scientific reports

OPEN

Impact of gene-environment interactions on atrial fibrillation and cardiac structure

Panlong Li[1], Xirui Zhu[1], Min Liu[4], Yanfeng Wang[1], Chun Huang[1], Junwei Sun[1], Shan Tian[5], Yuna Li[5], Yuan Qiao[5], Junting Yang[5], Shanshan Cao[5], Chaohua Cong[5], Lei Zhao[5], Meiyun Wang[2,3]✉, Jingjing Su[5]✉ & Dandan Tian[4]✉

Environmental pollution is a major burden of cardiovascular disease. The aim of the study was to investigate the interactions between combined environmental factors and genetic susceptibility on atrial fibrillation (AF) and cardiac structures. The study included 374,495 participants from the UK Biobank, utilizing genetic data and environmental variables (including air pollution, noise, greenspace and water quality). Polygenic risk score (PRS) was calculated to estimate individual genetic risk. Cox proportional hazard model was applied to estimate the impact of exposure factors on the risk of AF occurrence. The mediation analysis was applied to assess the relationship among environmental scores, AF and cardiac structures. Population attributable fraction (PAF) was employed to assess potential influence of mitigating unfavorable environment characteristics on AF. The results showed that the highest group of four domain scores exhibited 3.38–16.83% higher AF risk than the lowest. Individuals with higher scores in four domains and high PRS had hazard ratio (95%CI) of 2.76 (2.62, 2.91), 2.61 (2.47, 2.75), 2.86 (2.71, 3.02) and 2.84 (2.66, 3.02). Environmental factors could indirectly affect cardiac structures through AF. Up to 7.37% of AF cases could be preventable through environmental interventions. Our findings pointed that gene-environment interaction can increase AF risk, which further affect cardiac structures.

Keywords Atrial fibrillation, Genetic risk, Gene-environment interaction, Cardiac structure

Atrial fibrillation (AF) is the most prevalent arrhythmia produced by a disturbance in the heart's electrical signals. The incidence of AF exceeded that of 1997 by 31%, with 37.574 million cases worldwide, and is expected to continue to rise in the future, posing one of the greatest epidemiological challenges[1]. Some risk factors such as lifestyle, genetics, and environmental factors have been found to be associated with the occurrence of AF[2,3]. Environmental exposures of broad scope are considered modifiable risk factors for AF, making the investigation of their relationship with AF and implementation of preventive measures crucial strategy for reducing the burden of AF.

Notably, recent studies have provided convincing evidence showing that environment pollution, especially air pollution is major cause of worldwide illness burden, which could cause cardiovascular disease (CVD) and changes in heart morphology and function[4,5]. However, the association between AF risk and air pollution exposure was inconsistent[6–10]. In addition to air pollution, noise pollution, greenspace cover, and water quality are critical environmental factors. The association between noise exposure and AF risk remains inconclusive[10,11]. A study based on greening and AF showed that those with the highest greening rate had a 6% lower risk of developing AF than those with the lowest greening rate[12]. Among those environmental factors, the relationship between water quality and AF risk has not been fully explored.

Despite extensive research examining the association between environmental factors and AF risk, a comprehensive examination of environment factors has been lacking. Studies often suffered from limited sample sizes and underrepresentation of population. Integrating data from diverse environmental domains,

[1]School of Electrical and Information Engineering, Zhengzhou University of Light Industry, Zhengzhou, China. [2]Department of Medical Imaging, Henan Provincial People's Hospital & Zhengzhou University People's Hospital, Zhengzhou, China. [3]Laboratory of Brain Science and Brain-Like Intelligence Technology, Biomedical Research Institute, Henan Academy of Science, Henan, China. [4]Department of Hypertension, Henan Provincial People's Hospital & Zhengzhou University People's Hospital, Zhengzhou, Henan, China. [5]Department of Neurology, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China. ✉email: mywang@ha.edu.cn; jingjingsu2000@163.com; tinagra1989@163.com

such as air pollution, noise pollution, greenspace, and water quality, might generate a more comprehensive understanding of environmental exposure and its implications for AF. In addition, the joint analysis enabled us to comprehensively assess the joint impact of multiple risk factors and precisely quantify their contribution to AF prevalence, which is crucial for specifying the prevention strategy.

In addition to environmental factors, the genetic risk of AF has also been identified[3,13]. There is increasing evidence that genetic variants can alter personal vulnerability to environment[14]. However, it is unclear whether genetic susceptibility has a potential modifying effect on the association between joint exposure to environmental factors and AF.

Therefore, we carried out a large cohort research using comprehensive information on genetic variation and environment, which were provided by the UK Biobank, to investigate the relationship between exposure to different environmental factors and the incidence rate of AF, as well as the interaction between composite scores for different environmental domains and genetic predisposition, and their impact on cardiac structure. We have quantified the PAF in detail for each environmental domain to assess the potential effectiveness of preventive strategies (Fig. 1).
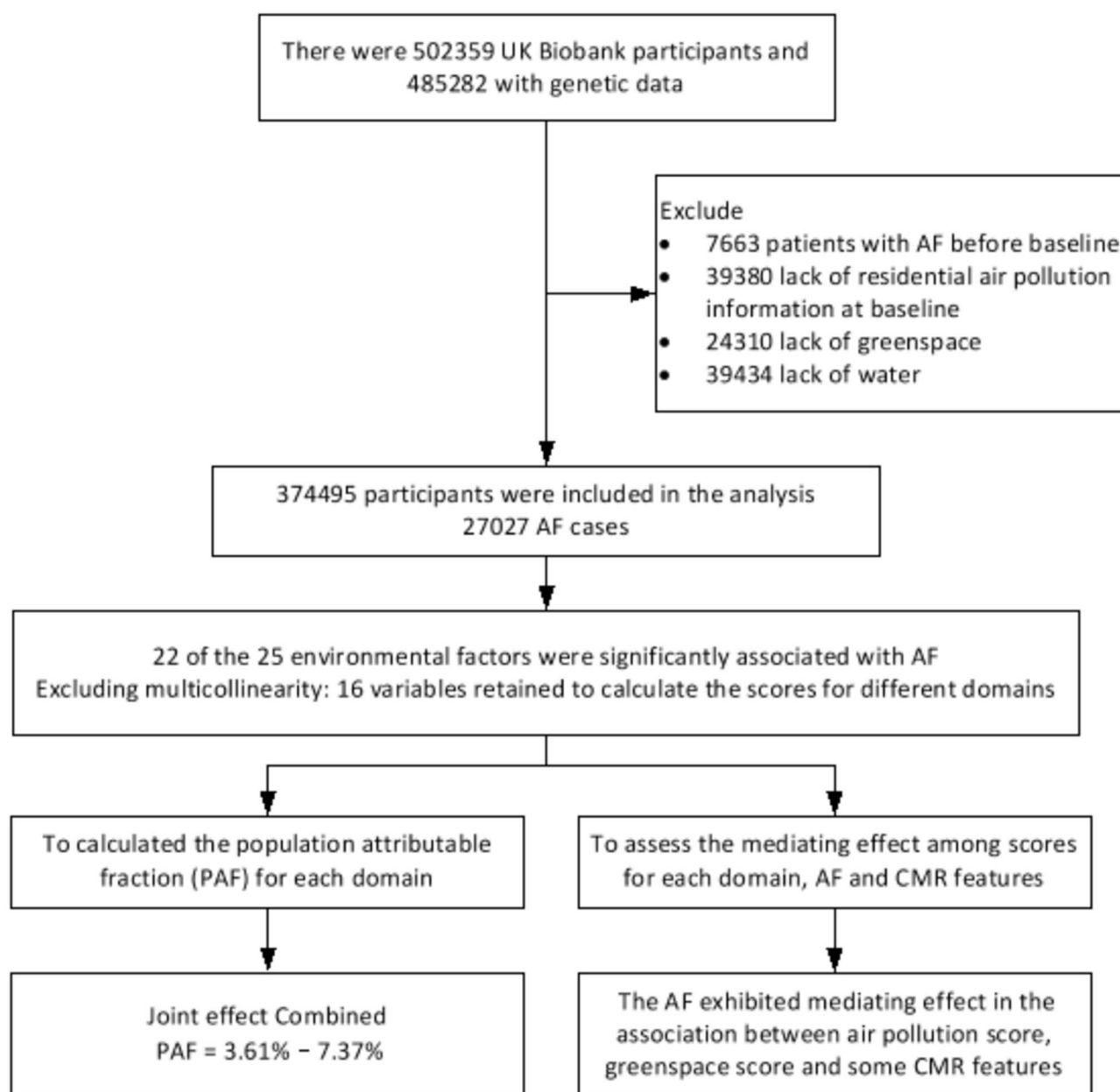

