scientific reports

OPEN



Association between exposure to volatile organic compounds and atherogenic index of plasma in NHANES 2011–2018

Lidian Meng, Yuqing Wang, Xisheng Wang, Menghui Mu & He Zheng

Volatile organic compounds (VOCs) are prevalent in daily life, yet the relationship between VOCs exposure and the atherogenic index of plasma (AIP) remains inadequately explored, especially in populations with high levels of exposure. This study aims to investigate the non-linear association between VOCs exposure and AIP in the U.S. adult population. Data from the National Health and Nutrition Examination Survey (NHANES) collected between 2011 and 2018 were analyzed. A range of statistical techniques, including Spearman's correlation analysis, weighted quantile sum (WQS), multivariate logistic regression, restricted cubic splines (RCS), stratified threshold models, and bayesian kernel machine regression (BKMR), were systematically employed to assess the relationship between high-dose VOCs exposure and AIP in U.S. adults. The study included 6,027 participants, with an average age of 37 (18-59), and 50.46% were male. Of these, 3,011 had elevated AIP levels. The Mann-Whitney U test compared VOCs exposure across quartiles (Q1–Q4). Spearman models revealed strong joint exposure effects between VOCs like IPMA3 and HMPMA (ρ = 0.97). WQS regression showed a positive association between VOCs and total cholesterol (TC) (β = 5.45, 95% Cl = 5.42–5.58, P = 0.04) and high-density lipoprotein cholesterol (HDL-C) (6=1.07, 95% CI=1.03-1.10, P=0.02). After adjusting for confounders, logistic regression revealed that VOCs such as 3-4MHA, 34DMA, AAMA, ATCA, CYMA, HEMA, and SBMA were linked to higher AIP. RCS analysis indicated a nonlinear association between VOCs and AIP. Stratified modeling found that ATCA was significantly and positively associated with AIP (OR = 1.60, 95% Cl = 1.20-2.14, p < 0.01), and that when ATCA levels exceeded 128.60 ng/mL, there was a 60% increased risk of elevated AIP. Higher urinary VOCs levels, particularly ATCA, are significantly associated with increased AIP, offering new insights into the potential link between VOCs exposure and cardiovascular disease.

Keywords Volatile organic compounds, Atherogenic index of plasma, Atherosclerosis, Weighted quantile sum regression, Bayesian kernel machine regression, NHANES

Abbreviations

VOCs	Volatile organic compounds
AIP	Atherogenic index of plasma
NHANES	National Health and Nutrition Examination Survey
WQS	Weighted Quantile Sum
RCS	Restricted cubic spline
BKMR	Bayesian kernel machine regression
2HPMA	N-acetyl-S-(2-hydroxypropyl)-l-cysteine
2MHA	2-methylhippuric acid
3-4MHAL	3- and 4-methylhippuric acid
3HPMA	N-acetyl-S-(3-hydroxypropyl)-l-cysteine
AAMA	N-acetyl-S-(2-carbamoylethyl)-l-cysteine
AMCC	N-acetyl-S-(N-methylcarbamoyl)-l-cysteine
ATCA	2-aminothiazoline-4-carboxylic acid
BPMA	N-acetyl-S-(n-propyl)-l-cysteine

School of Physical Education, Huaibei Normal University, Huaibei 23500, Anhui, China. [⊠]email: 303820823@qq.com

CEMA	N-acetyl-S-(2-carboxyethyl)-l-cysteine
CYMA	N-acetyl-S-(2-cyanoethyl)-l-cysteine
DHBMA	N-acetyl-S-(3,4-dihydroxybutyl)-l-cysteine
GAMA	N-acetyl-S-(2-carbamoyl-2-hydroxyethyl)-l-cysteine
HEMA	N-acetyl-S-(2-hydroxyethyl)-l-cysteine
HMPMA	N-acetyl-S-(3-hydroxypropyl-1-methyl)-l-cysteine
MA	Mandelic acid
MHBMA3	N-acetyl-S-(4-hydroxy-2-butenyl)-l-cysteine
PGA	Phenylglyoxylic acid
PHEMA	N-acetyl-S-(phenyl-2-hydroxyethyl)-l-cysteine
SBMA	N-acetyl-S-(benzyl)-l-cysteine
TTCA	2-Thioxothiazolidine-4-carboxylic acid
BMI	Body Mass Index
TC	Total Cholesterol
TG	Triglyceride
LDL-C	Low Density Lipoprotein -Cholesterol
HDL-C	High Density Lipoprotein -Cholesterol

Cardiovascular disease (CVD) remains one of the leading causes of death worldwide, claiming approximately 17.92 million lives annually¹. Atherosclerosis, a major contributor to CVD, involves the buildup of lipids and other substances within the inner walls of arteries. This accumulation, along with the proliferation of smooth muscle cells and collagen fibers, forms plaques that harden the arterial walls². The pathogenesis of atherosclerosis is complex, with established risk factors including hyperlipidemia, hypertension, smoking, diabetes, obesity, immune damage, and genetics³. Of particular interest is the association between elevated plasma triglycerides (TG) and an increased risk of atherosclerosis⁴.

Volatile organic compounds (VOCs) are a common class of air pollutants, primarily released through human activities and environmental processes. VOCs are known to possess mutagenic, neurotoxic, genotoxic, and carcinogenic properties⁵, and have been associated with various chronic diseases, including cardiovascular disease⁶, bronchitis⁷, and arthritis⁸. Jing et al. reported that exposure to benzene, ethylbenzene, and o-xylene significantly increased the incidence of nonfatal cardiovascular and cerebrovascular events in U.S. adults⁹. Exposure to VOCs affects the cardiovascular system through multiple mechanisms. Firstly, VOCs enter the body via the respiratory tract and quickly come into contact with the bloodstream, triggering oxidative stress responses that generate free radicals, leading to endothelial cell damage and impaired vascular function¹⁰. Endothelial dysfunction is an early pathological change in atherosclerosis, and oxidative stress is considered a key factor driving its progression. Additionally, VOCs can activate the immune system, increase inflammatory responses, and promote arterial wall thickening and plaque formation¹¹. VOCs exposure also affects lipid metabolism, with long-term exposure to chemicals such as benzene, toluene, and ethylbenzene significantly increasing blood levels of triglycerides and low-density lipoprotein cholesterol (LDL-C)⁸. These lipid metabolism abnormalities may accelerate the progression of atherosclerosis, further exacerbating the risk of cardiovascular disease (CVD)^{6,12} (Fig. 1).

