metabolic pathways and there is epidemiological data linking tryptophan levels independently with the development of hypertension, our mediation findings fill in this gap.

The precursor for vasoactive lipid mediators such as prostacyclin and thromboxane is PC ae C38:4, an ether phospholipid with an arachidonic acid side chain at the sn-2 position. Air pollution-induced membrane lipid peroxidation and phospholipase A_2 activation can abnormally increase these glycerophospholipids, resulting in derivatives of arachidonic acid with pro-inflammatory or vasoconstrictive properties that elevate blood pressure.

All things considered, these three metabolites provide an explanation for how air pollution mixes might raise the risk of hypertension by metabolic reprogramming from three different pathways: lipid signaling, amino acid metabolism, and mitochondrial energy metabolism. The combined mediation proportion of 32.6% suggests that their cumulative effect accounts for a considerable part of the mixture's effects, while additional routes may include direct oxidative damage and autonomic nervous system malfunction.

Compared with the recent findings reported by Dai et al. [17], which identified some acylcarnitines and glycerophospholipids as mediators of cardiovascular events in single-pollutant analyses but didn’t involve a mixture framework, include tryptophan, or assess the theoretical benefits of population interventions, our study expands methodologically—using quantile g-computation to handle mixed exposures, while incorporating untargeted metabolomics screening and high-dimensional multiple mediator causal models. This not only increases the representativeness of real-world mixed exposures but also reduces biased selection induced by predefining potential mediators. Moreover, by modeling treatments to determine population-attributable benefits, our work directly integrates molecular mediation findings into public health measures, which has seldom been documented in previous environmental exposomics research.

Three features of this study demonstrate its novelty. We employed a counterfactual paradigm to break down the overall effect and constructed a full causal mediation chain of "exposure mixture—untargeted metabolomics—disease outcomes." This makes it one of the few closed-loop pieces of evidence inside the exposure-omics study paradigm. All three of the metabolites we found have modifiable characteristics: tryptophan metabolism can be impacted by anti-inflammatory interventions or gut microbiota regulation; glycerophospholipid composition is closely linked to dietary fat intake; and hexanoylcarnitine can be influenced by diet patterns (such as the Mediterranean diet or limiting refined carbohydrates), exercise, and possible medications (such as carnitine palmitoyltransferase modulators). This shows that the metabolic mediators we uncovered are not merely indicators but might possibly be targets for precision prevention. Despite the hypothetical nature of the simulated intervention results, they quantitatively demonstrate the feasibility of employing population-level metabolism optimization to mitigate health harm from environmental exposure, offering data support for multi-sector (environmental policy + personalized prevention) joint interventions.

This study also has some shortcomings. First, observational studies are still unable to fully rule out residual confounding and reverse causation, despite the adoption of a prospective design, a rigorous causal mediation framework, comprehensive sensitivity analysis, and E-value evaluations. The mediation effect may be weakened since metabolites were only tested once at baseline and their long-term fluctuations were not recorded. Although time-weighted sensitivity studies produced reliable results, the assessment of air pollution exposure was based on baseline addresses, which do not accurately reflect individual activity patterns and interior microenvironments. While the metabolomics platform covered key metabolite categories, it didn't contain all metabolites (such as extremely short-lived reactive oxygen species indicators); therefore, future integration with lipidomics, proteomics, and other omics might extend pathway coverage. The research group comprised middle-aged and older persons from a single urban location; therefore, generalizing to younger populations or places with substantially differing pollution levels should be done cautiously. Lastly, the simulated intervention scenarios assumed that metabolite levels could be safely decreased to a given population percentile, but the safe range and successful ways for such treatments still need to be proven in clinical trials and translational research.

