

Air pollution below US regulatory standards and cardiovascular diseases using a double negative control approach

Received: 31 October 2023

Accepted: 23 August 2024

Published online: 30 September 2024

Yichen Wang^{1,2}✉, Mahdiah Danesh Yazdi³, Yaguang Wei^{1,4} & Joel D. Schwartz^{1,5}

Growing evidence suggests that long-term air pollution exposure is a risk factor for cardiovascular mortality and morbidity. However, few studies have investigated air pollution below current regulatory limits, and causal evidence is limited. We use a double negative control approach to examine the association between long-term exposure to air pollution at low concentration and cardiovascular hospitalizations among US Medicare beneficiaries aged ≥ 65 years between 2000 and 2016. The expected values of the negative outcome control (preceding-year hospitalizations) regressed on exposure and negative exposure control (subsequent-year exposure) are treated as a surrogate for omitted confounders. With analyses separately restricted to low-pollution areas ($\text{PM}_{2.5} < 9 \mu\text{g}/\text{m}^3$, $\text{NO}_2 < 75.2 \mu\text{g}/\text{m}^3$ [40 ppb], warm-season $\text{O}_3 < 88.2 \mu\text{g}/\text{m}^3$ [45 ppb]), we observed positive associations of the three pollutants with hospitalization rates of stroke, heart failure, and atrial fibrillation and flutter. The associations generally persisted in demographic subgroups. Stricter national air quality standards should be considered.

Long-term exposure to air pollution has been recognized as an important modifiable risk factor for cardiovascular diseases^{1,2}. An increasing number of epidemiological studies support positive associations between long-term air pollution and the occurrence of cardiovascular events, although specific cardiovascular outcomes have been less investigated relative to overall cardiovascular mortality and morbidity. Stroke, which is characterized by high incidence and mortality, is the second leading cause of death worldwide³. Researchers have reported that long-term exposure to air pollution, particularly fine particulate matter with an aerodynamic diameter $< 2.5 \mu\text{m}$ ($\text{PM}_{2.5}$), could be associated with an increased risk of hospitalization, incidence, and mortality due to stroke⁴. Heart failure (HF) and atrial fibrillation (AF) are other two major cardiovascular diseases. They are important risk factors for stroke onset⁵. Several studies demonstrated the adverse effect of long-term air pollution on the risk of HF^{6–8} and

AF^{9,10}, although these two endpoints have been understudied as primary outcomes of interest. Overall, the evidence for the hypothesized association, especially with HF and AF, remains scarce and inconsistent. In addition, with the predominant focus on $\text{PM}_{2.5}$, the potential cardiovascular effects of long-term exposure to other major air pollutants such as nitrogen dioxide (NO_2) and ozone (O_3) have been under-examined, and the correlations between different pollutants have also been overlooked. To conclude, the potential causal relationships between multiple air pollutants and specific cardiovascular events need to be further elucidated.

Most of the existing studies linking long-term exposure to air pollution to cardiovascular events examined the entire range of exposure. The average pollution levels may differ substantially by region and therefore partially account for the geographical differences in the estimated associations. Average pollution level differences

¹Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, MA, USA. ²School of the Environment, Yale University, New Haven, CT, USA. ³Program in Public Health, Department of Family, Population, & Preventive Medicine, Stony Brook University, Stony Brook, NY, USA. ⁴Department of Environmental Medicine and Climate Science, Icahn School of Medicine at Mount Sinai, New York, NY, USA. ⁵Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA. ✉e-mail: yichen.wang.yw@yale.edu

could only explain geographical differences if the association is not linear. There is a dearth of our understanding of the health impacts of air pollution at concentrations below regulatory standards, which has important implications for air pollution regulations in regions such as the United States (US) where populations experience generally low air pollutant exposures. Previous studies found the shape of the exposure–response (E–R) curves for long-term PM_{2.5} and all-cause and cardiovascular mortality to be curvilinear with no evidence of a threshold^{11,12}. According to several studies of large cohorts in the US^{9,13} and Europe^{14,15}, the risk of cardiovascular diseases per mass unit could persist and even become stronger at lower exposure levels below the annual limit values recommended by the US Environmental Protection Agency (EPA; 9 µg/m³ for PM_{2.5} and 53 ppb [99.6 µg/m³] for NO₂) and European Union (10 µg/m³ for PM_{2.5} and 20 µg/m³ for NO₂). Further research specifically at lower concentrations can help quantify the disease burden attributable to low-level air pollution and elucidate the true “safe” level. This could further inform recommendations for even stricter air quality guidelines, as suggested by the World Health Organization (5 µg/m³ for PM_{2.5} and 10 µg/m³ for NO₂).

Furthermore, all observational studies are subject to the possibility of confounding by omitted variables, and causal modeling methods that can capture some omitted confounders are therefore valuable. Propensity scores are widely adopted in air pollution research by balancing measured covariates across different levels of continuous exposure. However, this method is weakened by its stringent requirement for precisely specified regression of exposure on

measured covariates and its inability to control for unmeasured covariates. Negative controls have been suggested as a useful tool to enhance causal inference independently of covariate distributions and to tackle unmeasured confounding bias¹⁶. The negative exposure control is a variable known not to be causally related to the outcome of interest, while the negative outcome control is a variable known not to be caused by the exposure of interest. Both of them may share a common confounding mechanism with the exposure and outcome¹⁷. Therefore, they can serve as instruments for reducing bias by unmeasured confounders. In prior air pollution and health studies, researchers have used future air pollution as a negative control exposure^{18–21}, or a negative outcome due to causes other than primary exposure as a negative control outcome^{22,23}. More recently, double negative control adjustment has been employed to strengthen causal inference in studies examining short- and long-term effects of air pollution^{24–26}.

To address the research gaps, the present study used a double negative control approach to analyze the relationships between long-term exposure to PM_{2.5}, NO₂, and warm-season O₃ at low concentrations with risk of hospitalizations for three major cardiovascular diseases (stroke, HF, and AF) in the Medicare population aged ≥65 years across the contiguous US from 2000 to 2016. We focused on the areas where populations were consistently exposed to low pollutant concentration levels (PM_{2.5} < 9 µg/m³, NO₂ < 75.2 [40 ppb], warm-season O₃ < 88.2 µg/m³ [45 ppb]). Furthermore, we conducted stratified analyses to investigate potential susceptible demographic subpopulations.

Table 1 | Summary of ZIP code-level air pollution, meteorological covariates, and socioeconomic status (SES) covariates in the low pollution areas from 2000 through 2016

Covariates	PM _{2.5} < 9 µg/m ³	NO ₂ < 75.2 µg/m ³	Warm-season O ₃ < 88.2 µg/m ³
N (ZIP code)	4086	26,415	3708
N (beneficiaries)	7,988,816	63,639,732	12,992,320
Air pollution concentration			
PM _{2.5} (µg/m ³)	5.2 (1.6)	9.6 (3.0)	7.8 (2.8)
NO ₂ (µg/m ³)	23.2 (14.7)	27.6 (13.4)	29.7 (17.7)
Warm-season O ₃ (µg/m ³)	88.8 (14.5)	88.2 (10.6)	73.1 (7.7)
Meteorological covariates			
Summer average temperature (°C)	16.9 (4.0)	20.2 (3.7)	19.1 (5.1)
Winter average temperature (°C)	3.1 (6.0)	6.2 (5.8)	7.5 (8.7)
Summer average RH (%)	56.4 (14.7)	66.7 (9.4)	68.9 (7.1)
Winter average RH (%)	66.3 (11.0)	66.9 (7.1)	70.0 (7.2)
SES covariates			
Percent Black (%)	1.1 (3.1)	8.5 (16.0)	6.5 (13.8)
Percent Hispanic (%)	10.5 (16.7)	7.8 (14.4)	14.8 (22.2)
Median household income (\$)	46,759 (17,253)	48,419 (20,204)	52,232 (23,002)
Median house value (\$)	175,012 (133,349)	149,969 (124,583)	230,709 (191,296)
Percent owner occupied (%)	75.8 (11.3)	74.2 (14.0)	69.1 (18.5)
Percent education <high school (%)	22.0 (14.8)	28.1 (16.1)	24.4 (16.4)
Population density (persons/mi ²)	225 (751)	840 (2089)	3445 (10,150)
Percent ≥65 below poverty (%)	9.4 (7.9)	10.1 (7.7)	10.3 (8.4)
Percent annual HbA1c test (%)	82.0 (9.5)	83.5 (5.9)	84.6 (4.8)
Percent ambulatory visit (%)	78.1 (7.4)	79.8 (6.0)	76.6 (7.2)
Percent eye exam (%)	68.9 (7.7)	67.3 (6.7)	70.5 (6.3)
Percent LDL test (%)	75.6 (10.7)	78.6 (7.3)	80.6 (5.5)
Percent mammogram (%)	64.6 (9.0)	64.1 (7.2)	66.1 (7.3)
Distance to nearest hospital (km)	18.2 (15.9)	12.2 (10.6)	10.2 (10.6)
Lung cancer rate (‰)	0.4 (5.7)	0.4 (2.2)	0.4 (6.0)
Ever smokers (%)	48.0 (7.7)	47.3 (7.4)	47.6 (7.7)
Mean BMI (kg/m ²)	28.1 (3.1)	28.1 (2.5)	27.4 (1.7)