Fig. 1. Work flow for the participants selection and study design. AF, atrial fibrillation; PAF, population attributable fraction; CMR, cardiac magnetic resonance.

Results

The baseline characteristics of participants were exhibited in Table 1. We recorded 27,027 cases of AF among the 374,495 participants during a median of 13 years of follow-up. The spatial distribution of AF event was shown in Supplementary Fig. 1. Compared to controls, patients with AF were found to be elder (61.90 ± 6.09 years VS 56.11 ± 8.07 years), predominantly male (61.2% VS 38.8%), and tended to have lower income levels (Less than 18,000: 27.3% VS 19%) and less-educated (69% VS 62.8%). In addition, they had a higher proportion of former or current smokers (54.6% VS 44.3%), had higher alcohol scores (2.49 ± 2.77 VS 2.08 ± 2.28) and were fatter (BMI: 28.91 ± 5.35 VS 27.26 ± 4.69). All the observed differences between AF patients and controls were statistically significant (all $P < 0.001$).

In this study, of the 25 examined factors, 22 were significantly associated with AF (Fig. 2). Among these, 11 factors displayed potential adverse effects, while 11 were protective (Supplementary Table 1). The Spearman correlation among the significant variables was presented in Supplementary Fig. 2. In the course of the collinearity examination, we rigorously accounted for and excluded the following variables ($r^2 > 0.9$): In terms of noise pollution, since 24-hour noise pollution is more representative, we excluded 16 h noise pollution, daytime, evening and night-time noise pollution. In other areas, when comparing Greenspace percentage (buffer 1000 m) with Natural environment percentage (buffer 1000 m), the latter was more significant, and when comparing Water hardness USGS with CaCO3 concentration, the latter was also more significant, hence the former were excluded.

The correlation between scores in four domains (air pollution, noise pollution, greenspace, and water) and AF incidence was presented in Table 2. In the age-, sex- and assessment center- adjusted model, participants in the highest group of four domain scores exhibited 3.38-16.83% higher risk of developing AF than individuals in the lowest ($P < 0.05$). Our sensitivity analysis demonstrated consistent results between unweighted scores and AF risk. After adjusting for multi factors, the patterns of outcomes for the other factors aside from noise

Characteristics	Total population ($n = 374495$)	Atrial fibrillation		
		Yes ($n = 27027$)	No ($n = 347468$)	p-value
Age [years (mean $\pm$ SD)]	56.53 ± 8.08	61.90 ± 6.09	56.11 ± 8.07	$P < 0.001$
Follow-up time [months, median(IQR)]	156 (17)	107(76)	157 (17)	$P < 0.001$
Sex [n(%)]				
Female	204,832 (54.7)	10,476 (38.8)	194,356 (55.9)	$P < 0.001$
Male	169,663 (45.3)	16,551 (61.2)	153,112 (44.1)	$P < 0.001$
Household income [n(%)]				
Less than 18,000	73,446 (19.6)	7384 (27.3)	66,062 (19.0)	$P < 0.001$
18,000 to 30,999	81,969 (21.9)	6594 (24.4)	75,375 (21.7)	$P < 0.001$
31,000 to 51,999	138,612 (37.0)	9469 (35.0)	129,143 (37.2)	$P < 0.001$
52,000 to 100,000	63,716 (17.0)	2854 (10.6)	60,862 (17.5)	$P < 0.001$
Greater than 100,000	16,752 (4.5)	726 (2.7)	16,026 (4.6)	$P < 0.001$
Higher education level [n(%)]				
Yes	137,771 (36.8)	8389 (31.0)	129,382 (37.2)	$P < 0.001$
No	236,724 (63.2)	18,638 (69.0)	218,086 (62.8)	$P < 0.001$
Body mass index [kg/m²(mean $\pm$ SD)]	27.38 ± 4.76	28.91 ± 5.35	27.26 ± 4.69	$P < 0.001$
Alcohol Score (mean $\pm$ SD)	2.11 ± 2.32	2.49 ± 2.77	2.08 ± 2.28	$P < 0.001$
Smoking status [n(%)]				
Never smoking	205,550 (54.9)	12,283 (45.4)	193,267 (55.7)	$P < 0.001$
Previous smoking	130,093 (34.7)	11,826 (43.8)	118,267 (34.0)	$P < 0.001$
Current smoking	38,852 (10.4)	2918 (10.8)	35,934 (10.3)	$P < 0.001$
IPAQ activity group [n(%)]				
Low	116,666 (31.2)	7930 (29.3)	108,736 (31.3)	$P < 0.001$
Moderate	205,152 (54.8)	15,119 (55.9)	190,033 (54.7)	$P < 0.001$
High	52,677 (14.1)	3978 (14.7)	48,699 (14.0)	$P < 0.001$
Healthy diet score [n(%)]				
0–1	2700 (0.7)	218 (0.8)	2482 (0.7)	$P < 0.001$
2–3	135,407 (36.2)	10,757 (39.8)	124,650 (35.9)	$P < 0.001$
4–5	236,388 (63.1)	16,052 (59.4)	220,336 (63.4)	$P < 0.001$