Recent epidemiological studies have identified links between VOCs exposure and an increased risk of cardiovascular disease¹³. For example, Chen et al. found that exposure to VOCs is associated with serum lipid levels, such as total cholesterol and triglycerides¹⁴. Dong et al. found that both single and mixed exposures to VOCs are positively correlated with metabolic syndrome¹⁵. Han et al. found a significant positive correlation between VOCs exposure and cardiovascular diseases, suggesting that VOCs may be an important risk factor for CVD¹⁶. However, the VOCs metabolites included in these studies were not comprehensive, and they only considered the effects of single or mixed exposures. Therefore, this study thoroughly examined 20 urinary VOCs metabolites and also determined the dose-response relationship of certain key VOCs exposures. The choice



Fig. 1. Exposure to volatile organic compounds induces atherosclerosis in humans a research flowchart.

of VOCs metabolites in urine as biomarkers is based on several advantages. VOCs metabolites in urine can persist for an extended period, allowing for easier long-term exposure monitoring and offering better stability compared to the rapid clearance of blood samples¹⁷. VOCs are widely present in the air, industrial emissions, and tobacco smoke, and after inhalation or absorption, they are excreted through urine, directly reflecting individual exposure levels¹⁸. Additionally, urine collection is non-invasive, simple, and well-suited for large-scale epidemiological studies, making it convenient for sample storage and transport⁵.

Methods

Study population

This study utilized data from the National Health and Nutrition Examination Survey (NHANES) spanning seven consecutive years (2011–2012, 2013–2014, 2015–2016, and 2017–2018). The specific variable names in the database are as follows: Demographics data includes gender (RIAGENDR), age (RIDAGEYR), race (RIDRETH1), education level (DMDEDUC2), marital status (DMDMARTL), and ratio of family income to poverty (INDFMPIR); Examination data includes systolic (BPXSY1), diastolic (BPXD11), and body mass index (BMXBMI); Laboratory data includes high-density lipoprotein (LBDHDDSI), low-density lipoprotein (LBDLDLSI), triglycerides (LBDTRSI), total cholesterol (LBDTCSI), cotinine (LBXCOT), hydroxycotinine (LBXHCT), and the 20 urinary volatile organic compound metabolites (URXHP2, URX2MH, URX34M, URXDHB, URXHPM, URXAAM, URXAMC, URXATC, URXBPM, URXCEM, URXCYM, URXGAM, URXHEM, URXPMM, URXMB3, URXMAD, URXPHG, URXPHE, URXBMA, URXTTC); Questionnaire data includes alcohol use (ALQ130), physical activity (PAQ605), and cigarette smoking (SMQ040). The dataset comprised a total of 44,270 participants. Initial exclusion of 15,664 participants due to missing or untested serum data reduced the sample to 28,606. An additional 22,579 participants were excluded due to missing or undetected volatile organic compounds (VOCs), leaving a final study population of 6,027 participant (Fig. 2).

The National Health and Nutrition Examination Survey (NHANES), conducted biennially by the National Center for Health Statistics (NCHS), provides comprehensive data on the health and nutritional status of children and adults in the United States. The survey utilizes stratified multistage sampling, encompassing demographic, socioeconomic, dietary, and health-related information, as well as physical exams, including medical and laboratory tests¹⁹. The study was approved by the Ethics Review Board of the National Center for Health Statistics (Approval No. Protocol #2011-17, Protocol #2018-01). All participants provided written informed consent, and the study adhered to the principles outlined in the Declaration of Helsinki²⁰.

Detection of urinary VOCs metabolites

Urine samples were collected, processed, stored, and transported to the National Center for Environmental Health and the Centers for Disease Control and Prevention's Division of Laboratory Sciences in Atlanta, GA, for analysis. The detection of VOCs metabolites was performed using ultra-performance liquid chromatographyelectrospray ionization tandem mass spectrometry (UPLC-ESI/MSMS)²¹. Use UPLC^{*} HSS T3 column (1.8 μ m × 2.1 mm × 150 mm, Part No. 186,003,540, Waters Inc.) was used for separation with a mobile phase of 15 mM ammonium acetate and acetonitrile. The eluate was ionized using an electrospray interface to produce negative ions, which were then analyzed by mass spectrometry. Metabolite concentrations were determined by



Fig. 2. Screening flowchart for inclusion in the population (n = 6027).

comparing the relative response factors (ratios of natural analyte to stable isotope-labeled internal standards) to known standard concentrations.

Twenty VOCs metabolites were measured in this study, including: N-acetyl-S-(2-hydroxypropyl)-Lcysteine (2HPMA), 2-methylhippuric acid (2MHA), 3- and 4-methylhippuric acid (3-4MHA), N-acetyl-S-(3-hydroxypropyl)-L-cysteine (3HPMA), N-acetyl-S-(2-carbamoylethyl)-L-cysteine (AAMA), N-acetyl-S-(N-methylcarbamoyl)-L-cysteine (AMCC), 2-aminothiazoline-4-carboxylic acid (ATCA), N-acetyl-S-(n-propyl)-L-cysteine (BPMA), N-acetyl-S-(2-carboxyethyl)-L-cysteine (CEMA), N-acetyl-S-(2-cyanoethyl)-L-cysteine (CYMA), N-acetyl-S-(3,4-dihydroxybutyl)-L-cysteine (DHBMA), N-acetyl-S-(2-carbamoyl-2-hydroxyethyl)-L-cysteine (GAMA), N-acetyl-S-(2-hydroxyethyl)-L-cysteine (HEMA), N-acetyl-S-(3-hydroxypropyl-1-methyl)-L-cysteine (HMPMA), mandelic acid (MA), N-acetyl-S-(4-hydroxy-2butenyl)-L-cysteine (MHBMA3), phenylglyoxylic acid (PGA), N-acetyl-S-(phenyl-2-hydroxyethyl)-L-cysteine (PHEMA), N-acetyl-S-(benzyl)-L-cysteine (SBMA), and 2-thioxothiazolidine-4-carboxylic acid (TTCA). Detailed information on the metabolites is provided in Table S1. Further experimental details are available in the NHANES laboratory data files (https://www.cdc.gov/nchs/nhanes/index.htm).