Numbers in the table are presented as Mean (Standard Deviation) for ZIP code-level covariates.

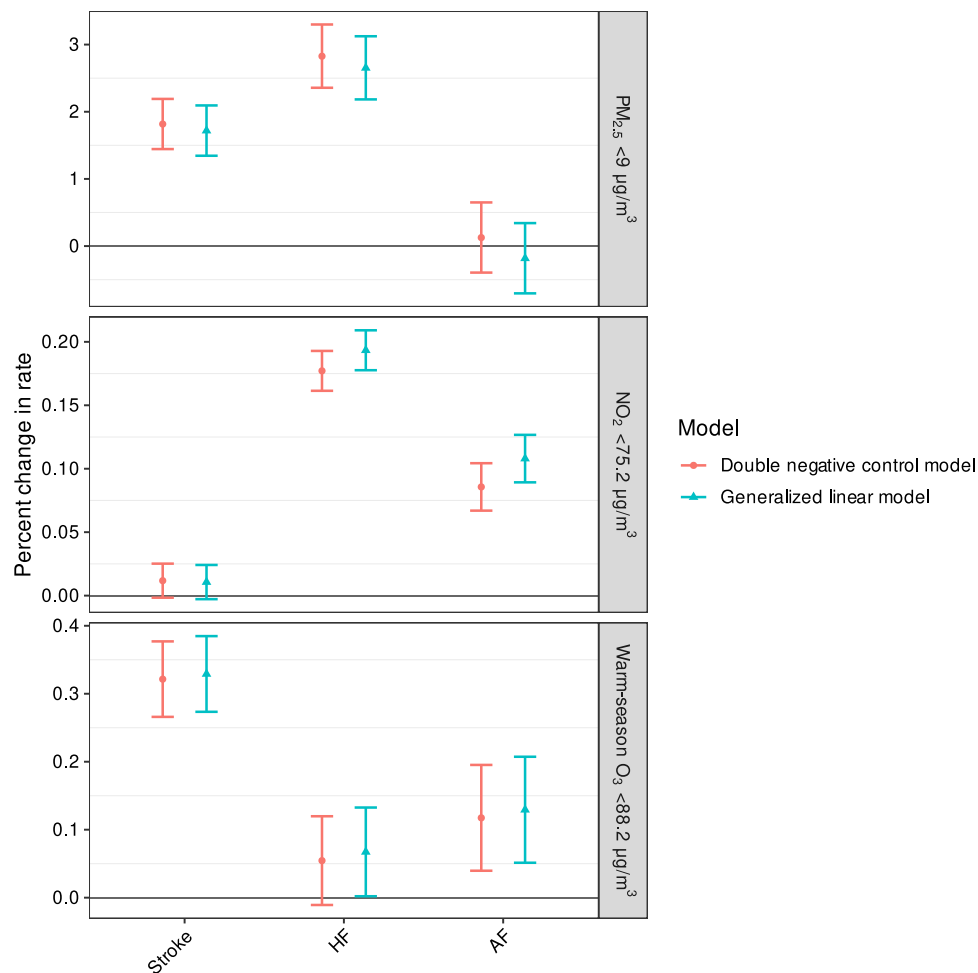


Fig. 1 | Percent change in hospitalization rate for cardiovascular diseases associated with 1-µg/m³ increase in long-term exposure to air pollution at low concentrations using double negative control models and generalized linear

models adjusted for co-pollutants. Error bars indicate the 95% confidence intervals. Source data are provided as a Source Data file.

Results

Table 1 shows the summary statistics of ZIP code-level air pollution and covariates in the low-pollution areas from 2000 through 2016. In low PM_{2.5} areas, the annual average concentrations of PM_{2.5}, NO₂, and warm-season O₃ were 5.2 ± 1.6 µg/m³, 23.2 ± 14.7 µg/m³, and 88.8 ± 14.5 µg/m³, respectively. In the areas with either NO₂ or O₃ deemed low in our analyses, the mean annual PM_{2.5} concentration was higher and closer to the typical range. The Pearson correlation coefficients (*r*) for three air pollutants are presented in Supplementary Table 1. We observed a moderate-to-low positive correlation between annual PM_{2.5} and NO₂ in low NO₂ areas (*r* = 0.38) and in low PM_{2.5} areas (*r* = 0.13). In contrast, there was a strong correlation between annual PM_{2.5} and NO₂ in areas with low warm-season O₃ (*r* = 0.66). Warm-season O₃ exhibited a moderate-to-low correlation with both annual PM_{2.5} and NO₂ in areas with low levels of PM_{2.5} and NO₂, while in areas with lower warm-season O₃, the correlations were negligible.

Supplementary Table 2 presents the total number of hospitalizations and the annual rate for stroke, HF, and AF in the low pollution areas during the study period. The annual hospitalization rate for stroke, HF, and AF among the Medicare participants were 0.87%, 0.84%, and 0.41%, respectively, in low PM_{2.5} areas where low NO₂ and warm-season O₃ exposures concurrently occurred. The corresponding hospitalization rates were similar in low O₃ areas. However, the hospitalization rates were higher in low NO₂ areas where people experienced more normal PM_{2.5} exposures. Nevertheless, the pattern of the hospitalization rates for each cardiovascular outcome within

demographic groups was generally similar across all the defined low-pollution areas. Overall, we observed higher annual hospitalization rates for stroke and HF among those aged 85 years and older and eligible for Medicaid. However, there were some inconsistencies in the pattern by sex and race across specific outcomes. While the annual hospitalization rate for stroke and HF was higher in males and black individuals, this was not seen for AF.

We compared the estimated associations of long-term exposures to PM_{2.5}, NO₂, and warm-season O₃ at low concentrations with the rates of hospitalizations for stroke, HF, and AF as determined from three-pollutant double negative control models and GLM (Fig. 1). The results from single-pollutant models are illustrated in Supplementary Fig. 1. Overall, the adjustments for co-pollutants resulted in stronger estimates for PM_{2.5}, while those for NO₂ and warm-season O₃ remained similar. When examining the associations between PM_{2.5} and all three outcomes, we found that the GLM yielded estimates that were modestly comparable but lower than those derived from the double negative control models. While both modeling approaches produced relatively similar estimates for the associations of NO₂ and warm-season O₃ with AF, there were slight differences in the estimates for stroke and HF. All the numeric results can be found in Supplementary Table 3.