Table 1. Baseline characteristics of participants in this study. SD: stand deviation; Higher education level: College/University degree and/or Other professional qualifications e.g.: nursing, teaching.
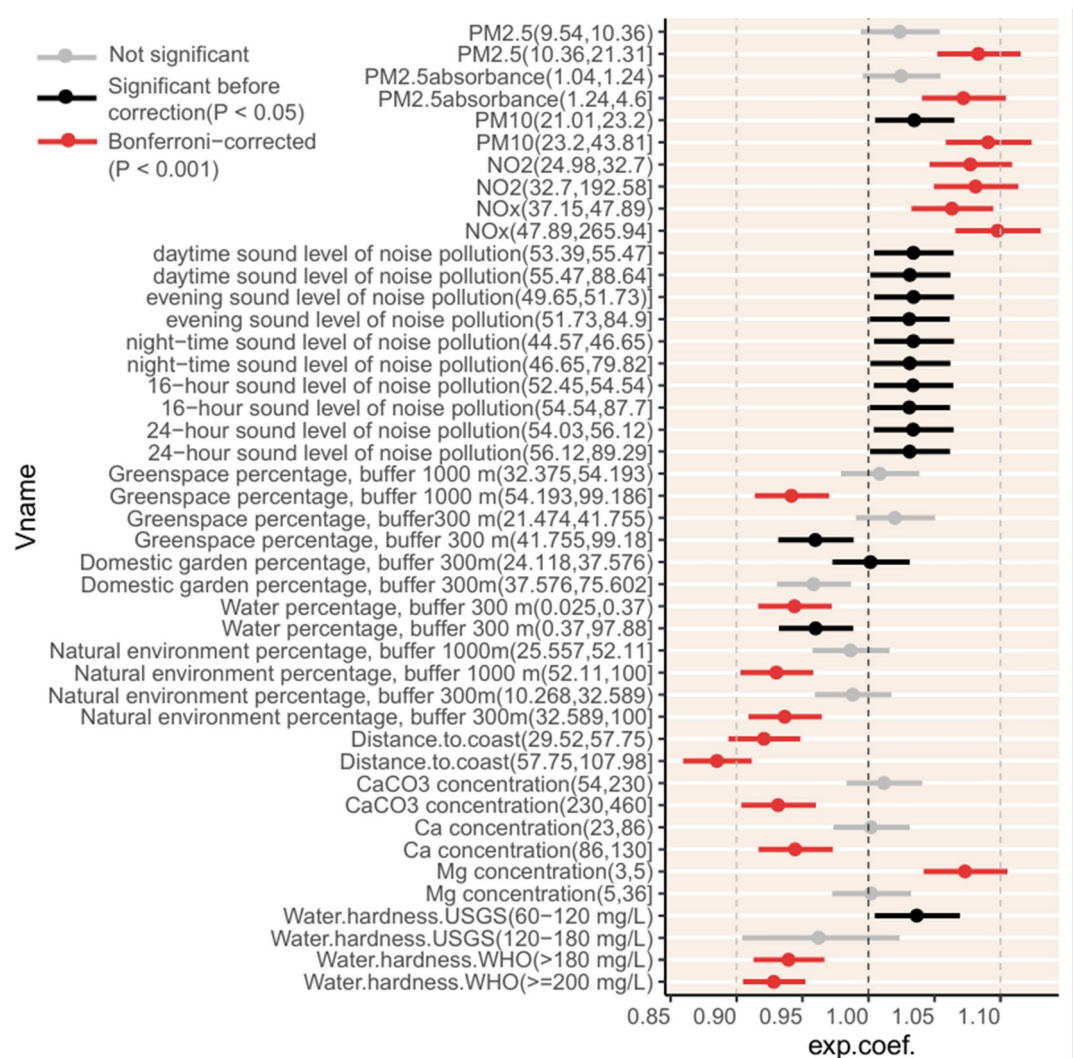
Fig. 2. Associations between significant environmental variables and the incidence of atrial fibrillation. NO2, nitrogen dioxide; NOx, nitrogen oxides; PM2.5, particular matter with aerodynamic diameter ≤ 2.5 μm; PM10, particular matter with an aerodynamic diameter ≤ 10 μm; PM2.5–10, particular matter with an aerodynamic diameter between 2.5 and 10 μm. Each point represents a hazard ratio, accompanied by a horizontal line indicating the corresponding 95% confidence interval.

Exposures	Subgroup	HR	CI5	CI95	P	P trend
Air pollution	Favourable	1	–	–	Reference	9.12×10^{-11}
	Intermediate	1.0627	1.0329	1.0934	4.86×10^{-5}	
	Unfavorable	1.1013	1.0692	1.1344	9.06×10^{-11}	
Noise pollution	Favourable	1	–	–	Reference	0.024733
	Intermediate	1.0311	1.0014	1.0618	0.040143	
	Unfavorable	1.0338	1.0042	1.0644	0.024982	
Greenspace	Favourable	1	–	–	Reference	1×10^{-20}
	Intermediate	1.0666	1.0357	1.0985	1.7×10^{-5}	
	Unfavorable	1.1683	1.1348	1.2028	1×10^{-20}	
Water	Favourable	1	–	–	Reference	8.14×10^{-11}
	Intermediate	1.1090	1.0743	1.1448	1.72×10^{-10}	
	Unfavorable	1.1206	1.0851	1.1572	4.15×10^{-12}	

Table 2. Relationships between atrial fibrillation and the four domains. HR, hazard ratio; CI, confidence interval.

pollution were nearly identical (Supplementary Table 2). The spatial distribution of each score were shown in Supplementary Fig. 3.

For the genetic correlation analysis (Supplementary Table 3), we found a significant and statistically significant correlation between PRS and AF risk (P $_{trend}$ < 2×10^{-16}). Compared with participants with low genetic risk, participants with medium to high genetic risk had a risk of AF that was approximately 53% (48-58%) or 152% (144-162%) higher. The same association was also shown in the sensitivity analysis using white samples and PRS based on independent datasets.

In the hierarchical analysis (Fig. 3), compared to individuals with lower scores in four domains and low PRS, those scoring higher and exhibiting higher PRS had HRs(95%CI) of 2.76(2.62,2.91), 2.61(2.47,2.75), 2.86(2.71,3.02) and 2.84(2.66,3.02). Furthermore, an interaction was found between the scores of the four domains and PRS (*P* < 0.001).

The results of Lasso regression and 10-fold cross-validation are presented in supplementary Fig. 4, screening out six Cardiac magnetic resonance (CMR) features significantly associated with AF. Table 3 summarized the relationship among scores for four domains, the occurrence of AF, and CMR features. We discovered that higher air pollution score was linked to lower left atrial ejection fraction (LAEF), in part due to a mediating mediated by AF. The study also found that AF partly mediated of the relationship between greenspace score and descending aorta minimum area. We did not view a relationship among AF-mediated noise pollution score, water score and CMR features. Forevermore, the diagnosis time of most AF is later than the baseline exposure assessment and earlier than the CMR feature measurement (Supplementary Fig. 5).

According to PAF estimates (Table 4), it is suggested that 3.61% of AF cases could be prevented by eliminating solely adverse factors (model1). When both moderate and adverse factors (model2) were eliminated, the potential for prevention increases to 7.37% of cases. In the conservative model (model 1), greenspace exerted the most significant preventive impact on AF, reducing it by 2.25%. Air pollution (0.53%), water (0.80%), and noise pollution (0.03%) followed incidence in descending order of influence. Under the optimistic estimates (model 2), greenspace (3.46%) contributed the most to AF prevention, with water (2.57%), air pollution (0.87%) and noise pollution (0.47%) trailing behind. Moreover, the consistency of the findings persisted when we limited participants' follow-up to periods exceeding six years (Supplementary Table 4).