6. Conclusion

This study, based on a large-scale prospective cohort, is the first to integrate individual-level air pollution mixture assessments with untargeted metabolomics, using high-dimensional causal mediation analysis to reveal that three metabolites—hexanoylcarnitine, tryptophan, and PC ae C38:4—act as core molecular mediators linking atmospheric pollutant mixtures to the onset of hypertension. When combined, they account for around one-third of the mixture's overall mediation impact. Controlling these important metabolites at optimal levels by population-level methods may possibly prevent between 15% and 23% of instances of pollution-related hypertension, according to a simulated intervention study. This discovery shifts the effects of air pollution on cardiovascular health from epidemiological associations to practical metabolic targets, explaining the molecular link driven by environmental factors in hypertension from the "exposure–metabolism–disease" perspective. It also provides crucial scientific evidence for cardiovascular prevention strategies that combine environmental management with targeted metabolic interventions.

Future research can expand in the following directions: verify the reproducibility and universality of these metabolic mediation pathways in populations with different pollution levels, ethnicities, and ages; conduct randomized controlled trials or pragmatic clinical trials to assess the effectiveness of dietary, exercise, or drug interventions targeting specific metabolites (like acylcarnitine levels) on blood pressure control; integrate multi-omics platforms, including proteomics, microbiomics, and epigenomics, to systematically map multidimensional molecular networks affected by exposure; develop personal wearable exposure monitoring and longitudinal repeated metabolomic measurements to capture time windows of exposure-metabolism fluctuations and dynamic mediation effects; explore the interactions between genetic susceptibility, environmental exposure, and metabolic pathways to further identify high-risk subgroups for more precise risk stratification and targeted prevention.

CRediT authorship contribution statement

Zhu Jia: Writing – review & editing, Methodology, Conceptualization.

Declaration of generative AI and AI-assisted technologies in the writing process

The free versions of Grammarly were used for spell-checking and grammar corrections. After using these tools, the author(s) reviewed and edited the content as needed and took full responsibility for the content of the publication.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of competing interests

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

Acknowledgment

We thank the reviewer for their knowledgeable contributions to the manuscript.

REFERENCES

[1] GBD 2019 Risk Factors Collaborators. Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet. 2020;396(10258):1223-1249. DOI: 10.1016/S0140-6736(20)30752-2

[2] NCD Risk Factor Collaboration (NCD-RisC). Worldwide trends in hypertension prevalence and progress in treatment and control from 1990 to 2019: a pooled analysis of 1201 population-representative studies with 104 million participants. Lancet. 2021;398(10304):957-980. DOI: 10.1016/S0140-6736(21)01330-1

[3] Brook RD, Rajagopalan S, Pope CA 3rd, et al. Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the American Heart Association. Circulation. 2010;121(21):2331-2378. DOI: 10.1161/CIR.0b013e3181dbece1

[4] Cai Y, Zhang B, Ke W, et al. Associations of Short-Term and Long-Term Exposure to Ambient Air Pollutants With Hypertension: A Systematic Review and Meta-Analysis. Hypertension. 2016;68(1):62-70. DOI: 10.1161/HYPERTENSIONAHA.116.07218

[5] Chen R, Yin P, Meng X, et al. Fine particulate air pollution and daily mortality. A nationwide analysis in 272 Chinese cities. Am J Respir Crit Care Med. 2017;196(1):73-81. DOI: 10.1164/rccm.201609-1862OC

[6] Stafoggia M, Breitner S, Hampel R, et al. Statistical Approaches to Address Multi-Pollutant Mixtures and Multiple Exposures: the State of the Science. Curr Environ Health Rep. 2017;4(4):481-490. DOI: 10.1007/s40572-017-0162-z

[7] Bobb JF, Valeri L, Claus Henn B, et al. Bayesian kernel machine regression for estimating the health effects of multi-pollutant mixtures. Biostatistics. 2015;16(3):493-508. DOI: 10.1093/biostatistics/kxu058

[8] Keil AP, Buckley JP, O'Brien KM, et al. A quantile-based g-computation approach to addressing the effects of exposure mixtures. Environ Health Perspect. 2020;128(4):47004. DOI: 10.1289/EHP5838