In this study, we focused on the results adjusted for co-pollutants using double negative control adjustment. For long-term PM_{2.5} exposure below 9 µg/m³, each 1-µg/m³ increase in PM_{2.5} was associated with the percent increases of 1.82% (95% confidence interval [CI]: 1.44%,

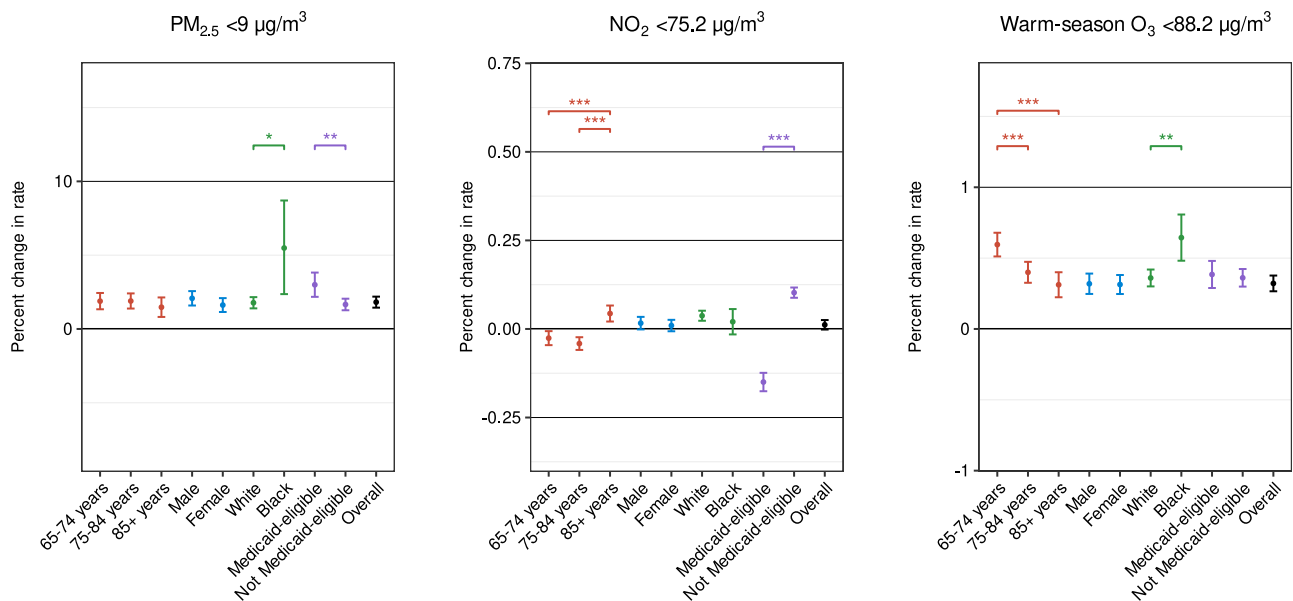


Fig. 2 | Percent change in hospitalization rate for stroke associated with 1- $\mu\text{g}/\text{m}^3$ increase in long-term exposure to air pollution at low concentrations in stratified analyses in three-pollutant double negative control models.

Statistical significance is calculated via a two-tailed t-test (* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$). Error bars indicate the 95% confidence intervals. Source data are provided as a Source Data file.

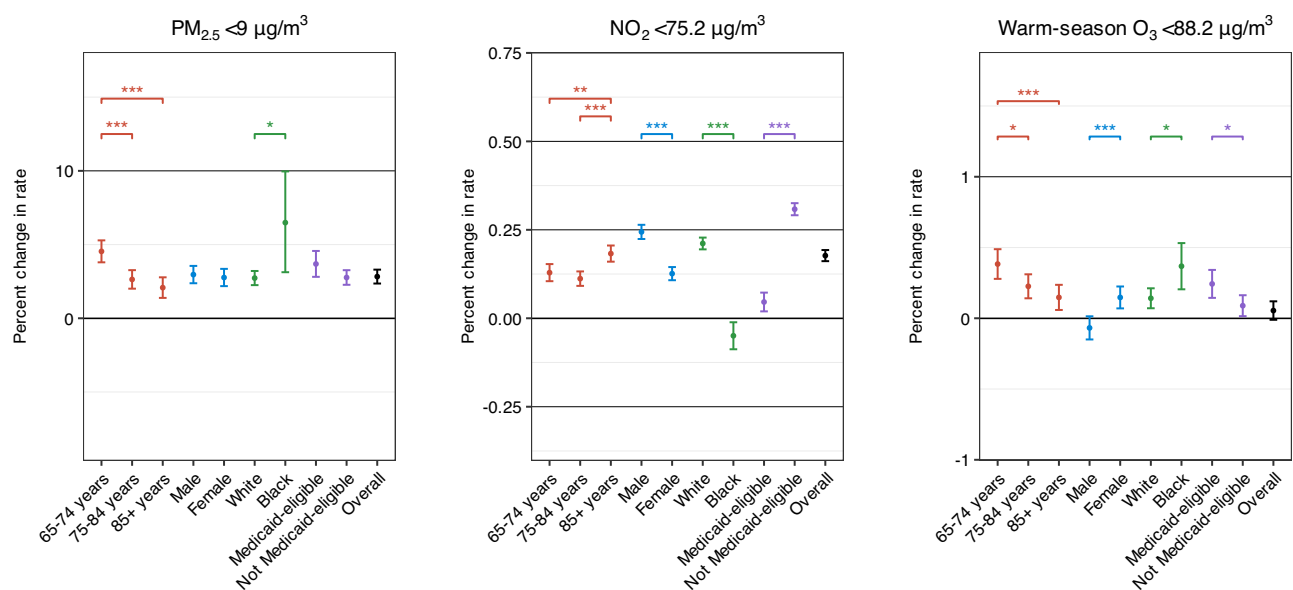


Fig. 3 | Percent change in hospitalization rate for heart failure associated with 1- $\mu\text{g}/\text{m}^3$ increase in long-term exposure to air pollution at low concentrations in stratified analyses in three-pollutant models using double negative control

adjustment. Statistical significance is calculated via a two-tailed t-test (* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$). Error bars indicate the 95% confidence intervals. Source data are provided as a Source Data file.

2.19%), 2.83% (95% CI: 2.36%, 3.30%), and 0.13% (95% CI: -0.39%, 0.65%) in the hospitalization rates for stroke, HF, and AF, respectively. For each 1- $\mu\text{g}/\text{m}^3$ increase in annual NO_2 below 75.2 $\mu\text{g}/\text{m}^3$, the percent increases in the hospitalization rates for stroke, HF, and AF were 0.01% (95% CI: -0.002%, 0.03%), 0.18% (95% CI: 0.16%, 0.19%), and 0.09% (95% CI: 0.07%, 0.10%), respectively. For long-term exposure to warm-season O_3 below 88.2 $\mu\text{g}/\text{m}^3$, we found adverse associations with the three outcomes with percent increases in the hospitalization rates of 0.32% (95% CI: 0.27%, 0.38%), 0.05% (95% CI: -0.01%, 0.12%), and 0.12% (95% CI: 0.04%, 0.20%) per 1- $\mu\text{g}/\text{m}^3$ increase in warm-season O_3 . The estimates remained very similar after excluding additional confounding adjustments from the prediction model for the negative outcome control (Supplementary Table 4).

We conducted stratified analyses by individual demographic characteristics to identify the vulnerable subgroups. The results of the stratified analyses for stroke, HF, and AF from three-pollutant models are shown in Figs. 2–4, respectively. The observed positive associations in the overall analyses generally persisted in demographic subgroups. Similar patterns of the potential effect modification by demographics were found in both single- and three-pollutant models, despite some differences in the magnitude and statistical significance of the subgroup-specific effect estimates (Supplementary Figs. 2–4 and Tables 5–7).