Discussion

In our study, we evaluated the association between exposure to different environmental factors and AF risk. We calculated scores for four domains including air pollution, noise pollution, greenspace and water to assess combined impact of each domain. The results presented that scores for four domains were tied to an increased risk of AF. We noticed that the connection between scores for four domains and AF incidence was increased by inherited predisposition to AF. Through mediation analysis we found that air pollution score and greenspace score could indirectly affect some CMR features through the onset of AF. Overall, we estimated that 3.61–7.37% of AF cases could be potentially preventable through environmental interventions.

Earlier investigations have mainly aimed at analyzing the impact of one aspect of environment on AF[15], and the results have been disputed[16,17], and largely ignored the effects arising from other environment factors. We used a large sample size, considered relatively comprehensive environmental factors and their impact on AF. Our analysis identified 22 factors associated with AF. Consistent with previous research, PM$_{2.5}$[18,19], PM$_{10}$[6], NO$_2$[20,21], and noise[22] were identified as significant risk factors, while greenspace[12]. emerged as a potentially protective element. The study also uncovered less-explored factors, such as PM$_{2.5}$ absorbance, natural environment percentage, domestic garden percentage, domestic water composition, and water hardness.

In consideration of people are exposed to mixed environment at the same time, it becomes imperative to evaluate joint exposures of various environmental factors[23]. The continuous scores provided a more complete measure of combined exposure to environmental factors in each fields[24]. Though the degree of effects of four environment domains on the occurrence of AF is different, unfavorable environment could promote the occurrence of AF.

We further investigated the potential interplay between comprehensive scores and genetic risk about AF. We discerned that genetic susceptibility to AF alters the relevance between each score and AF risk, with an additive interaction. Compared individuals possessing lower genetic predisposition, the link between each score and AF risk was more pronounced in those with higher genetic predisposition. In our study, certain genes associated with AF, such as IL-6, was also linked to markers of inflammatory through gene-environment interactions, which in turn is a causative risk factor for AF[25]. Regardless, our results indicated that individuals at high genetical risk of AF are at greater susceptible from environment.

AF was a mediating factor between air pollution score, greenspace score and some CMR features. We observed that AF resulted in smaller LAEF, larger descending aorta minimum volume. The decrease of LAEF is an early sign of the development of HF events[26], It is suggested that environmental exposure may accelerate the deterioration process of cardiac function through AF. Previous studies have shown that the aortic stiffness increases in patients with AF[27], and green space exposure may indirectly improve aortic elasticity by reducing the occurrence of AF. Plant volatiles related to green Spaces can improve vascular endothelial function[28], which may function independently of AF. In summary, exposure to such environmental risk factors was highly connected with the occurrence of AF and structural heart changes.

When solely considering eliminating adverse factors, the study indicated that the scarcity of greenspace was the leading PAF contributor to AF, accounting for 2.25%. When more exhaustive risk factor elimination is implemented, ecological greening still exhibited the greatest potential, resulting in a 3.46% reduction in AF incidence. High-quality green spaces have been scientifically proven to offer multiple health benefits[29], including stress reduction[30], increased engagement in physical activities[31], reduced exposure to air pollution[32], and a
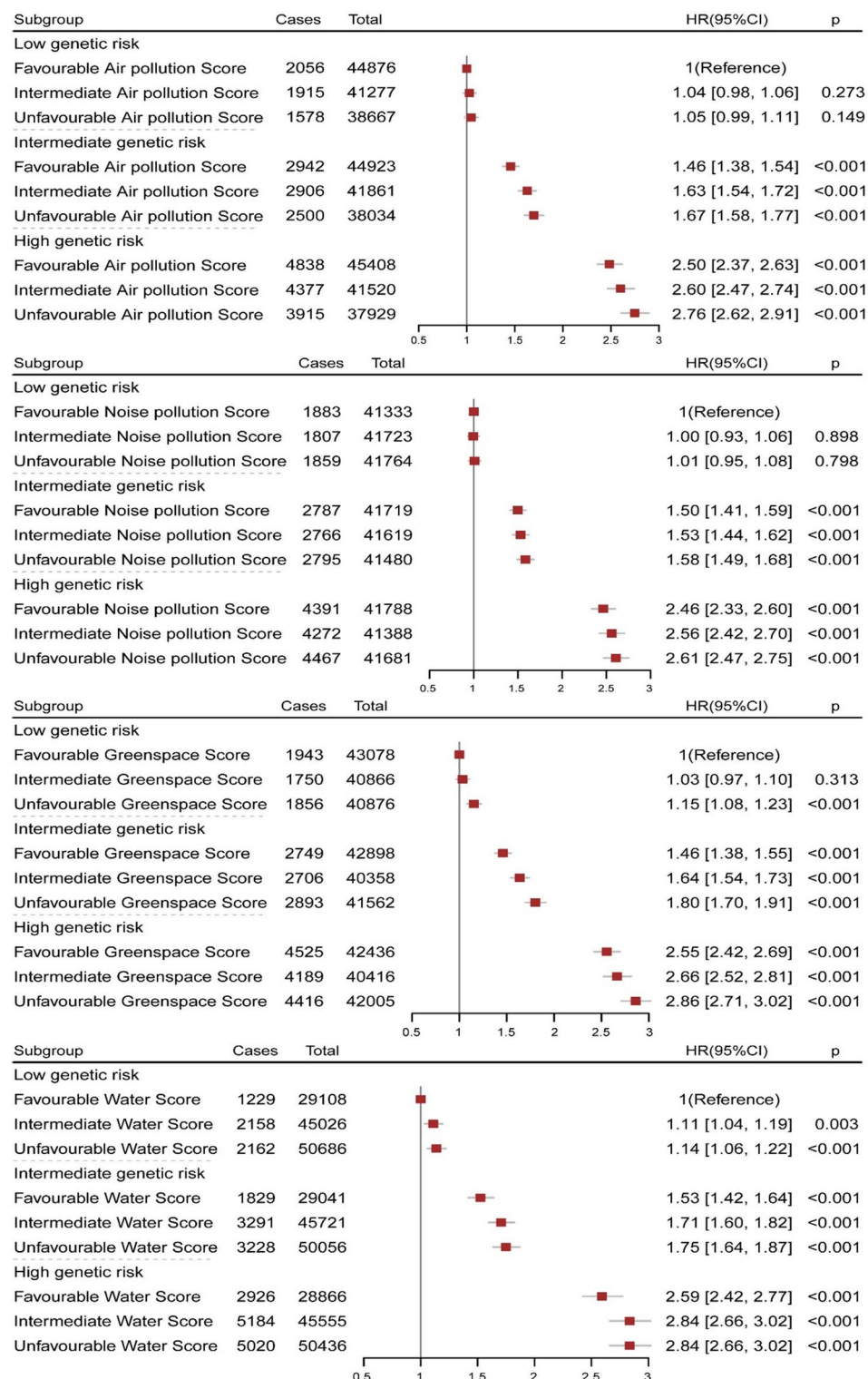
Fig. 3. Risk of atrial fibrillation stratified by four domains within each genetic risk category. HR, hazard ratio; CI, confidence interval.

decreased risk of cardiometabolic diseases[33]. These advantages may be closely associated with a decline in AF prevalence.