Atherosclerosis risk assessment

The atherogenic index of plasma (AIP) was selected for evaluating atherosclerosis risk in this study due to its simplicity and strong association with cardiovascular outcomes. AIP, defined as the logarithmic ratio of triglycerides (TG) to high-density lipoprotein cholesterol (HDL-C) (Eq. 1), has been shown to be a more effective predictor of atherosclerotic risk than individual lipid components like total cholesterol or LDL- C^{22} . Elevated AIP values are linked to a higher risk of atherosclerosis and cardiovascular diseases, while lower values suggest a more favorable lipid profile. In addition, AIP has been widely used in clinical and epidemiological studies for its ability to predict atherosclerotic progression. These characteristics make AIP a reliable and clinically relevant marker for this study's risk assessment.

$$AIP = \log \frac{TG}{HDL - C} \tag{1}$$

Serum samples were collected, processed, stored, and transported to the Advanced Research and Diagnostic Laboratory (ARDL) at the University of Minnesota for analysis. Triglyceride levels were measured using a Roche Cobas 600 chemistry analyzer (USA 46250-0446)²³. This method employs microbial lipoprotein lipase to rapidly hydrolyze triglycerides into glycerol, which is then oxidized into dihydroxyacetone phosphate and hydrogen peroxide. The hydrogen peroxide reacts with 4-aminophenazone and 4-chlorophenol in the presence of peroxidase, producing a red dye (Trinder's reaction), the intensity of which is proportional to the triglyceride concentration and is measured photometrically. For further experimental details, please refer to the NHANES laboratory data file (https://www.cdc.gov/nchs/nhanes/index.htm).

Other covariates

In addition to VOCs exposure, several covariates known to influence atherosclerosis risk were considered in this study. These included sex (male, female), race/ethnicity (Mexican American, other Hispanic, non-Hispanic white, non-Hispanic black, other races), education level (incomplete elementary school, grades 9–11, high school graduate, AA degree, bachelor's degree or higher), and marital status (married, widowed, divorced, unmarried, living with a partner). Other factors, including alcohol consumption (defined as the average number of alcoholic drinks consumed per day over the past 12 months, where one drink is equivalent to a 12 oz. serving of beer, a 5 oz. glass of wine, or 1.5 oz. of distilled spirits), smoking status (every day, some days, not at all), and physical activity (yes, no), were also considered. Socioeconomic factors were assessed using the household poverty-to-income ratio (PIR), categorized as low (PIR \leq 1), medium (1 < PIR \leq 3), and high (PIR > 3). Body mass index (BMI) was categorized as normal (18.5–23.9 kg/m²), overweight (25.0–29.9 kg/m²), or obese (\geq 30.0 kg/m²). Blood pressure (systolic and diastolic) and cotinine levels, a marker for tobacco exposure, were also measured. These covariates were included in the analysis to control for their potential effects on atherosclerosis risk.

Statistical analysis

Statistical analyses were performed using appropriate methods for the data types. Continuous variables were assessed using the Kruskal-Wallis test, with results presented as median and interquartile range [M (P25, P75)], as this non-parametric test is suitable for comparing distributions across multiple independent groups without assuming a normal distribution. Categorical variables were analyzed with the Chi-square test, presented as frequencies and percentages [n (%)], to test associations between categorical variables.

To explore potential co-exposure effects of volatile organic compounds (VOCs), Spearman rank correlation was employed, as it is ideal for assessing monotonic relationships between ordinal or non-normally distributed continuous variables²⁴. Next, weighted quantile sum (WQS) regression models were applied to evaluate the combined exposure effects of the 20 VOCs across the groups. WQS regression is particularly useful for assessing the joint effects of correlated exposures, enabling the identification of their collective impact on health outcomes while accounting for inter-exposure correlations²⁵. The association between VOCs and AIP was analyzed using multi-model logistic regression. In Model 1, odds ratios were calculated using univariate logistic regression as the baseline to assess the unadjusted effect of VOCs on AIP. Model 2 adjusted for potential confounders, including sex, age, race, marital status, education level, and poverty-to-income ratio, factors known to influence health outcomes. Model 3 further adjusted for physical activity, systolic and diastolic blood pressure, cigarette smoking, alcohol consumption, and cotinine levels, based on Model 2, to account for behavioral and physiological factors that may confound the VOCs-AIP relationship. Nonlinear relationships between VOCs and AIP were

examined using restricted cubic spline (RCS) analysis, which is effective for modeling nonlinear associations and uncovering more complex exposure-outcome relationships that may not follow a simple linear trend²⁶. Finally, Bayesian kernel machine regression (BKMR) was employed to evaluate the overall effect of the 20 VOCs in urine on AIP, utilizing 10,000 Markov chain Monte Carlo (MCMC) iterations²⁷. BKMR is an advanced statistical method that enables flexible modeling of complex, high-dimensional relationships between multiple exposures and outcomes, particularly when interactions between exposures are uncertain or non-linear. This method provides robust estimates of the exposure-response relationship, while accounting for uncertainty and correlation between exposures.

Statistical analyses were performed by SPSS (version 27.00, https://www.ibm.com/spss, The IMB), R language and R package (version 4.4.1, http://www.R-project.org, The R Foundation) and OriginPro (version 2024b, www .originlab.com, The OriginLab). Differences were considered statistically significant at a two-sided significance level of P < 0.05.

Results

Demographic characteristics of the study population

Participants were divided into quartiles based on the atherosclerotic index of plasma (AIP): Q1 (0.11–0.43), Q2 (0.43–0.52), Q3 (0.52–0.63), and Q4 (0.63–1.17). Among the 6,027 participants in the study, the average age was 37 years (ranging from 18 to 59 years), with 3,041 (50.46%) identifying as male and 2,986 (49.54%) as female. Compared to those in Q1, participants in Q4 tended to be older, had a higher proportion of males, and were more likely to identify as non-Hispanic white and married. Additionally, individuals in Q4 exhibited lower education levels, higher rates of smoking and alcohol consumption, and less physical activity. They also had a higher body fat percentage and blood pressure, along with elevated levels of total cholesterol (TC), triglycerides (TG), and low-density lipoprotein cholesterol (LDL-C), while showing lower levels of high-density lipoprotein cholesterol (HDL-C) (P<0.01). However, poverty rates did not differ significantly among the groups (P>0.05) (Table 1).