We found a significantly larger effect of long-term $\text{PM}_{2.5}$ below 9 $\mu\text{g}/\text{m}^3$ on all three outcomes for black people compared to white people. In the association of $\text{PM}_{2.5}$ with stroke and AF, we identified

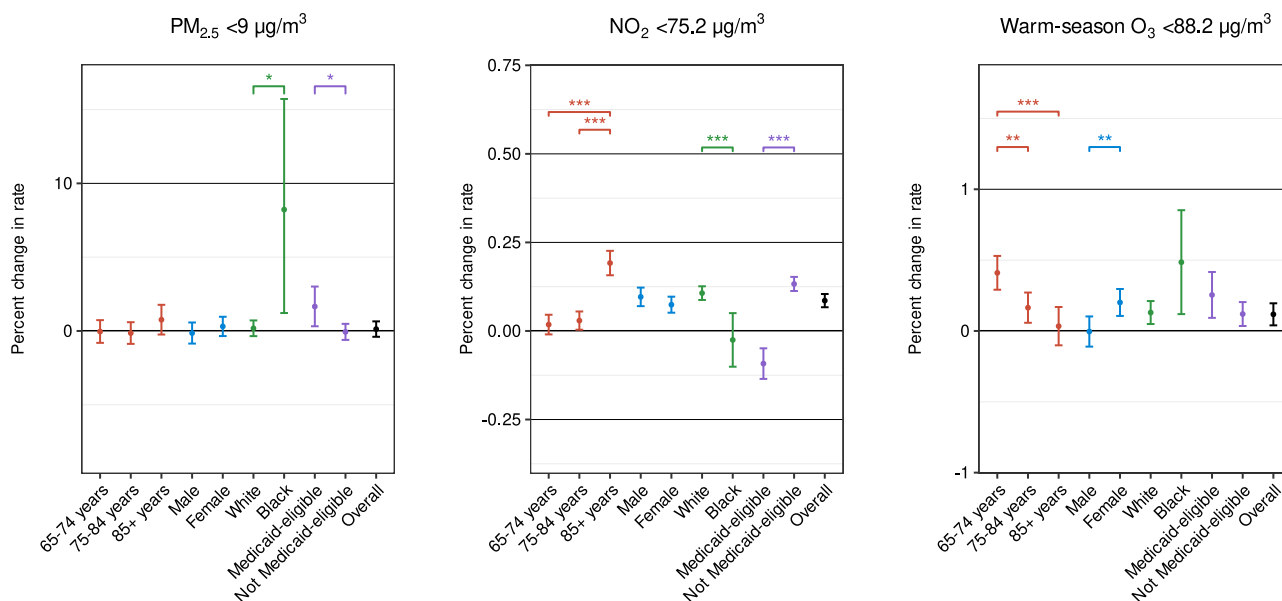


Fig. 4 | Percent change in hospitalization rate for atrial fibrillation and flutter associated with 1- $\mu\text{g}/\text{m}^3$ increase in long-term exposure to air pollution at low concentrations in stratified analyses in three-pollutant models using double

negative control adjustment. Statistical significance is calculated via a two-tailed t-test (* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$). Error bars indicate the 95% confidence intervals. Source data are provided as a Source Data file.

Medicaid eligibility as a significant modifier, with a higher risk seen in individuals who were eligible for Medicaid than those who were not. In addition, age modified the $\text{PM}_{2.5}$ association for HF with a stronger effect in the younger group (aged 65–74 years), but this modification pattern was not observed for stroke or AF. In contrast, we found no evidence of any effect modification by sex on the association of all outcomes in relation to $\text{PM}_{2.5}$.

For long-term exposure to NO_2 below $75.2 \mu\text{g}/\text{m}^3$, individuals aged over 84 years and those who were not Medicaid-eligible were at greater risk of stroke. We observed similar effect modification patterns by age and Medicaid eligibility in the associations of HF and AF with NO_2 . Regarding the modification by sex, males were at greater NO_2 -associated risk of HF compared to females. At the same time, white people exhibited a significantly higher NO_2 -associated risk of HF and AF compared to black people.

In terms of long-term exposure to warm-season O_3 below $88.2 \mu\text{g}/\text{m}^3$, individuals aged 65–74 years had greater risks of all three outcomes compared to older age groups. Black individuals were found to be more susceptible to stroke and HF, while females were more susceptible to HF and AF. Additionally, individuals eligible for Medicaid were at greater risk of HF compared to those who were not.

The E–R curves for the main associations from three-pollutant GLM models with natural spline function are provided in Supplementary Fig. 5. For the relationship between $\text{PM}_{2.5}$ with stroke and HF, a positive association for stroke and HF was apparent down to the lowest concentrations. The effect size for AF was more complex, with a positive association beyond levels of $5 \mu\text{g}/\text{m}^3$, but negative at lower concentrations. The E–R curves for NO_2 displayed an almost positively linear shape for HF with a steeper slope below $20 \mu\text{g}/\text{m}^3$, however, appeared negative before linearly increasing at around $25 \mu\text{g}/\text{m}^3$ for stroke and AF. For warm-season O_3 , the E–R curve showed a linear positive association with stroke, while the curves for HF and AF depicted a non-linear relationship, with effects increasing when approaching the highest concentrations.

Discussion

Among US Medicare participants, we found that long-term exposure to low-level $\text{PM}_{2.5}$ ($<9 \mu\text{g}/\text{m}^3$), NO_2 ($<75.2 \mu\text{g}/\text{m}^3$), and warm-season O_3 ($<88.2 \mu\text{g}/\text{m}^3$) could be positively associated with increased rate of

hospitalizations for stroke in three-pollutant models that accounted for correlations between co-existing air pollutants and controlled for unmeasured confounders using negative controls. $\text{PM}_{2.5}$ and NO_2 were most strongly associated with HF, whereas the strongest effect of warm-season O_3 was seen on stroke. Black people and Medicaid-eligible people appeared to be more vulnerable to the cardiovascular risk attributable to $\text{PM}_{2.5}$ and warm-season O_3 . The youngest-old and females were also found to be more vulnerable to the warm-season O_3 -related risk. However, the NO_2 -related risk showed contradictory effect modification patterns.

We designed a pair of negative control exposure and outcome variables in an attempt to capture uncontrolled confounding. The selection of negative exposure control (or negative outcome control) relies on the absence of a causal relationship with the true outcome (or exposure) due to chronological order. However, correlations between them are likely present due to their relations with unmeasured confounders. If the assumption of linearity between unmeasured covariates with exposure and negative exposure control holds, regressing the negative outcome control on the exposure and negative exposure control is expected to reveal confounded associations. Therefore, adjusting for the expected counts of hospitalization in the preceding year aims to capture confounding bias and strengthen the causal interpretation of our observed associations. Given that concurrent-year air pollution exposure and subsequent-year exposure can be highly correlated, an alternative assumption of the two variables sharing the same magnitude of a correlation with the omitted confounders is rendered more reasonable. The GLM method yielded comparable results with the double negative control approach, exhibiting only slight differences in the effect size estimates. Such discrepancies may be attributable to unadjusted confounding bias. The consistent findings derived from these two statistical methods demonstrate the robustness of our results to different model adjustments, suggesting that any omitted confounding bias is small, and negative in the case of $\text{PM}_{2.5}$ but positive for NO_2 . A previous study reported that greater control for SES resulted in increased effect sizes for $\text{PM}_{2.5}$ ²⁷.

Our study has a special emphasis on long-term exposure to low-level air pollution below the annual US EPA limits. While a growing number of prior studies have revealed increased health risks at lower

levels of air pollution exposure under regulatory standards, most have focused on all-cause and cardiovascular mortality^{11,12,28,29}. However, the available evidence concerning cardiovascular disease risk at these lower pollution levels remains limited. In consistence with our findings, several previous studies of the Medicare population have found a steeper E–R curve for a range of cardiovascular outcomes when restricted to lower exposures^{9,13,26,30}. Specifically, we found no threshold below which cardiovascular effects were absent in our E–R curves for PM_{2.5} in relation to stroke and HF. In addition, the positive curves for NO₂ and HF displayed a steeper increase in risk below 20 µg/m³. Our E–R curves for warm-season O₃ also showed an increasing tendency at higher concentrations, which was the most pronounced for stroke second to AF. These findings indicate that substantial health benefits can likely be obtained by lowering ambient air pollution levels even at low concentrations. Similarly, in a large population-based Canadian cohort, Bai et al.⁷ found the concentration–response curves for congestive HF with long-term exposure to PM_{2.5} and NO₂ to be supralinear with no discernable threshold values. They also observed a sublinear relationship for O₃ with an indicative threshold. A meta-analysis of 102 coefficients from 53 cohort studies reporting associations with all-cause or cause-specific mortality found a steeper E–R curve at lower PM_{2.5} concentrations for cardiovascular mortality, which also supports these findings¹². Overall, our findings indicate the need to reassess the current air quality guidelines and tighten pollution control policies and measures.