The detrimental impact of poor water quality for AF was significant, with a PAF as high as 2.57%. Existing research indicated hard water consumption may have a potentially positive role in CVD prevention[34,35]. Our

Exposure	Outcome	Total effect		Mediation effect	
		Effect size	P	Effect size	P
Air pollution score	LA Minimum volume	0.063	0.087	0.033	0.003**
	LAEF	– 0.09	**0.001 *****	– 0.017	**0.003****
	Minimum volume of RA	– 0.096	0.071	0.68	0.004**
	Ascending aorta minimum area	– 0.734	0.182	0.065	0.009**
	Descending aorta minimum area	0.108	0.581	0.033	0.005**
	LV average myocardial wall thickness AHA 9	– 0.002	0.402	0.001	0.005**
Noise pollution score	LA Minimum volume	0.054	0.636	0.018	0.537
	LAEF	– 0.039	0.586	0.009	0.54
	Minimum volume of RA	0.061	0.68	0.012	0.538
	Ascending aorta minimum area	0.333	0.777	0.035	0.542
	Descending aorta minimum area	– 0.306	0.508	0.018	0.535
	LV average myocardial wall thickness AHA 9	– 0.002	0.781	0	0.526
Greenspace score	LA Minimum volume	– 0.05	0.122	0.025	0.002**
	LAEF	– 0.016	0.409	– 0.013	0.002**
	Minimum volume of RA	– 0.305	0.001 ***	0.016	0.003**
	Ascending aorta minimum area	0.058	0.881	0.048	0.009**
	Descending aorta minimum area	0.499	**0.001****	0.025	**0.007****
	LV average myocardial wall thickness AHA 9	0.003	0.125	0	0.004**
Water score	LA Minimum volume	– 0.04	0.6	0.017	0.427
	LAEF	– 0.087	0.065	– 0.008	0.428
	Minimum volume of RA	– 0.079	0.351	0.011	0.424
	Ascending aorta minimum area	1.49	0.075	0.032	0.447
	Descending aorta minimum area	0.579	0.146	0.016	0.45
	LV average myocardial wall thickness AHA 9	0.001	0.834	0	0.449

Table 3. Mediation analysis of the mediation effect of atrial fibrillation in abnormal CMR induced by the four domains. Significant values are in [bold]. AF, atrial fibrillation; CI, confidence interval; LAEF, LA ejection fraction.

Domains	Model1			Model2		
	Unweighted PAF	Communality	Weighted PAF	Unweighted PAF	Communality	Weighted PAF
Air pollution	0.79%	45.59%	0.53%	1.11%	62.15%	0.87%
Noise pollution	0.50%	0.40%	0.03%	1.07%	18.95%	0.47%
Greenspace	3.05%	54.29%	2.25%	5.40%	41.13%	3.46%
Water	1.24%	41.53%	0.80%	6.06%	17.96%	2.57%
Overall weighted PAF			3.61%			7.37%

Table 4. Weighted and unweighted PAF for the four domains. PAF, population attributable fraction. In Model 1, we transformed adverse factors into moderate and favorable ones. In Model 2, we converted both adverse and moderate factors into favorable ones.

study findings supported the hypothesis that hard water consumption may have a preventive effect on CVD. However, further validation and confirmation are required.

The PAF contribution of air pollution to AF was 0.87%. Air pollution exposure could cause atrial enlargement, oxidative stress, and autonomic nervous imbalance[36], which could further alters intercellular coupling and gap connection function, followed by structural and electrical changes in the atriums, ultimately leading to AF[37]. Exposure to NO_2 might induce mitochondrial dysfunction, thereby increasing arrhythmia risk[38]. Notably, AF itself is a major risk factor for ischemic stroke, and recent evidence suggests that environmental factors may exacerbate this pathway. For instance, Święczkowski et al. demonstrated that short-term exposure to $PM_{2.5}$ and NO_2 significantly increased ischemic stroke incidence, particularly in vulnerable subgroups such as non-elderly women[39]. This aligns with our findings on air pollution-induced AF, suggesting a potential synergistic effect where pollution-triggered AF may further elevate stroke risk. The interplay between environmental exposures, AF, and stroke underscores the need for integrated prevention strategies targeting both air quality and cardiovascular

health. This not only underscored the strong association between air pollution and AF, but also highlighted the critical importance of improving air quality in the prevention and management of this condition.

The PAF proportion of noise-related factors in AF was relatively low, potentially due to the high correlation in noise data, which restricts the number of contributing factors in noise score. However, our study demonstrated a significant decline in AF incidence when noise pollution is more effectively mitigated compared to solely addressing unfavorable characteristics. Previous observational studies have also demonstrated a link between noise pollution and AF[40,41], further emphasizing the significant influence of noise on the disease's onset.

Our study, for the first time, simultaneously investigated the impact of multiple environmental factors on the risk of AF, controlling for highly correlated variables, and employed PAF weights in each domain to expose their non-independence. However, the intricate relationship among these factors might transcend mere consideration of shared risk. Though limitations, the study's results are encouraging, suggesting the potential for environmental improvements in AF prevention.

There were a few potential limitations to acknowledge. Firstly, some environmental data available from the UK Biobank encompassed limited time points, changes in environment throughout the follow-up period could not be evaluated. Secondly, lack of data on certain air-related exposures like CO, SO_2, and O_3, which could potentially influence AF risk. Moreover, our study not considered information about environment at locations other than the subject's home addresses, potentially leading to exposure measurement errors. Lastly, the participants in this study included only the British aged between 40 and 69 years, so generalizing conclusions of PAF estimates to other age and other ethnic groups need to be interpreted carefully. Although all death cases were retained for analysis in this study, the existence of competitive risk may affect the interpretation of the results through unmeasured confounding factors.

Materials and methods
Study design and populations
The data we used comes from UK Biobank, which is a world-leading biomedical database. Additional information regarding study protocol has been discussed elsewhere[42]. In brief, the study conducted 2006 to 2010, encompassed more than 0.5 million British aged 40–69. Participants took part in a baseline assessment at a center, where data regarding lifestyle, physical fitness measurements, and biological samples were gathered. The UK Biobank study obtained approval from the Northwest Multicentre Research Ethics Committee. Each of the participants offered knowledgeable written informed consent. This study was performed in accordance with the principles of the Declaration of Helsinki. Our study excluded participants with missing genetic and environmental data, as well as those who had already suffered from AF prior to the baseline. The final analysis included 374,495 participants.

Assessment of outcomes
The source of AF report was from self-reports, related medication, hospital admission records, and death registration (Supplementary Table 5). Hospital admission records were determined by linking records with Health Event Statistics for England and Wales and the Scottish Onset Record, enabling exact determination of the date of first recording for each diagnosis. AF status was ascertained using ICD-10 codes I48 from hospital registry data (Data-field 41270).

Environmental factors
We collected data on environment that were measured or derived at baseline and divided into 4 categories: (1) residential air pollution: particulate matter (PM) with diameters ≤ 2.5 μm ($PM_{2.5}$), $PM_{2.5}$ absorbance, ≤ 10 μm (PM_{10}), ranging from 2.5 to 10 μm ($PM_{2.5-10}$), and nitrogen oxides (NO_2 and NO_x). (2) residential noise pollution: 16 h and 24 h noise pollution, daytime, evening and night-time noise pollution. (3) greenspace and coastal proximity: natural environment, greenspace, domestic garden and water percentage (at 1000 m and 300 m home location buffers), and distance from home location to the coast. (4) water minerals content: domestic water concentration of calcium (Ca), calcium carbonate ($CaCO_3$) and magnesium (Mg), and the water hardness classification. The detail of estimation of environmental factors in Supplementary Text 1.