Spearman correlation modeling between 20 VOCs

A Spearman correlation model was employed to illustrate the co-exposure effects among volatile organic compounds (VOCs) through heat maps. The analysis revealed strong correlations between several VOCs pairs: IPMA3 and HMPMA (ρ =0.97), IPMA3 and 3HPMA (ρ =0.97), and HMPMA and 3HPMA (ρ =0.96). Additionally, other notable correlations included 3-4MHA and 2MHA (ρ =0.96), PGA and MA (ρ =0.96), CEMA and 3HPMA (ρ =0.95), GAMA and AAMA (ρ =0.95), HMPMA and CEMA (ρ =0.94), as well as IPMA3 and CYMA (ρ =0.94). These findings indicate that many VOCs are interrelated, suggesting potential common sources of pollution that may affect human health. Furthermore, the co-exposure effects of VOCs were found to be generally consistent across the four subgroups (Fig. 3).

Differences in exposure dose of VOCs across the four subgroups

After transforming the doses of volatile organic compounds (VOCs) into logarithmic form, the differences in exposure doses across the four groups (Q1, Q2, Q3, and Q4) were assessed using the Kruskal-Wallis test (KW). The results indicated significant differences (P<0.01) in 17 VOCs, including 2MHA, 34DMA, 3HPMA, AAMA, AMCC, ATCA, BPMA, CEMA, CYMA, HEMA, HMPMA, IPMA3, MA, PGA, PHEMA, 3-4MHA, and 2HPMA. However, no significant differences were observed for GAMA (P=0.064), SBMA (P=0.089), and TTCA (P=0.136) (Fig. 4).

WQS regression model of VOCs and AIP

A weighted quantile sum (WQS) regression model was developed using the R package "gWQS" to assess the combined effects of VOCs exposure on the atherogenic index of plasma (AIP), as illustrated in Fig. 4. The analysis indicated that participants in the Q4 group exhibited higher VOCs levels, demonstrating a positive correlation between VOCs and total cholesterol (TC) (β =5.45, 95% CI=5.42–5.58, P=0.04) as well as high-density lipoprotein cholesterol (HDL-C) (β =1.07, 95% CI=1.03–1.10, P=0.02) (Fig. 5).

Within the TC analysis, the metabolites AMCC (w=0.40), TTCA (w=0.12), and 2MHA (w=0.11) were identified as having the most significant weights. In the case of HDL-C, TTCA (w=0.27), SBMA (w=0.26), and ATCA (w=0.21) also exhibited relatively high weights. Additionally, in the other three subgroups, TTCA and AMCC maintained their significance (Table S2).

Weighted multi-model logistic regression model for VOCs and AIP

The analysis of urinary volatile organic compounds (VOCs) and their association with the atherogenic index (AIP) was presented in Model 2, which adjusted for demographic variables, and Model 3, which utilized stepwise regression modeling (Fig. 6a-c).

After accounting for potential confounders such as sex, age, race, marital status, education, poverty-to-income ratio, physical activity, blood pressure levels, smoking status, alcohol consumption, and cotinine levels, several VOCs demonstrated a significant correlation with AIP. Specifically, 3-4MHA (OR=1.36, 95% CI=1.05-1.75, P=0.02), 34DMA (OR=1.93, 95% CI=1.22-3.05, P<0.01), AAMA (OR=1.67, 95% CI=1.24-2.26, P<0.01), ATCA (OR=1.47, 95% CI=1.26-1.73, P<0.01), CYMA (OR=1.30, 95% CI=1.21-1.51, P<0.01), HEMA (OR=1.55, 95% CI=1.21-1.98, P<0.01), and SBMA (OR=1.43, 95% CI=1.21-1.70, P<0.01) were significantly associated with AIP. Conversely, although 2MHA, AMCC, BPMA, CEMA, and IPMA showed some statistical significance, their odds ratios were less than 1 (Table S3).

The dose-response relationship between specific VOCs and AIP was further analyzed using restricted cubic splines (RCS) (Fig. 6d). The analysis revealed that the inflection point for ATCA was 128.60 ng/mL. Above