[9] Zhang Z, Laden F, Forman JP, et al. Long-term exposure to particulate matter and self-reported hypertension: a prospective analysis in the Nurses' Health Study. Environ Health Perspect. 2016;124(9):1414-1420. DOI: 10.1289/EHP163

[10] Wild CP. Complementing the genome with an "exposome": the outstanding challenge of environmental exposure measurement in molecular epidemiology. Cancer Epidemiol Biomarkers Prev. 2005;14(8):1847-1850. DOI: 10.1158/1055-9965.EPI-05-0456

[11] Nassan FL, Kelly RS, Kosheleva A, Koutrakis P, Vokonas PS, Lasky-Su JA, Schwartz JD. Metabolomic signatures of the long-term exposure to air pollution and temperature. Environ Health. 2021 Jan 7;20(1):3. doi: 10.1186/s12940-020-00683-x.

[12] Liang D, Moutinho JL, Golan R, Yu T, Ladva CN, Niedzwiecki M, Walker DI, Sarnat SE, Chang HH, Greenwald R, Jones DP, Russell AG, Sarnat JA. Use of high-resolution metabolomics for the identification of metabolic signals associated with traffic-related air pollution. Environ Int. 2018 Nov;120:145-154. doi: 10.1016/j.envint.2018.07.044.

[13] Nassan FL, Kelly RS, Kosheleva A, Koutrakis P, Vokonas PS, Lasky-Su JA, Schwartz JD. Metabolomic signatures of the long-term exposure to air pollution and temperature. Environ Health. 2021 Jan 7;20(1):3. doi: 10.1186/s12940-020-00683-x.

[14] Guasch-Ferré M, Toledo E, Li J, et al. Plasma metabolites from choline pathway and risk of cardiovascular disease in the PREDIMED study. J Am Heart Assoc. 2017;6(11):e006524. DOI: 10.1161/JAHA.117.006524

[15] Zheng Y, Yu B, Alexander D, et al. Metabolomics and incident hypertension among blacks: the atherosclerosis risk in communities study. Hypertension. 2013;62(2):398-403. DOI: 10.1161/HYPERTENSIONAHA.113.01166

[16] Cheng S, Rhee EP, Larson MG, et al. Metabolite profiling identifies pathways associated with metabolic risk in humans. Circulation. 2012;125(18):2222-2231. DOI: 10.1161/CIRCULATIONAHA.111.067827

[17] Cristina Menni, Sarah J. Metrustry, Robert P. Mohney, Sean Beevers, Ben Barratt, Tim D. Spector, Frank J. Kelly, Ana M. Valdes, Circulating Levels of Antioxidant Vitamins Correlate with Better Lung Function and Reduced Exposure to Ambient Pollution, American Journal of Respiratory and Critical Care

Medicine, Volume 191, Issue 10, May 2015, Pages 1203–1207, https://doi.org/10.1164/rccm.201411-2059LE

[18] Valeri L, VanderWeele TJ. Mediation analysis allowing for exposure-mediator interactions and causal interpretation: theoretical assumptions and implementation with SAS and SPSS macros. Psychol Methods. 2013;18(2):137-150. DOI: 10.1037/a0031034

[19] Tang H, Huang J, Lin H, Zhang X, Yang Q, Luo N, Lin M, Tian C, Wu S, Hong J, Wen J, Jiang L, Chen P, Chen X, Tang J, Zhang Y, Yi K, Tan X, Chen Y. The global burden and biomarkers of cardiovascular disease attributable to ambient particulate matter pollution. J Transl Med. 2025 Mar 22;23(1):359. doi: 10.1186/s12967-025-06375-9.

[20] Münzel T, Gori T, Al-Kindi S, et al. Effects of gaseous and solid constituents of air pollution on endothelial function. Eur Heart J. 2018;39(38):3543-3550. DOI: 10.1093/eurheartj/ehy481