Genotype data, QC, and PRS
The details of genotyping, imputation process, phasing and quality control have been previously documented[43](detail in Supplementary Text 2). The calculation of PRS for AF utilized a partially independent subset of SNPs above a specific threshold for the P-value of GWAS association[44]. To extract SNPs for inclusion in the PRS calculation, we obtained GWAS statistics of AF from the IEU Open GWAS Project (https://gwas.mrcieu.ac.uk/). The GWAS of ebi-a-GCST006414 with status available, most recent, and largest sample size was chosen, which involved 1,030,836 European individuals, among which there were 60,620 cases and 970,216 controls. We filtered GWAS statistics results that did not possess variants with $p \geq 5 \times 10^{-10}$. Then, we used linkage disequilibrium to identify independently correlated variants ($r^2 < 0.001$).

An additive genetic model containing 73 SNP (detailed in Supplementary Table 6) was selected for PRS calculation and carried on the standardization processing. The PRS was computed using the following formula:

$$PRS_j = \left[\sum_{ij} \beta_i \times G_{ij} - mean\,(PRS) \right] / \text{SD}\,(PRS),$$ where i represents the i_{th} SNP, j represents the j_{th} individual, β represents the statistical coefficient of SNP, and G represents the number of observed effective alleles. The procedure mentioned above was executed on PRSlice-2 in UKB Research Analysis Platform.

We adopted the previously published PRS of AF based on independent datasets to test the reliability of our PRS-related results (Marston, Nicholas A et al. "A polygenic risk score predicts atrial fibrillation in

cardiovascular disease." European heart journal vol. 44,3 (2023): 221–231. doi:https://doi.org/10.1093/eurhear tj/ehac460). Download the SNPS and corresponding weights of AF risk in the published PRS articles (detailed in Supplementary Table 7). Upload the weights of each SNP in the downloaded file to the UK Biobank Research Analysis Platform for further processing. The plink software in the platform is used to calculate the text-based PRS of each UKB participant.

Measurements of covariates

This study comprehensively examined several potential confounding factors form demographic, social, economic and life style points, including age, sex, average total household income before tax, qualifications, smoking status, alcohol score, activity group, and healthy diet scores. Detail in Supplementary Text 3.

Statistical analysis

The study recorded the survival time for per participant from their initial recruitment date until the occurrence of AF, a competitive event such as death, or censorship, whichever happened first. The Cox proportional hazards models were applied to assess relationships between individual environmental factors and AF, adjusting for age, gender and assessment center, estimating hazard ratios (HRs) and corresponding 95% confidence intervals (CIs). Continuous variables were categorized into tertiles, with the lowest third serving as a reference point for the analysis. Twenty-five variables were collected, so the total number of tests was 50. This method was chosen to be consistent with the common practice of analyzing dose-response relationships in epidemiological studies. Using the lowest quantile as a reference group helps us interpret the findings more intuitively, where a HR greater than 1 means that the other groups have increased risk relative to the baseline low-risk group. The proportional hazards assumption was tested by analyzing the relationship between time and standardized Schoenfeld residuals[45]. A conservative Bonferroni correction method was applied to determine a more significance threshold ($P < 0.001$). This was calculated by dividing the conventional significance level of 0.05 by total number of tests (50). We assessed the multicollinearity among the significant variables using the Spearman correlation coefficient, determining the presence of potential multicollinearity when the Spearman correlation coefficient (r^2) between variables exceeded 0.9. For variable pairs with $r^2 > 0.9$, we excluded the variable with lower statistical significance to ensure the robustness of our analysis.

We categorized the variables significantly associated with AF into four domains: air pollution, noise pollution, greenspace and coastal proximity, and water minerals content. Protective factors (HR < 1) were subjected to a reverse coding approach, serving as an indicator of potential detrimental effects, with each factor demonstrating harmfulness (HR > 1) awarded a score of 1. In the Cox models, where risk factors within each domain were mutually adjusted and baseline age, sex, and assessment center were controlled, we gained the β-coefficients for each variable, which used to generate weighted risk scores for each domain. The initial binary variables were multiplied by their respective β coefficients, summing the products, and finally dividing by the total sum of the β coefficients. The resulting score served as a measure of an individual's exposure to a number of risk factors, with higher scores indicating a higher level of exposure. We further classified scores into tertiles: favorable, moderate, and unfavorable.

To estimate the impact of risk scores in each domain on AF, we adopted multiple Cox regression models. First of all, we examined the relationship between risk scores in four domains and AF, adjusted according to age, gender and assessment center. Then, we further included the collected confounding factors as covariates in the Cox regression model. Furthermore, we conducted the same Cox regression analysis using unweighted risk scores to further verify the robustness of the results.

To assess the impact of genetic factors, we evaluated the relationship between PRS and the risk of AF using a PRS based on the UK Biobank dataset. Two models were utilized: the first model adjusted only for age, sex, assessment center, and genotyping measurement, while the second model further adjusted for all collected confounding factors. In the sensitivity analysis, we extracted only White samples for an independent analysis and also assessed the relationship between PRS and the risk of AF using a PRS based on an independent dataset.

To explore the potential influence of genetic predisposition on the relevance between AF incidence and each score, we divided participants into 9 groups based on levels of genetic risk (low, medium, and high) and the tertiles of each score. We evaluated the HR of each group of AF, and adjusted for age, gender, assessment center and genotype measurement as covariates in the Cox model. We also looked at the interplay between genes and each domain on AF incidence by including continuous composite scores in four domains and PRS cross-product terms into the model.

To investigate the correlation among environment, AF, and changes in CMR features, and considering the multicollinearity problem in the association analysis between CMR features and AF, the contents of CMR image acquisition and analysis are supplemented in Supplementary Text 4. Wwe used Lasso regression to screen out CMR features that were significantly associated with AF. This step is a necessary prerequisite for mediation analysis and meets the requirements of the Baron & Kenny pathway model, that is, there must be significant associations among exposure–mediation— outcome[46]. We constructed a Lasso regression model with AF as the dependent variable and CMR features as the independent variables. The optimal penalty parameter λ was selected through 10-fold cross-validation to minimize the model prediction error. The data were divided into 10 parts, with 9 parts used as the training set and 1 part as the validation set in each iteration. After repeating this process 10 times, the average prediction error was calculated. The λ value that minimized the average prediction error was chosen as the optimal parameter. As the penalty parameter λ increased, some variables with weaker associations with AF had their coefficients shrink to zero earlier and were automatically screened out.

Previous studies have not focused on the changes in CMR characteristics caused by environmental factors influencing the risk of AF. Therefore, this study hypothesized that environmental exposure could indirectly affect CMR characteristics by increasing the risk of AF, and this study tested the mediating role among them.

The mechanism pathways of environmental exposure affecting CMR characteristics through AF were examined through three key steps: Firstly, an exposure-outcome total effects model (CMR ~ environmental exposure) was established to evaluate the direct impact of environmental factors on CMR parameters; Secondly, the exposure - mediation model (AF - environmental exposure) and the mediation - outcome model (CMR - AF + environmental exposure) were respectively fitted to obtain the path coefficients; Finally, the mediating effect size was calculated by the product method (path a× path b), and the 95% confidence interval was calculated by Bootstrap sampling (1000 times) for statistical inference. In order to prove the rationality of the mediation analysis, a three-stage density curve graph was used to display the temporal data distribution density of exposure - mediation - outcome.