Variables	Q1 (<i>n</i> =1503)	Q2 (<i>n</i> =1513)	Q3 (<i>n</i> =1508)	Q4 (n=1503)	Total (<i>n</i> = 6027)	P-value
Age (years)	23 (12, 52)	30 (15, 57)	42 (22, 61)	46 (32, 60)	37 (18, 59)	< 0.001**
Sex, n (%)						< 0.001**
Men	702 (46.71)	657 (43.42)	753 (49.93)	929 (61.81)	3041 (50.46)	
Women	801 (53.29)	856 (56.58)	755 (50.07)	574 (38.19)	2986 (49.54)	
Race, n (%)						< 0.001**
Mexican American	212 (14.11)	219 (14.47)	245 (16.25)	230 (15.3)	906 (15.03)	
Other Hispanic	106 (7.05)	139 (9.19)	135 (8.95)	198 (13.17)	578 (9.59)	
Non-Hispanic White	504 (33.53)	493 (32.58)	506 (33.55)	555 (36.93)	2058 (34.15)	
Non-Hispanic Black	447 (29.74)	384 (25.38)	338 (22.41)	265 (17.63)	1434 (23.79)	
Other Race	234 (15.57)	278 (18.37)	284 (18.83)	255 (16.97)	1051 (17.44)	
Education, n (%)						< 0.001**
Less than 9th grade	46 (3.06)	73 (4.82)	109 (7.23)	161 (10.71)	389 (6.45)	
9-11th grade	98 (6.52)	101 (6.68)	145 (9.62)	185 (12.31)	529 (8.78)	
High school graduate	180 (11.98)	212 (14.01)	251 (16.64)	325 (21.62)	968 (16.06)	
AA degree	940 (62.54)	849 (56.11)	720 (47.75)	551 (36.66)	3060 (50.77)	
College graduate or above	239 (15.9)	278 (18.37)	283 (18.77)	281 (18.70)	1081 (17.94)	
Marital status, n (%)						< 0.001**
Married	355 (23.62)	469 (31)	626 (41.51)	714 (47.5)	2164 (35.91)	
Widowed	68 (4.52)	85 (5.62)	80 (5.31)	77 (5.12)	310 (5.14)	
Divorced	92 (6.12)	103 (6.81)	115 (7.63)	161 (10.71)	471 (7.81)	
Never married	918 (61.08)	787 (52.02)	599 (39.72)	428 (28.48)	2732 (45.33)	
Living with partner	70 (4.66)	69 (4.56)	88 (5.84)	123 (8.18)	350 (5.81)	
Drinking (cup)	2 (1, 3)	2 (1, 3)	2 (1, 3)	2 (1, 4)	2 (1, 3)	0.008**
Smoking, n (%)						< 0.001**
Every day	451 (30.01)	497 (32.85)	559 (37.07)	671 (44.64)	2178 (36.14)	
Some days	41 (2.73)	45 (2.97)	44 (2.92)	41 (2.73)	171 (2.84)	
Not at all	1011 (67.27)	971 (64.18)	905 (60.01)	791 (52.63)	3678 (61.03)	
Vigorous work activity, n (%)						< 0.001**
Yes	585 (38.92)	549 (36.29)	426 (28.25)	365 (24.28)	1925 (31.94)	
No	917 (61.01)	964 (63.71)	1082 (71.75)	1139 (75.78)	4102 (68.06)	
Poverty to income ratio	2.01 (1.04, 4.03)	2.18 (1.08, 4.07)	1.82 (1.06, 3.65)	1.71 (0.97, 3.40)	1.88 (1.02, 3.82)	0.779
BMI (Kg/m ²)	23.3 (20.5, 27.2)	25.60 (22.10, 30.10)	28.20 (24.70, 33.50)	29.80 (26.40, 34.98)	26.20 (21.75, 31.30)	< 0.001**
Systolic pressure (mmHg)	114 (104, 126)	116 (106, 128)	122 (112, 134)	122 (112, 134)	116 (106, 130)	< 0.001**
Diastolic pressure (mmHg)	66 (58, 74)	68 (60, 74)	70 (62, 78)	72 (66, 80)	68 (60, 76)	< 0.001**
TC (mmol/L)	3.96 (3.44, 4.53)	4.34 (3.80, 5.02)	4.78 (4.24, 5.38)	5.48 (4.84, 6.10)	4.55 (3.90, 5.30)	< 0.001**
TG (mmol/L)	0.62 (0.47, 0.84)	0.86 (0.64, 1.12)	1.14 (0.86, 1.51)	1.67 (1.23, 2.25)	1.03 (0.69, 1.52)	< 0.001**
LDL-C (mmol/L)	1.97 (1.60, 2.33)	2.48 (2.12, 2.92)	2.95 (2.53, 3.44)	3.54 (3.01, 4.14)	2.72 (2.12, 3.36)	< 0.001**
HDL-C (mmol/L)	1.68 (1.47, 1.94)	1.42 (1.27, 1.63)	1.27 (1.11, 1.45)	1.06 (0.96, 1.22)	1.32 (1.11, 1.58)	< 0.001**
Cotinine (ng/mL)	0.04 (0.01, 2.13)	0.03 (0.01, 0.25)	0.04 (0.01, 1.08)	0.05 (0.01, 79.38)	0.04 (0.01,0.81)	< 0.001**

Table 1. Baseline characteristics of the study population. ^aBMI, Body Mass Index; TC, Total Cholesterol; TG,Triglyceride; LDL-C, Low Density Lipoprotein -Cholesterol; HDL-C, High Density Lipoprotein -Cholesterol.*P < 0.05, **P < 0.01.

this level, ATCA showed a significant positive association with AIP (OR=1.60, 95% CI=1.20-2.14, P<0.01), indicating that when ATCA levels exceed 128.60 ng/mL, the risk of elevated AIP increases by 60%. Similarly, the inflection point for SBMA was 6.81 ng/mL, and above this level, the risk of elevated AIP increases by 36%. (OR=1.36, 95% CI=1.08-1.70, P<0.01). However, no significant associations were found between 34DMA or AAMA and AIP (Table 2).

Subgroup analysis

To further investigate the variables that may contribute to an increased risk of atherosclerosis, subgroup analyses were conducted. In the age group analysis, high exposure was significantly positively associated with the occurrence of events in the 30–59 years group (OR=1.55, P<0.001), while no significant effects were observed in the 18–29 years and >60 years groups (P=0.153 and P=0.370, respectively). Regarding gender differences, exposure was significantly associated with event occurrence in men (OR=1.64, P<0.001), but inversely related in women (OR=0.81, P=0.015). Racial analysis revealed significant associations between exposure and event occurrence in Mexican Americans and non-Hispanic Blacks (P<0.001), while no significant differences were

а



Fig. 3. Co-exposure effects between VOCs. (**a**) total exposure effect; (**b**) Q1 group exposure effect; (**c**) Q2 group exposure effect; (**d**) Q3 group exposure effect; (**e**) Q4 group exposure effect.

observed in other racial groups. In the body mass index (BMI) analysis, exposure was positively correlated with event occurrence in individuals with a BMI < 25 (OR=1.51, P<0.001), whereas no significant correlation was found in those with a BMI > 25. Smoking status also significantly influenced the relationship between exposure and event occurrence, with smokers showing an OR of 2.70 (P=0.002), while no significant association was observed in non-smokers (OR=1.08, P=0.525). The high exposure group in smokers showed a significantly



ig i. Detrieen group dose amerenees in voor

increased risk of atherosclerosis (OR=2.70, P=0.002), whereas no significant change was observed in nonsmokers (OR=1.08, P=0.525) (Table 3).

Robustness test

A Bayesian kernel machine regression (BKMR) model was developed using the R package "bkmr" to evaluate the combined effects of volatile organic compound (VOCs) exposure on the atherogenic index (AIP). The model analyzed the dose levels of 20 VOCs in urine in relation to total cholesterol (TC) and high-density lipoprotein cholesterol (HDL-C). The results indicated a positive correlation between VOCs exposure and TC among participants at the 50th percentile level. Conversely, a negative correlation was observed for HDL-C in the same group. Furthermore, a relationship between VOCs and AIP was confirmed, with sensitivity test results aligning with the main analysis findings (Fig. 7).

Discussion

In this study, we investigated the potential association between volatile organic compounds (VOCs) and the plasma atherogenic index (AIP), with a particular focus on the effect of 2-aminothiazoline-4-carboxylic acid (ATCA) on atherosclerosis risk. Utilizing data from the NHANES database (2011–2018), our findings indicate that long-term exposure to VOCs, especially ATCA, may elevate the risk of atherosclerosis, providing new evidence linking VOCs to cardiovascular disease.