This study also supplemented the limited epidemiologic evidence regarding the long-term effects of multiple air pollutants on cause-specific cardiovascular morbidity. Our findings of adverse associations are in accordance with some of the existing literature, despite some slight difference in the statistical significance. Prior studies of the Medicare population using diverse methodologies and different ranges of exposure have reported significant positive associations of our studied outcomes with annual PM_{2.5}, NO₂, and warm-season O₃^{9,13,31}. A review and meta-analysis identified five studies of long-term exposure to PM_{2.5} and stroke incidence from North America and Europe and found a 6.4% (95% CI: 2.1%, 10.9%) increase in the hazard for each 5-µg/m³ increase in PM_{2.5}³². A more recent review article reported that each 10-µg/m³ increase in long-term PM_{2.5} exposure could be associated with an increased risk of 13% (95% CI: 11%, 15%) for incident stroke, synthesizing the results of fourteen studies across the globe³³. Additionally, other studies conducted in Canada^{7,34}, the UK^{6,8}, and Sweden³⁵ indicated an increased risk of HF associated with PM_{2.5} at relatively low exposures. According to state-of-the-art evidence, the odds ratios of HF associated with each 10-µg/m³ increase in long-term PM_{2.5} and NO₂ exposure were estimated to be 1.019 (95% CI: 1.008–1.030) and 1.012 (95% CI: 1.007–1.017), respectively³⁶. Yue et al.¹⁰ conducted a systematic review and meta-analysis to quantify the association between air pollutants and AF based on eighteen studies. They indicated that exposure to all air pollutants including PM_{2.5}, NO₂, and O₃ had a deleterious impact on AF onset in the general population. By contrast, several other studies reported null relationships between air pollution and the risk of these outcomes^{37–40}.

It is worth noting that direct comparisons across these studies might be challenging because of potentially heterogeneous air pollution ranges and diverse demographic characteristics of study populations. In addition, in the presence of correlations between air pollutants, considering the influence of co-pollutants in the air pollution mixture is crucial for the validity of the estimated association. Recent studies have increasingly emphasized the importance of utilizing multi-pollutant models to better disentangle the individual effect of a certain pollutant^{6,13,31,35,37}, which is the most widely used way to adjust for confounding bias by co-pollutants⁴¹. In contrast, results obtained from single-pollutant models in other studies are more likely confounded by the impact of other pollutants that share similar sources with the pollutant under investigation^{7,34,39,40}. Therefore, our

use of multi-pollutant models likely yields more accurate estimates in describing the individual cardiovascular effect of each air pollutant.

Multiple pathophysiological mechanisms have been proposed to explain the detrimental cardiovascular effects of air pollution. It is widely accepted that air pollution can trigger systemic inflammation, oxidative stress reactions, and dysfunction of the autonomic nervous system¹. The autonomic imbalance can further result in increases in cardiac frequency and arterial pressure, and a reduction in heart rate variability⁴². Numerous experimental studies have demonstrated that these responses may further instigate endothelial dysfunction, atherosclerosis, and vascular dysfunction^{42,43}. Another plausible mechanism underlying the onset of cardiovascular diseases is that inhaled irritants can traverse the pulmonary epithelium and directly enter the blood circulation and cardiac organs, which may alter blood coagulability and contribute to thrombus formation⁴⁴. A higher PM_{2.5}- and NO₂-associated risk appeared to be seen for HF hospitalization possibly because it was the common consequence of most cardiovascular diseases, especially for elderly people.

Environmental justice is an increasing concern and we found evidence that independent of differences in exposure, some disadvantaged groups had worse responses to any given level of air pollution. Specifically, we identified Medicaid eligibility as a positive modifier of the association of low-level PM_{2.5} and warm-season O₃ with at least one studied cardiovascular outcome. This suggests a greater vulnerability for lower-SES individuals even when residing in low-pollution regions, as Medicaid coverage is provided for low-income elderly beneficiaries to expand their healthcare access⁴⁵. Low SES has been determined as a significant risk factor for cardiovascular diseases because socio-economically disadvantaged individuals tend to have poorer health, higher psychosocial stress, and a propensity for unhealthy behaviors and lifestyles⁴⁶. In addition to Medicaid eligibility, we found that the effect sizes for effects of PM_{2.5} and warm-season O₃ on all outcomes were more pronounced for Black individuals compared to white individuals. The tendency of a higher susceptibility among Blacks is consistent with much of the existing evidence^{13,47}. Black populations have been disproportionately affected by the detrimental health impacts of historic discrimination and ongoing racial segregation, and this study demonstrates additional susceptibility to air pollution. Additionally, while we observed increased susceptibility to warm-season O₃ in individuals aged 65–74 years, the specific underlying reasons for this pattern remain unclear. It is likely that a lower baseline risk in this age group may influence these findings. We also found greater susceptibility among females compared to males, which is similar to the sex difference reported in some previous studies and could be explained by physiological differences^{48,49}. This pattern and the specific reasons are worth attention in future research.

In terms of the adverse effects of NO₂, our results indicated that people aged ≥85 years, males, white people, and those who were not Medicaid-eligible may be more vulnerable to at least one cardiovascular disease we studied. First, an increased risk in the oldest group is understandable, given that advanced age significantly drives the deterioration of cardiovascular functionality in older people⁵⁰. Relative to age differences, sex as a potential modifier of cardiovascular risk in relation to air pollution as well as the relevant biological mechanisms has been more underappreciated. While some researchers found a more prominent NO₂-attributed cardiovascular risk among males^{51,52}, which is comparable to our finding for HF, there is no consensus on this question^{53,54}. Our findings of a higher susceptibility among the very elderly and males are not conclusive, but we think that paying more attention to these questions can be meaningful to improve the distribution of preventive medical care in the future. Interestingly, when we looked at the modification by race and Medicaid eligibility, the greater susceptibility for NO₂ seen in white individuals and non-Medicaid eligible individuals contrasts with our findings for PM_{2.5} and warm-season O₃. Such inconsistent results in the modifying roles of

demographics and SES exist in the literature examining the association between air pollution and cardiovascular health, which may have to do with specific air pollutants and outcomes^{9,55,56}. For example, a study utilizing US nationwide survey data found an adverse association between PM_{2.5} and hypertension among non-Hispanic white adults but a nearly null association among non-Hispanic black and Hispanic adults, although the latter two groups are generally thought to experience higher exposure and are more vulnerable⁵⁶. Moreover, there is also a controversy over the presence and direction of modification by community-level SES in the current empirical-based literature, which could be explained by discrepancies in underlying vulnerable factors in diverse neighborhood samples^{55,57–59}. In fact, the specious modification patterns we found for NO₂ are unlikely but still possible. As a pollutant predominantly coming from urban origins and often transported on a local scale, NO₂ can vary by urbanicity level⁶⁰. It is reasonable to assume that NO₂ might be more of a proxy for commercial activities, since its emissions from other major sources (e.g., diesel traffic, fuel combustion, power plants) have been reduced in recent years^{61,62}. Therefore, the observed higher vulnerability in white and non-Medicaid eligible individuals might be partially accounted for by their higher access to urbanization or commercial activities. In addition, we should also note that our estimate is a measure relative to the baseline risk and does not necessarily represent the magnitude of its absolute attributable risk. For example, the lower baseline risk of hospitalization rates in non-Medicaid eligible beneficiaries might have exaggerated the magnitude of relative risk, although the difference in rate is unlikely to be the major explanation.