To quantify the proportion of AF reduction that could be attributed to replacing risk factors with beneficial factors, we estimated the PAF for each domain. Using the Cox regression model adjusted for age, gender, assessment center, and all collected confounding factors, we calculated the HR for each domain. Then, we generated PAFs for each domain using the formula: $PAF = (P_{pop} \times (HR - 1))/[P_{pop} \times (HR - 1) + 1]$, where P_{pop} represents the exposure rates in the total population. PAF is commonly employed for categorical variables[47]. Consequently, we employed two distinct models: the first model incorporated both advantageous and moderate factors from four domains, eliminating adverse ones, to yield a conservative estimate. The second model retained only the advantageous factors, eliminating both moderate and adverse elements, intending to provide a more optimistic forecast. First, PAFs for each domain were generated in Cox models that were adjusted for baseline age, gender, and the assessment center. To mitigate the PAF potential overestimation due to interaction among risk factors, we employed principal component analysis (PCA) to compute the commonality among the factors applying it to determine the weights of each PAF. Both individual-weighted PAFs and combined-weighted PAFs were calculated based on the PCA-derived factor weights. we included only individuals who were monitored for over six years in our sensitivity analysis to reduce possible influence of reverse causality on observed association.

Conclusions

In this retrospective study, our findings revealed the association between AF and multiple environmental factors. The results indicated air pollution, noise pollution, inadequate green spaces, and water quality deterioration were associated with an increased risk of AF. The gene-environment interactions could increase the risk of AF. AF occurs could further affect the cardiac structure. The PAF estimated underscores the significance of prioritizing improvements in greenspace and water quality. Enhanced air quality and reduced noise pollution could further contribute to risk reduction.

Data availability

This research has been conducted using the UK Biobank, a publicly accessible database. Data were used under application number 94,885. The datasets are available to researchers through an open application via https://www.ukbiobank.ac.uk/enable-your-research/apply-for-access.

References

1. Lippi, G., Sanchis-Gomar, F. & Cervellin, G. Global epidemiology of atrial fibrillation: an increasing epidemic and public health challenge. *Int. J. Stroke* **16**, 217–221. https://doi.org/10.1177/1747493019897870 (2021).
2. Kornej, J., Borschel, C. S., Benjamin, E. & Schnabel, R. B. Epidemiology of atrial fibrillation in the 21st century novel methods and new insights. *Circ. Res.* **127**, 4–20. https://doi.org/10.1161/circresaha.120.316340 (2020).
3. Wang, B. et al. Integrative omics approach to identifying genes associated with atrial fibrillation. *Circ. Res.* **126**, 350–360. https://doi.org/10.1161/circresaha.119.315179 (2020).
4. de Bont, J. et al. Ambient air pollution and cardiovascular diseases: an umbrella review of systematic reviews and meta-analyses. *J. Intern. Med.* **291**, 779–800. https://doi.org/10.1111/joim.13467 (2022).
5. Hu, J. J. et al. Association of long-term exposure to ambient air pollutants with cardiac structure and cardiovascular function in Chinese adults. *Ecotoxicol. Environ. Saf.* **249**, 475. https://doi.org/10.1016/j.ecoenv.2022.114382 (2023).
6. Stafoggia, M. et al. Short-term effects of particulate matter on cardiovascular morbidity in Italy: a National analysis. *Eur. J. Prev. Cardiol.* **29**, 1202–1211. https://doi.org/10.1093/eurjpc/zwaa084 (2022).
7. Dahlquist, M. et al. Short-term ambient air pollution exposure and risk of atrial fibrillation in patients with intracardiac devices. *Environ. Epidemiol.* **6**, 23. https://doi.org/10.1097/ee9.0000000000000215 (2022).
8. Stockfelt, L. et al. Long-term effects of total and source-specific particulate air pollution on incident cardiovascular disease in Gothenburg, Sweden. *Environ. Res.* **158**, 61–71. https://doi.org/10.1016/j.envres.2017.05.036 (2017).
9. Xue, X. et al. Hourly air pollution exposure and the onset of symptomatic arrhythmia: an individual-level case-crossover study in 322 Chinese cities. *Can. Med. Assoc. J.* **195**, E601–E611. https://doi.org/10.1503/cmaj.220929 (2023).
10. Andersen, Z. J. et al. Long-Term exposure to road traffic noise and air pollution, and incident atrial fibrillation in the Danish nurse cohort. *Environ. Health Perspect.* **129**, 36. https://doi.org/10.1289/ehp8090 (2021).
11. Andersson, E. M. et al. Road traffic noise, air pollution and cardiovascular events in a Swedish cohort. *Environ. Res.* **185**, 452. https://doi.org/10.1016/j.envres.2020.109446 (2020).
12. Wang, K. et al. Relationship of neighborhood greenness to heart disease in 249 405 US medicare beneficiaries. *J. Am. Heart Assoc.* **8**, 41. https://doi.org/10.1161/jaha.118.010258 (2019).
13. Roselli, C. et al. Multi-ethnic genome-wide association study for atrial fibrillation. *Nat. Genet.* **50**, 1225. https://doi.org/10.1038/s41588-018-0133-9 (2018).
14. Weng, Z. K. et al. Associations of genetic risk factors and air pollution with incident hypertension among participants in the UK biobank study. *Chemosphere* **299**, 123. https://doi.org/10.1016/j.chemosphere.2022.134398 (2022).
15. Ma, Y. et al. Air pollution, genetic susceptibility, and the risk of atrial fibrillation: a large prospective cohort study. *Proc. Natl. Acad. Sci. U. S. A.* **120**, 745. https://doi.org/10.1073/pnas.2302708120 (2023).
16. Lin, C. W. et al. Risk and economic cost of hospitalization due to atrial fibrillation caused by air pollution: a multi-city time series analysis. *Environ. Sci. Eur.* **35**, 475. https://doi.org/10.1186/s12302-022-00709-w (2023).