Initially, Spearman's correlation analysis demonstrated a joint exposure effect among different VOCs, revealing a significant correlation between IPMA3 and HMPMA (ρ =0.97). This suggests that co-exposure



Fig. 5. WQS regression modeling reveals the weighting of VOCs dose in blood TG, HDL-C. (**a**) WQS regression for group Q1; (**b**) WQS regression for group Q2; (**c**) WQS regression for group Q3; (**d**) WQS regression for group Q4.

to multiple VOCs may have a synergistic impact on atherosclerosis development. Further analysis through weighted quantile sum (WQS) regression modeling identified significant positive correlations between VOCs and total cholesterol (TC) as well as high-density lipoprotein cholesterol (HDL-C), indicating their potential influence on lipid metabolism. These findings are consistent with previous studies and reinforce the hypothesis that VOCs are a risk factor for cardiovascular disease²⁸.

Notably, our stratified threshold model indicated a significant increase in the risk of elevated AIP when ATCA levels exceeded 128.60 ng/mL (OR = 1.60). This finding has important clinical implications, suggesting that ATCA may be a key risk factor for atherosclerosis. As a biomarker of cyanide metabolism, ATCA may contribute to atherogenesis through oxidative stress and inflammatory mechanisms²⁹. Prior research has established a link between cyanide exposure and cardiovascular disease and our study reveals, the specific role ATCA may play in this pathological process.



Fig. 6. Weighted Multimodal Logistic Regression Models of VOCs and AIP with RCS Analysis. (a) Model 1: Multimodal Logistic Regression of VOCs and AIP; (b) Model 2: Multimodal Logistic Regression of VOCs and AIP with additional corrections for sex, age, race, marital status, educational attainment, and poverty-toincome ratio; (c) Multimodal Logistic Regression of VOCs with AIP, additionally adjusted for physical activity, systolic blood pressure, diastolic blood pressure, smoking, alcohol consumption, and cotinine levels; (d) Nonlinear trends in 34DMA, AAMA, SBMA, ATCA.

Variable	β	S.E	Z	OR (95%CI)	P-value
34DMA < 321.75 ng/mL	-0.19	0.09	-2.05	0.83 (0.69–0.99)	0.040*
$34 DMA \!\geq\! 321.75 \text{ ng/mL}$	-0.12	0.17	-0.68	0.89 (0.64–1.24)	0.497
AAMA<50.77 ng/mL	-0.36	0.14	-2.69	0.69 (0.53-0.91)	0.007**
AAMA \geq 50.77 ng/mL	0.10	0.13	0.78	1.11 (0.86–1.43)	0.535
SBMA < 6.81 ng/mL	-0.16	0.14	-1.21	0.85 (0.65-1.11)	0.225
SBMA≥6.81 ng/mL	0.30	0.12	2.63	1.36 (1.08–1.70)	0.009**
ATCA < 128.60 ng/mL	0.18	0.12	1.44	1.20 (0.94–1.53)	0.050*
ATCA≥128.60 ng/mL	0.47	0.15	3.16	1.60 (1.20-2.14)	0.002**

Table 2. Segmented binary logistic regression model for dose-response effect. ^a β , Ratio; *S.E.*, Standard Error; *Z*, *Z*-score; *OR*, Odds ratio; *95%CI*, Odds ratio. ^{*}*P*<0.05, ^{**}*P*<0.01.

Subgroup	n (%)	Low exposure (Events/Total)	High exposure (Events/Total)	OR (95%CI)	P-value	P _{interaction}
Age (years)						< 0.001**
18-29	2514 (41.71)	649/709	1682/1805	1.26 (0.92–1.74)	0.153	
30-59	2158 (35.81)	483/943	444/1215	1.55 (0.96-1.65)	< 0.001**	
>60	1355 (22.48)	303/697	302/658	1.10 (0.89–1.37)	0.370	
Sex						< 0.001**
Men	3041 (50.46)	766/1393	1099/1648	1.64 (1.41–1.90)	< 0.001**	
Women	2986 (49.54)	669/956	1329/2030	0.81 (0.69-0.96)	0.015*	
Race						0.671
Mexican American	906 (15.03)	159/288	449/618	2.16 (1.61-2.89)	< 0.001**	
Other Hispanic	578 (9.59)	133/215	241/363	1.22 (0.86–1.73)	0.271	
Non-Hispanic White	2058 (34.15)	558/967	661/1091	1.13 (0.94–1.34)	0.184	
Non-Hispanic Black	1434 (23.79)	389/537	704/897	1.39 (1.08–1.78)	0.009**	
Other Race	1051 (17.44)	196/342	373/709	0.83 (0.64–1.07)	0.152	
Body mass index (kg/m ²)						0.432
< 25	2586 (42.91)	582/823	1383/1763	1.51 (1.25–1.82)	< 0.001	
25-30	1609 (26.70)	359/716	441/893	0.98 (0.80-1.20)	0.879	
> 30	1832 (30.39)	468/800	578/1032	0.90 (0.75-1.09)	0.285	
Smoking						0.007**
Yes	2349 (55.09)	359/1034	992/1315	2.70 (1.43-3.07)	0.002**	
No	2707 (44.91)	207/1011	410/1696	1.08 (0.86-1.35)	0.525	

Table 3. Subgroup analysis of covariates influencing atherosclerosis risk across different exposure levels. ^a*OR*, Odds ratio; 95%*CI*, Odds ratio. ^{*}*P* < 0.05, ^{**}*P* < 0.01.



Fig. 7. BKMR model fitting of urinary VOCs dose to AIP. (a) VOCs & TC; (b) VOCs & HDL-C.

While previous studies have shown that VOCs are common air pollutants associated with conditions such as diabetes³⁰, obesity³¹, decreased lung function³², and autoimmune diseases³³, the exact mechanisms through which VOCs affect cardiovascular health remain unclear³⁴. Smoking is known to increase the concentration of various VOCs in urine and blood^{35,36}, which can subsequently affect blood lipid levels³⁷. A substantial body of research has explored the relationship between smoking, VOCs metabolites, and serum lipids. Smoking has been shown to significantly impact lipid metabolism, influencing both the respiratory system and systemic metabolism through the release of harmful VOCs³⁸. Studies indicate that smoking not only increases serum levels of total cholesterol (TC) and low-density lipoprotein cholesterol (LDL-C), but is also closely associated with a decrease in high-density lipoprotein cholesterol (HDL-C) levels³⁹. Moreover, oxidative stress induced by smoking may exacerbate lipid metabolic abnormalities, thus increasing the risk of cardiovascular diseases (CVD)⁴⁰. Volatile organic compounds, as major metabolic products of smoking, have been identified as important biomarkers for revealing the metabolic effects of smoking⁴¹. Kampa and Marilena suggested that exposure to high levels of air pollutants, including VOCs, may lead to chronic inflammation, thereby increasing the incidence of atherosclerosis and other cardiovascular conditions⁴². Additionally, research by Chen et al. indicated that chronic exposure to VOCs significantly elevates plasma LDL-C levels, promoting the development of atherosclerosis¹⁴.