Our study has multiple strengths. Foremost is the use of a double negative control approach. This methodology provides an alternative tool to instrumental variables to control for omitted confounding and thus enhance the credibility of the estimated associations. We also thoroughly considered a variety of cardiovascular risk factors to reinforce the confounding adjustment. Another notable strength is that we leveraged the data from the Medicare population. The data that we used was from a very large nationwide cohort, which ensured sufficient statistical power and increased the generalizability of our results to the population that suffers over three quarters of the deaths in the US. Furthermore, the exposure data were derived from high-quality models with a fine resolution and satisfactory predictive accuracy, further assuring the reliability of our analyses. Moreover, compared to restricting the analyses to low exposures in ZIP code-year combinations in prior Medicare studies^{13,63}, the selection criteria applied in this study are somewhat more rigorous by imposing low-exposure constraints over the 17-year study duration. Hence, the possibility of mistakenly including the individuals impacted by past higher exposures was reduced, despite that the exposure history due to migration and travel patterns was not fully accounted for. Last, we attempted to address the correlations among air pollutants and more accurately estimate the independent effect of each exposure by constructing both single- and three-pollutant models.

Some limitations of this study should also be cautioned. First, we may not generalize the conclusions to younger populations or highly polluted regions. Second, there could be residual or unmeasured confounding by omitted cardiovascular risk factors such as diet when the assumption of the same magnitude of linear correlations of them with true exposure and subsequent-year exposure is violated. However, we considered a series of major confounders, ranging from possible meteorological conditions, and area-level health behavioral factors, to socioeconomic measures, which should have captured most of the confounding bias. It is noteworthy that we controlled for co-exposures to other air pollutants using the three-pollutant models as well. Admittedly, the moderate correlation between annual PM_{2.5} and NO₂ concentrations may indicate potential collinearity and the risk of over-controlling issues. Third, the ZIP code-level air pollution data derived from exposure models may not fully represent true personal

exposures. Specifically, our exposure metrics did not account for the exposures occurring distant from the participants' residences. However, the National Human Activity Pattern Survey reported that US adults spent 69% of their time at home and 8% of the time immediately outside their home⁶⁴. Older people may spend even more time at home, implying that the exposure misclassification would be relatively minor. Another concern is that the variations in personal exposures caused by different indoor activity patterns and building features might not be captured by the neighborhood metrics. Nevertheless, the resulting error is likely a Berksonian exposure error and may cause little bias⁶⁵. In addition, ambient concentration serves as an instrumental variable for personal exposure and thus personal behavioral factors which were not available would not confound the association between ambient exposure and outcomes⁶⁶. That is neighborhood level pollution can be correlated with neighborhood level covariates, but if e.g. neighborhoods with a high intake of saturated fats had higher exposure to a pollutant, a vegetarian living in the neighborhood would get the same ambient exposure, despite not eating any saturated fats. Therefore, the confounding is with neighborhood characteristics, not individual ones. Some residual prediction errors of exposure models may be present, but they should be minimal because we studied low air pollutant concentrations. Last, we accessed hospital discharge diagnoses from the administrative Medicare database as the morbidity measure, which may not capture some cases with milder symptoms. This outcome classification might be differential because it can be related to SES factors such as healthcare accessibility.

In conclusion, using a double negative control approach, we found positive associations of long-term exposure to PM_{2.5}, NO₂, warm-season O₃ at low concentrations with the hospitalization rate of stroke, HF, and AF in US Medicare older adults. Our findings suggest that the current National Ambient Air Quality Standards (NAAQS) for annual PM_{2.5} and NO₂ may not be adequate to minimize the cardiovascular disease burden. Future guidelines for warm-season O₃ could be warranted.

Methods

Study population and outcome assessment

We used data from a national cohort of fee-for-service (FFS) Medicare beneficiaries aged 65 years and older across the contiguous US from January 1st, 2000 to December 31st, 2016. The beneficiaries were followed up from January 1st of the year after their Medicare enrollment until the development of the outcome of interest, death, censoring, or the end of the follow-up time. In this study, we restricted the analyses to the individuals who were consistently exposed to low-level annual air pollution for the entire period (2000–2016) with certain thresholds (PM_{2.5} < 9 µg/m³, NO₂ < 75.2 µg/m³ [40 ppb], warm-season O₃ < 88.2 µg/m³ [45 ppb]). Therefore, separate datasets were created for each pollutant according to its specified threshold. We further restricted the datasets to ZIP code areas with more than 100 beneficiaries.

Beneficiary records were provided by the Medicare denominator file from the Centers for Medicare and Medicaid Services, which contained information on age, self-reported sex, self-reported race, Medicaid eligibility, date of death, and residential ZIP code for each beneficiary. Information on age, Medicaid eligibility, and residential ZIP code are updated each year. We obtained the hospital discharge claims of Medicare enrollees from the Medicare Provider Analysis and Review (MEDPAR) file. The International Classification of Diseases (ICD) codes were used to identify the primary discharge diagnosis for each of our three cardiovascular outcomes of interest: stroke (ICD-9 codes: 430–438, ICD-10 codes: I60–I69), heart failure (ICD-9 code: 428, ICD-10 code: I50; hereafter referred to as HF), and atrial fibrillation and flutter (ICD-9 code: 427.3, ICD-10 code: I48; hereafter referred to as AF). For each cardiovascular outcome, we computed the ZIP code-level annual counts based on the beneficiaries' residential

addresses. All hospitalizations from each beneficiary occurring after enrollment during the multi-year follow-up period were counted as cases.

This study was approved by the institutional review board at Harvard T. H. Chan School of Public Health. Our study was exempt from consent requirements as it is considered non-human subject research.

Exposure assessment

We obtained the daily concentrations of ambient PM_{2.5}, NO₂, and O₃ at 1 km × 1 km spatial resolution across the contiguous US from three ensemble prediction models that combined multiple machine learning algorithms^{67,68,69}. The exposure models incorporated meteorological variables, chemical transport model simulations, land-use features, and satellite remote sensing data. They were well validated using 10-fold cross-validation. We aggregated the daily predictions of PM_{2.5} and NO₂ to annual averages. For long-term O₃, we calculated its warm-season levels based on the daily predictions from April 1st through September 30th, since the health impacts of O₃ are suggested to be more observable during warm seasons compared to throughout the year^{13,31,54}. We then computed the ZIP code-level exposures by averaging the 1 km × 1 km grid cell predictions whose centroids were within the boundary of ZIP code polygons or assigning the nearest grid cell predictions for the ZIP codes that do not have polygon representations. Annual average exposures were then linked to Medicare beneficiaries based on their residential ZIP codes for each calendar year over the study period.

For each exposure, we limited our dataset to the ZIP code areas where the populations were always exposed to low-concentration air pollution below thresholds we set over the study period of 2000–2016. The threshold was determined for each pollutant individually to more directly assess its specific health effects and allow for greater statistical power given the different distributions of exposures for the different pollutants. We chose 9 µg/m³ as the threshold for annual average PM_{2.5} concentration, which is the latest limit set by the US EPA on February 7, 2024 to substitute the previous NAAQS of 12 µg/m³. For NO₂, we chose an annual limit of 75.2 µg/m³ [40 ppb] for our analysis, well below the NAAQS of 99.6 µg/m³ [53 ppb], as the annual NO₂ concentrations in the US rarely exceeded this standard. Although there is no formal annual regulatory standard for long-term O₃, we selected 88.2 µg/m³ [45 ppb] as the threshold value to define low-level warm-season O₃, which has been chosen as a plausible pollution target in previous studies to evaluate its effectiveness in reducing health risk^{70,71}. We did not examine lower thresholds due to potentially insufficient statistical power from fewer observations.