17. Jones, J. S. et al. Long-term exposure to low-concentration pm 2.5 and heart disease in older men in Perth, Australia: the health in men study. *Environ. Epidemiol.* **7**, 748. https://doi.org/10.1097/ee9.0000000000000255 (2023).
18. Kim, I. S. et al. Long-term pm < sub > 2.5 exposure and the clinical application of machine learning for predicting incident atrial fibrillation. *Sci. Rep.* **10**, 25. https://doi.org/10.1038/s41598-020-73537-8 (2020).
19. Sun, T. et al. Long-term exposure to air pollution and increased risk of atrial fibrillation prevalence in China. *Int. J. Cardiol.* **378**, 130–137. https://doi.org/10.1016/j.ijcard.2023.02.039 (2023).
20. Zhou, C. B. et al. More Obvious association between short-term ambient nitrogen dioxide and atrial fibrillation outpatient visits in cool seasons: A hospital-based study in Northwestern China. *Environ. Res.* **212**, 12. https://doi.org/10.1016/j.envres.2022.113220 (2022).
21. Liu, C. et al. Application of smart devices in investigating the effects of air pollution on atrial fibrillation onset. *Npj Digit. Med.* **6**, 785. https://doi.org/10.1038/s41746-023-00788-w (2023).
22. Hahad, O. et al. Annoyance to different noise sources is associated with atrial fibrillation in the Gutenberg health study. *Int. J. Cardiol.* **264**, 79–84. https://doi.org/10.1016/j.ijcard.2018.03.126 (2018).
23. Cao, R. et al. The establishment of air quality health index in China: a comparative analysis of methodological approaches. *Environ. Res.* **215**, 23. https://doi.org/10.1016/j.envres.2022.114264 (2022).
24. Lin, Y. J. et al. Long-term exposure to ambient air pollutants and their interaction with physical activity on insomnia: A prospective cohort study. *Environ. Res.* **224**, 74. https://doi.org/10.1016/j.envres.2023.115495 (2023).
25. Bennett, M., Nault, I., Koehle, M. & Wilton, S. Air pollution and arrhythmias. *Can. J. Cardiol.* **39**, 1253–1262. https://doi.org/10.1016/j.cjca.2023.03.023 (2023).
26. Andersen, D. M. et al. Measures of left atrial function predict incident heart failure in a low-risk general population: the Copenhagen City heart study. *Eur. J. Heart Fail.* **24**, 483–493. https://doi.org/10.1002/ejhf.2406 (2022).
27. Chen, S. C. et al. Association of Brachial-Ankle pulse wave velocity with cardiovascular events in atrial fibrillation. *Am. J. Hypertens.* **29**, 348–356. https://doi.org/10.1093/ajh/hpv124 (2016).
28. Ramos-Campo, J. Biochemical responses and physical performance during high-intensity resistance circuit training in hypoxia and normoxia. *Eur. J. Appl. Physiol.* **117**, 809–818. https://doi.org/10.1007/s00421-017-3571-7 (2017).
29. Martin, L. et al. Nature contact, nature connectedness and associations with health, wellbeing and pro-environmental behaviours. *J. Environ. Psychol.* **68**, 36. https://doi.org/10.1016/j.jenvp.2020.101389 (2020).
30. Kondo, M. C., Jacoby, S. F. & South, E. C. Does spending time outdoors reduce stress? A review of real-time stress response to outdoor environments. *Health Place.* **51**, 136–150. https://doi.org/10.1016/j.healthplace.2018.03.001 (2018).
31. Marquet, O. et al. GPS-based activity space exposure to greenness and walkability is associated with increased accelerometer-based physical activity. *Environ. Int.* **165**, 47. https://doi.org/10.1016/j.envint.2022.107317 (2022).
32. Hou, J. et al. Residential greenness attenuated associations of long-term exposure to air pollution with biomarkers of advanced fibrosis. *Environ. Sci. Pollut Res.* **29**, 977–988. https://doi.org/10.1007/s11356-021-15676-7 (2022).
33. Brown, S. C. et al. Neighborhood greenness and chronic health conditions in medicare beneficiaries. *Am. J. Prev. Med.* **51**, 78–89. https://doi.org/10.1016/j.amepre.2016.02.008 (2016).
34. Knezovic, N. J., Memic, M., Mabic, M., Huremovic, J. & Mikulic, I. Correlation between water hardness and cardiovascular diseases in Mostar City, Bosnia and Herzegovina. *J. Water Health.* **12**, 817–823. https://doi.org/10.2166/wh.2014.129 (2014).
35. Bykowska-Derda, A. et al. The relationship between mortality from cardiovascular diseases and total drinking water hardness: systematic review with Meta-Analysis. *Foods* **12**, 74. https://doi.org/10.3390/foods12173255 (2023).
36. Han, B. et al. Associations of exposure to fine particulate matter mass and constituents with systemic inflammation: a Cross-Sectional study of urban older adults in China. *Environ. Sci. Technol.* **56**, 7244–7255. https://doi.org/10.1021/acs.est.1c04488 (2022).
37. Su, K. N. et al. Atrial AMP-activated protein kinase is critical for prevention of dysregulation of electrical excitability and atrial fibrillation. *Jci Insight.* **7**, 475. https://doi.org/10.1172/jci.insight.141213 (2022).
38. Karoui, A. et al. Nitrogen dioxide inhalation exposures induce cardiac mitochondrial reactive oxygen species production, impair mitochondrial function and promote coronary endothelial dysfunction. *Int. J. Environ. Res. Public. Health.* **17**, 742. https://doi.org/10.3390/ijerph17155526 (2020).
39. Swieczkowski, M. et al. Association between exposure to air pollution and increased ischaemic stroke incidence: a retrospective population-based cohort study (EP-PARTICLES study). *Eur. J. Prev. Cardiol.* **32**, 276–287. https://doi.org/10.1093/eurjpc/zwae301 (2024).
40. Thacher, J. D. et al. Long-term exposure to transportation noise and risk for atrial fibrillation: A Danish nationwide cohort study. *Environ. Res.* **207**, 475. https://doi.org/10.1016/j.envres.2021.112167 (2022).
41. Hahad, O. et al. Noise annoyance and risk of prevalent and incident atrial fibrillation-A sex-specific analysis. *Front. Public. Health.* **10**, 475. https://doi.org/10.3389/fpubh.2022.1061328 (2022).
42. Sudlow, C. et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* **12**, 856. https://doi.org/10.1371/journal.pmed.1001779 (2015).
43. Bycroft, C. et al. The UK biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203. https://doi.org/10.1038/s41586-018-0579-z (2018).
44. Choi, S. W., Mak, T. S. H. & O'Reilly, P. F. Tutorial: a guide to performing polygenic risk score analyses. *Nat. Protoc.* **15**, 2759–2772. https://doi.org/10.1038/s41596-020-0353-1 (2020).
45. Schoenfeld, D. Partial residuals for the proportional hazards regression model. *Biometrika* **69**, 239–241. https://doi.org/10.1093/biomet/69.1.239 (1982).
46. Baron, R. M. & Kenny, D. A. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J. Personal. Soc. Psychol.* **51**, 1173–1182. https://doi.org/10.1037/0022-3514.51.6.1173 (1986).
47. Ritchie, K. et al. Designing prevention programmes to reduce incidence of dementia: prospective cohort study of modifiable risk factors. *Br. Med. J.* **341**, 896. https://doi.org/10.1136/bmj.c3885 (2010).

Acknowledgements

Author contributions

D.T., J.S. and M.W. designed the study, oversaw the completion, and revised the manuscript. P.L. and X.Z. did data processing, data analyses, and wrote the manuscript. S.T., Y.L., Y.Q., J.Y., S.C., C.C. and L.Z. contributed to

the data collection, data processing, and manuscript revision. M.L., Y.W., C.H., J.S. participated in the investigation of the article, revision of manuscript, and results visualization. All authors read and approved the final manuscript.

Competing interests
The authors declare no competing interests.

Ethics approval and consent to participate
The UK Biobank study obtained approval from the Northwest Multicentre Research Ethics Committee. Each of the participants offered knowledgeable written informed consent. This study was performed in accordance with the principles of the Declaration of Helsinki.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-00921-7.

Correspondence and requests for materials should be addressed to M.W., J.S. or D.T.

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.