ATCA, as a product of cyanide metabolism, is closely related to oxidative stress and cardiovascular health. Cyanide is a well-known environmental toxin, and its detrimental effects on the cardiovascular system have been documented, particularly among smokers and industrial workers⁴³. For example, Lin et al. showed that ATCA, a biomarker of cyanide exposure, may reflect the significant effect of long-term low-dose cyanide exposure on atherogenesis⁴⁴. These studies support our finding of a positive correlation between ATCA and AIP.

This study systematically evaluated the complex association between VOCs and AIP using multiple statistical models (e.g., WQS, RCS, BKMR). The findings not only underscore the potential role of ATCA as a risk factor for atherosclerosis but also provide a theoretical basis for developing public health policies. Given the impact of environmental exposures on cardiovascular health, future research should focus on interventions to reduce VOCs exposure and their potential benefits in mitigating cardiovascular disease risk.

However, this study has several limitations. First, as a cross-sectional investigation, we were unable to establish a causal relationship between VOCs and AIP. Second, while multiple confounders were adjusted for, certain variables (e.g., diet, genetic susceptibility) were not considered, which may have influenced the results. Additionally, a single measurement of VOCs in urine may not accurately reflect long-term exposure. Future studies should aim for longer follow-up periods to validate these findings further. Furthermore, this study excluded participants for whom VOCs metabolites were not detected, which may have resulted in unclear classification of the study sample and potentially affected the generalizability of the results. Therefore, future research should consider more comprehensive sample collection and analysis to enhance the representativeness of the data.

Conclusion

In summary, elevated levels of volatile organic compounds (VOCs) in urine, particularly 2-aminothiazoline-4-carboxylic acid (ATCA), were significantly and positively correlated with an increase in the plasma atherogenic index (AIP). This suggests that higher exposure to VOCs may be associated with an increased risk of atherosclerosis, offering important insights for further investigation into the relationship between VOCs and cardiovascular diseases.

Data availability

All data have been included in the article and supplementary information, and all data are available on the open database NHANES: https://wwwn.cdc.gov/nchs/nhanes/Default.aspx.

Received: 24 September 2024; Accepted: 10 March 2025 Published online: 16 March 2025

References

- 1. Roth, G. A. et al. Global, regional, and National burden of cardiovascular diseases for 10 causes, 1990 to 2015. *J. Am. Coll. Cardiol.* **70** (1), 1–25 (2017).
- 2. Gotto, M. Jr. Triglyceride as a risk factor for coronary artery disease. Am. J. Cardiol. 82, 22-25 (1998).
- Huo, R. R. et al. Interacting and joint effects of triglyceride-glucose index (TyG) and body mass index on stroke risk and the mediating role of TyG in middle-aged and older Chinese adults: a nationwide prospective cohort study. *Cardiovasc. Diabetol.* 23 (1), 30 (2024).
- 4. Toth, P. P. Triglycerides and atherosclerosis: bringing the association into sharper focus*. J. Am. Coll. Cardiol. 77 (24), 3042–3045 (2021).
- 5. McGraw, K. E. et al. Exposure to volatile organic compounds-acrolein, 1, 3-butadiene, and crotonaldehyde-is associated with vascular dysfunction. *Environ. Res.* **196**, 110903 (2021).
- He, Y. et al. Exposure to BTEX is associated with cardiovascular disease, dyslipidemia and leukocytosis in National US population. Sci. Total Environ. 919, 170639 (2024).
- 7. Paterson, C. A. et al. Indoor PM2. 5, VOCs and asthma outcomes: A systematic review in adults and their home environments. *Environ. Res.* 202, 111631 (2021).
- 8. Ogbodo, J. et al. Volatile organic compounds: a Proinflammatory activator in autoimmune diseases. *Front. Immunol.* **13**, 928379 (2022).
- 9. Jing et al. Association of the blood levels of specific volatile organic compounds with nonfatal cardio-cerebrovascular events in US adults. *BMC Public. Health.* 24 (1), 616 (2024).
- Rehman, A. et al. Cell-free Circulating mitochondrial DNA: an emerging biomarker for airborne particulate matter associated with cardiovascular diseases. Free Radic. Biol. Med. 195, 103–120 (2023).