Covariates

We considered a variety of SES covariates at the ZIP code tabulation-area (ZCTA) level as important predictors for cardiovascular disease⁷², including percent of the population self-reporting as Black, percent of the population self-reporting as Hispanic, percent of the population ≥65 years of age living in poverty, population density, percent of the population ≥65 years of age who had not graduated from high school, median home value, median household income, and percent of owner-occupied housing unit. These data were obtained from the US Census Bureau 2000 and 2010 Census Summary File 3 and the American Community Survey from 2011 through 2016. To account for long-term smoking behaviors, we included lung cancer hospitalization rates as a surrogate measure for each ZIP code from the MEDPAR file. We also accessed county-level data on the yearly percentage of residents who ever smoked and mean body mass index (BMI) from the Centers for Disease Control and Prevention (CDC) Behavioral Risk Factor Surveillance System (BRFSS)⁷³. These county-level lifestyle data were assigned to ZIP codes. Additionally, from the Dartmouth Atlas of Health Care⁷⁴,

we obtained several access-to-care covariates in each hospital service area, and further assigned them to ZIP codes: proportion of Medicare beneficiaries with at least 1 hemoglobinA1c test per year, proportion of diabetic beneficiaries who had a lipid panel test in a year, proportion of beneficiaries who had an eye examination in a year, proportion of beneficiaries with at least 1 ambulatory doctor visits in a year, and proportion of female beneficiaries who had a mammogram during a 2-year period. We also calculated the distance from the centroid of each ZIP code to the nearest hospital, a proxy for healthcare accessibility, using data on hospital locations derived from an ESRI dataset⁷⁵. Given that seasonal meteorological conditions have been known to impact cardiovascular health^{76,77}, we assessed the average temperature and relative humidity (RH) during the summer (June–August) and the winter (December–February) for each ZIP code and each year based on the 4 km Gridded Surface Meteorological (gridMET) dataset^{78,79}.

Missing values for all area-level risk factors were filled in using linear interpolation and extrapolation. Any other missingness accounting for <1% of the observations was assumed to be random and was excluded from our analyses.

Statistical analysis

In this study, we analyzed the association between long-term exposure to low-level air pollution and hospitalization rate of major cardiovascular diseases among the US Medicare population. As aforementioned, the analysis was restricted to the low pollution ZIP code areas with more than 100 Medicare beneficiaries. We used a double negative control strategy¹⁶, which has been recommended to address unmeasured confounding in observational settings, to enhance the causal evidence of a potential relationship⁷⁹. The detailed descriptions of this double negative control approach can be found elsewhere²⁶. A summary of the principles is given below. First, we consider a quasi-Poisson regression model to obtain the unbiased association between the exposure (A) and the outcome (Y), adjusting for unmeasured confounders (U):

$$\ln[E(Y)] = \beta_{Y0} + \beta_{YA}A + \beta_{YU}U \quad (1)$$

The negative exposure control (Z) and negative outcome control (W) are designed to capture confounding bias introduced by U . In this study, we chose the exposure to air pollution in the year after cause-specific hospitalizations as Z . It cannot lead to the hospitalization outcome in the concurrent year, however, it could be influenced by unmeasured or measured confounders that are correlated with air pollution level in the year of the hospitalization outcome. Similarly, we defined the count of cause-specific hospitalizations in the year before exposure as W , as it is by no means affected by the exposure in the concurrent year but may be correlated to omitted confounders. Given the hypothesized correlations of U with A and Z , and non-causality between A and W , the formulas (2) and (3) can be derived:

$$E(U) = \beta_{U0} + \beta_{UA}A + \beta_{UZ}Z \quad (2)$$

$$\ln[E(W)] = \beta_{W0} + \beta_{WU}U \quad (3)$$

If we substitute U with its expected value regressed on A and Z from the formula (2), the formula (1) can be interpreted into:

$$\ln[E(Y)] = (\beta_{Y0} + \beta_{YU}\beta_{U0}) + (\beta_{YA} + \beta_{YU}\beta_{UA})A + \beta_{YU}\beta_{UZ}Z \quad (4)$$

where $\beta_{YU}\beta_{UA}$ is exactly equal to the bias due to unmeasured confounders. Thus, if the equation $\beta_{UZ} = \beta_{UA}$ holds, the subtraction between the coefficient of A and the coefficient of Z will yield a causal effect of A on Y .

If we substitute U with its expected value again in the formula (3), W as a surrogate for U can be predicted by A and Z based on:

$$\ln[E(W)] = (\beta_{W0} + \beta_{WU}\beta_{U0}) + \beta_{WU}\beta_{UA}A + \beta_{WU}\beta_{UZ}Z \quad (5)$$

Alternatively, assuming the linear correlations of U with A and Z , which renders the formulas (2) and (5) valid, we can mitigate the confounding effect of U by including the predicted W in the outcome regression model.

To summarize, unmeasured confounding bias can be captured if either of the following two assumptions holds true:

- (1) We assume that U is linearly correlated with both A and Z . Although W is unlikely to link to exposure variables by its definition, a correlation of W with A and Z can be introduced due to a connection with U . Therefore, the predicted value of W by regressing it on A and Z represents the part of U that is related to A and Z . As a surrogate for U , adjusting for predicted W is equivalent to removing omitted confounding bias.
- (2) Alternatively, we assume that U has the same magnitude of correlations with A and Z ($\beta_{UZ} = \beta_{UA}$). According to the formula (4), if $\beta_{UZ} = \beta_{UA}$, the causal effect of A on Y can be derived by subtracting the coefficient for Z from the coefficient for A . If either assumption holds, omitted confounder U is controlled for.

Conversely, violations may occur if neither of these two assumptions is satisfied.

In both the model used to predict the negative outcome control and the outcome regression model, we adjusted for a variety of area-level risk factors for cardiovascular diseases selected prior, including SES, behavioral, and meteorological covariates which are described in the covariates section, to relax our assumptions and to mitigate potential uneliminated confounding bias as comprehensively as possible. Confounding bias by other unmeasured area-level and individual-level factors was assumed to be addressed given the afore-described assumptions. We also included the admission year as a categorical indicator in the models to control for the time trends of omitted confounders that might drive an association. We analyzed the effect of each air pollutant separately using both a single-pollutant model and a three-pollutant model. A directed acyclic graph for the double negative control approach considering measured confounders altogether is shown in Supplementary Fig. 6. The performance of the outcome regression models is evaluated as satisfactory through computing the Quasi-Akaike information criteria and pseudo- R^2 values (Supplementary Table 8).

As a secondary analysis, we repeated the main analyses using generalized linear models (GLM) without the negative controls. To examine the shape of E-R curves, we further applied natural spline functions with three degrees of freedom to the GLM adjusted for co-pollutants.

We examined the potential effect measure modification by individual demographic characteristics, namely, age (65–74 years, 75–84 years, 85+ years), sex (male or female), race (White or Black), and Medicaid eligibility (yes or no), using stratified analyses. We conducted comparisons of coefficients within the strata of each factor to detect any statistically significant differences, assuming the difference between the coefficients to follow a normal distribution with a mean of zero and a variance of the sum of the strata variances. To assess the robustness of the results, we repeated the primary analysis by removing confounding adjustments from the prediction model for the negative outcome control. In the above analyses, we reported the effect as the percent change in hospitalization rate and its 95% CIs for each cardiovascular outcome per $\mu\text{g}/\text{m}^3$ increase in annual exposure to $\text{PM}_{2.5}$ and per ppb increase in annual exposure to NO_2 and O_3 .

All analyses were performed using R software version 4.2.3 on the Research Computing Environment as part of Research Computer at

Harvard University Faculty of Arts and Sciences. A two-sided P -value < 0.05 was considered statistically significant.