- 11. Riggs, D. W. et al. Environmental exposure to volatile organic compounds is associated with endothelial injury. *Toxicol. Appl. Pharmcol.* **437**, 115877 (2022).
- 12. Xu, T. et al. Gas chromatography-mass spectrometry pilot study to identify volatile organic compound biomarkers of childhood obesity with dyslipidemia in exhaled breath. J. Translational Intern. Med. 11 (1), 81–89 (2023).
- Nalini, M. et al. Volatile organic compounds and mortality from ischemic heart disease: A case-cohort study. Am. J. Prev. Cardiol. 19, 100700 (2024).
- 14. Chen, W. Y. et al. The association between exposure to volatile organic compounds and serum lipids in the US adult population. *Lipids Health Dis.* **22** (1), 129 (2023).
- 15. Dong, R. et al. Association of volatile organic compound exposure with metabolic syndrome and its components: a nationwide cross-sectional study. *BMC Public. Health.* **24** (1), 671 (2024).
- Han, S. et al. Associations between specific volatile organic chemical exposures and cardiovascular disease risks: insights from NHANES. Front. Public. Health. 12, 1378444 (2024).
- 17. Pal, V. et al. Diurnal variability in urinary volatile organic compound metabolites and its association with oxidative stress biomarkers. Sci. Total Environ. 818, 151704 (2022).
- Wang, B. L. et al. Unmetabolized VOCs in urine as biomarkers of low level exposure in indoor environments. J. Occup. Health. 49 (2), 104–110 (2007).
- 19. Ahluwalia, N. et al. Update on NHANES dietary data: focus on collection, release, analytical considerations, and uses to inform public policy. *Adv. Nutr.* 7 (1), 121–134 (2016).
- World Medical Association. World medical association declaration of Helsinki. Ethical principles for medical research involving human subjects. Bull. World Health Organ. 79 (4), 373 (2001).
- Alwis, K. et al. Simultaneous analysis of 28 urinary VOC metabolites using ultra high performance liquid chromatography coupled with electrospray ionization tandem mass spectrometry (UPLC-ESI/MSMS). Analytica chimica acta. 750, 152–160 (2012).
- Fernández-Macías, J. C. et al. Atherogenic index of plasma: novel predictive biomarker for cardiovascular illnesses. Arch. Med. Res. 50 (5), 285–294 (2019).
- Friedewald, W. T., Robert, I., Levy & Donald, S. Fredrickson. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. *Clin. Chem.* 18 (6), 499–502 (1972).
- 24. Zar, J. H. Significance testing of the spearman rank correlation coefficient. J. Am. Stat. Assoc. 67, 578-580 (1972).
- 25. Lee, M. J. et al. A generalized weighted quantile sum approach for analyzing correlated data in the presence of interactions. *Biom. J.* **61** (4), 934–954 (2019).
- Marrie, R., Ann, N. V., Dawson & Allan Garland. Quantile regression and restricted cubic splines are useful for exploring relationships between continuous variables. J. Clin. Epidemiol. 62 (5), 511–517 (2009).
- 27. Bobb, J. F. et al. Bayesian kernel machine regression for estimating the health effects of multi-pollutant mixtures. *Biostatistics*. **16** (3), 493–508 (2015).
- 28. Tan, L. et al. Associations of individual and mixture exposure to volatile organic compounds with metabolic syndrome and its components among US adults. *Chemosphere* 347, 140683 (2024).
- 29. Delporte, C. et al. Myeloperoxidase-catalyzed oxidation of cyanide to cyanate: a potential carbamylation route involved in the formation of atherosclerotic plaques? J. Biol. Chem. 293 (17), 6374-6386 (2018).
- Nguyen, H. et al. Environmental science and pollution research role of heavy metal concentrations and vitamin intake from food in depression: a National cross-sectional study (2009–2017). Environ. Sci. Pollut. Res. 29 (3), 4574–4586 (2022).
- 31. Mendy, A. et al. Urinary volatile organic compound metabolites and reduced lung function in US adults. *Respir. Med.* **205**, 107053 (2022).
- 32. Lei, T. et al. The exposure to volatile organic chemicals associates positively with rheumatoid arthritis: a cross-sectional study from the NHANES program. *Front. Immunol.* 14, 1098683 (2023).
- 33. Nguyen, P. et al. Liver lipid metabolism. Journal of animal physiology and animal nutrition. 92 (3) 272-283 (2008).
- 34. Zhang, Y. et al. Association between exposure to a mixture of phenols, pesticides, and phthalates and obesity: comparison of three statistical models. *Environ. Int.* **123**, 325–336 (2019).
- Jo, W. K. & Ku-Won, P. Utilization of breath analysis for exposure estimates of benzene associated with active smoking. *Environ. Res.* 83 (2), 180–187 (2000).
- 36. Van Berkel, J. J. B. N. et al. Development of accurate classification method based on the analysis of volatile organic compounds from human exhaled air. J. Chromatogr. B. 861 (1), 101–107 (2008).
- Prakash, V. et al. The smoking-dyslipidaemia dyad: A potent synergistic risk for atherosclerotic coronary artery disease. JRSM Cardiovasc. Disease. 10, 2048004020980945 (2021).
- 38. Kuang, H. et al. Volatile organic compounds from second-hand smoke May increase susceptibility of children through oxidative stress damage. *Environ. Res.* 207, 112227 (2022).
- 39. Arsenault, B. J. et al. Beyond low-density lipoprotein cholesterol: respective contributions of non-high-density lipoprotein cholesterol levels, triglycerides, and the total cholesterol/high-density lipoprotein cholesterol ratio to coronary heart disease risk in apparently healthy men and women. *J. Am. Coll. Cardiol.* **55** (1), 35–41 (2009).
- 40. Rotariu, D. et al. Oxidative stress-Complex pathological issues concerning the hallmark of cardiovascular and metabolic disorders. *Biomed. Pharmacother.* **152**, 113238 (2022).
- 41. Shen, Q. et al. A review of disrupted biological response associated with volatile organic compound exposure: insight into identification of biomarkers. *Sci. Total Environ.*, **948**, 174924. (2024).
- 42. Kampa, M. & Castanas, E. Human health effects of air pollution. Environ. Pollut. 151 (2), 362-367 (2008).
- 43. Hardy, H. L. et al. Thiocyanate effect following industrial cyanide exposure: report of two cases. N. Engl. J. Med. 242, 968–972 (1950).
- 44. Lin, C. Y. et al. Investigating the association between urinary levels of acrylonitrile metabolite N-acetyl-S-(2-cyanoethyl)-Lcysteine and the oxidative stress product 8-hydroxydeoxyguanosine in adolescents and young adults. *Environ. Pollut.* **239**, 493–498 (2018).

Acknowledgements

Thanks to all the researchers and participants of the National Health and Nutrition Examination Survey (NHANES), conducted by the National Center for Health Statistics (NCHS) of the Centers for Disease Control and Prevention (CDC). I also extend my gratitude to Linfeng Zhang (College of Humanities and Social Sciences, Xi'an Jiaotong University) for her invaluable assistance over the years.

Author contributions

L.M. Conceptualization, methodology, data analysis, visualization, writing. Y.W. data analysis, Calculate, inspect. X.W. Calculate, inspect. M.M. Calculate, inspect. H.Z. Provided guidance, funding acquisition, supervision, review.All authors reviewed the manuscript.

Declarations

Ethics approval and consent to participate

All data were obtained from NHANES, which was approved by the Ethics Review Board of the National Center for Health Statistics (Protocol #2018-01, Protocol #2018-01), all subjects signed a written letter of information, and the study process was in accordance with the Declaration of Helsinki.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/1 0.1038/s41598-025-93833-5.

Correspondence and requests for materials should be addressed to H.Z.

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

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

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

© The Author(s) 2025