Reporting summary

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

Data availability

The Medicare data are available under restricted access for the requirements from the Center for Medicare and Medicaid Services (CMS). Therefore, the authors do not have permission to share Medicare data. Interested investigators can obtain it by applying for their own Data Use Agreement to the CMS and the CMS dissemination contractor will process the data request. The air pollution data are available at the NASA SEDAC website (<https://sedac.ciesin.columbia.edu/data/collection/aqdh/sets/browse>). The SES data from US census and American Community Survey are available at <https://www.census.gov/data/datasets/2000/dec/summary-file-3.html>, <https://www.census.gov/data/datasets/2010/dec/summary-file-1.html>, and <https://www.census.gov/data/developers/data-sets/acs-1year.html>. The BRFSS data are available at https://www.cdc.gov/brfss/annual_data/annual_data.htm. The data from the Dartmouth Atlas of Health Care are available at <https://data.dartmouthatlas.org/>. The air pollution data are freely available online at the NASA SEDAC website (<https://sedac.ciesin.columbia.edu/data/collection/aqdh/sets/browse>). Source data of the figures are available with this paper. Source data are provided with this paper.

Code availability

The R codes of this study are publicly available from https://github.com/Yichen0430/air_pollution_negative_control.

References

1. Brook, R. D. et al. Particulate matter air pollution and cardiovascular disease: an update to the scientific statement from the American Heart Association. *Circulation* **121**, 2331–2378 (2010).
2. Cosselman, K. E., Navas-Acien, A. & Kaufman, J. D. Environmental factors in cardiovascular disease. *Nat. Rev. Cardiol.* **12**, 627–642 (2015).
3. Hankey, G. J. The global and regional burden of stroke. *Lancet Glob. Health* **1**, e239–e240 (2013).
4. Kulick, E. R., Kaufman, J. D. & Sack, C. Ambient air pollution and stroke: an updated review. *Stroke* **54**, 882–893 (2023).
5. Abraham, J. M. & Connolly, S. J. Atrial fibrillation in heart failure: Stroke risk stratification and anticoagulation. *Heart Fail Rev.* **19**, 305–313 (2014).
6. Atkinson, R. W. et al. Long-term exposure to outdoor air pollution and incidence of cardiovascular diseases. *Epidemiology* **24**, 44–53 (2013).
7. Bai, L. et al. Exposure to ambient air pollution and the incidence of congestive heart failure and acute myocardial infarction: A population-based study of 5.1 million Canadian adults living in Ontario. *Environ. Int.* **132**, 105004 (2019).
8. Wang, M. et al. Joint exposure to various ambient air pollutants and incident heart failure: a prospective analysis in UK Biobank. *Eur. Heart J.* **42**, 1582–1591 (2021).
9. Danesh Yazdi, M. et al. Long-term association of air pollution and hospital admissions among medicare participants using a doubly robust additive model. *Circulation* **143**, 1584–1596 (2021).
10. Yue, C., Yang, F., Li, F. & Chen, Y. Association between air pollutants and atrial fibrillation in general population: A systematic review and meta-analysis. *Ecotoxicol. Environ. Saf.* **208**, 111508 (2021).
11. Burnett, R. et al. Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter. *Proc. Natl Acad. Sci.* **115**, 9592–9597 (2018).

12. Vodonos, A., Awad, Y. A. & Schwartz, J. The concentration-response between long-term PM_{2.5} exposure and mortality: A meta-regression approach. *Environ. Res.* **166**, 677–689 (2018).
13. Jin, T. et al. Associations between long-term air pollution exposure and the incidence of cardiovascular diseases among American older adults. *Environ. Int.* **170**, 107594 (2022).
14. Brunekreef, B. et al. Mortality and morbidity effects of long-term exposure to low-level PM_{2.5}, BC, NO₂, and O₃: an analysis of European Cohorts in the ELAPSE project. *Res Rep. Health Eff. Inst.* **2021**, 208 (2021).
15. Stafoggia, M. et al. Long-term exposure to low ambient air pollution concentrations and mortality among 28 million people: results from seven large European cohorts within the ELAPSE project. *Lancet Planet. Health* **6**, e9–e18 (2022).
16. Lipsitch, M., Tchetgen, E. T. & Cohen, T. Negative controls: a tool for detecting confounding and bias in observational studies. *Epidemiology* **21**, 383–388 (2010).
17. Shi, X., Miao, W., Nelson, J. C. & Tchetgen Tchetgen, E. J. Multiply robust causal inference with double-negative control adjustment for categorical unmeasured confounding. *J. R. Stat. Soc. Ser. B: Stat. Methodol.* **82**, 521–540 (2020).
18. Flanders, W. D., Strickland, M. J. & Klein, M. A new method for partial correction of residual confounding in time-series and other observational studies. *Am. J. Epidemiol.* **185**, 941–949 (2017).
19. Flanders, W. D. et al. A method for detection of residual confounding in time-series and other observational studies. *Epidemiol. (Camb., Mass.)* **22**, 59 (2011).
20. Schwartz, J., Bind, M.-A. & Koutrakis, P. Estimating causal effects of local air pollution on daily deaths: effect of low levels. *Environ. Health Perspect.* **125**, 23–29 (2017).
21. Yu, Y. et al. Identification and estimation of causal effects using a negative-control exposure in time-series studies with applications to environmental epidemiology. *Am. J. Epidemiol.* **190**, 468–476 (2021).
22. Sheppard, L., Levy, D., Norris, G., Larson, T. V. & Koenig, J. Q. Effects of ambient air pollution on nonelderly asthma hospital admissions in Seattle, Washington, 1987–1994. *Epidemiology* **10**, 23 (1999).
23. Yitshak-Sade, M., Kloog, I., Zanobetti, A. & Schwartz, J. D. Estimating the causal effect of annual PM_{2.5} exposure on mortality rates in the Northeastern and mid-Atlantic states. *Environ. Epidemiol.* **3**, e052 (2019).
24. Liu, R. A., Wei, Y., Qiu, X., Kosheleva, A. & Schwartz, J. D. Short term exposure to air pollution and mortality in the US: a double negative control analysis. *Environ. Health* **21**, 81 (2022).
25. Miao W., Shi X. & Tchetgen E. T. A confounding bridge approach for double negative control inference on causal effects. [arXiv.org https://arxiv.org/abs/1808.04945v3](https://arxiv.org/abs/1808.04945v3) (2018).
26. Schwartz, J., Wei, Y., Dominici, F. & Yazdi, M. D. Effects of low-level air pollution exposures on hospital admission for myocardial infarction using multiple causal models. *Environ. Res.* **232**, 116203 (2023).
27. Krewski, D. et al. Extended follow-up and spatial analysis of the American Cancer Society Study Linking Particulate Air Pollution and Mortality. *Res Rep. Health Eff. Inst.* **140**, 5–114 (2009).
28. Di, Q. et al. Air pollution and mortality in the medicare population. *N. Engl. J. Med.* **376**, 2513–2522 (2017).
29. Shi, L. et al. Low-concentration PM_{2.5} and mortality: estimating acute and chronic effects in a population-based study. *Environ. Health Perspect.* **124**, 46–52 (2016).
30. Wang, Y., Qiu, X., Wei, Y. & Schwartz, J. D. Long-term exposure to ambient PM_{2.5} and hospitalizations for myocardial infarction among US Residents: A Difference-in-Differences analysis. *JAHA* **12**, e029428 (2023).
31. Ma, T. et al. Long-term air pollution exposure and incident stroke in American older adults: a national cohort study. *Glob. Epidemiol.* **4**, 100073 (2022).
32. Scheers, H., Jacobs, L., Casas, L., Nemery, B. & Nawrot, T. S. Long-term exposure to particulate matter air pollution is a risk factor for stroke: Meta-analytical evidence. *Stroke* **46**, 3058–3066 (2015).
33. Alexeeff, S. E., Liao, N. S., Liu, X., Van Den Eeden, S. K. & Sidney, S. Long-term PM_{2.5} exposure and risks of ischemic heart disease and stroke events: Review and meta-analysis. *JAHA* **10**, e016890 (2021).
34. To, T. et al. Chronic disease prevalence in women and air pollution — A 30-year longitudinal cohort study. *Environ. Int.* **80**, 26–32 (2015).
35. 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 (2017).
36. Zhang, D. et al. Air pollution exposure and heart failure: a systematic review and meta-analysis. *Sci. Total Environ.* **872**, 162191 (2023).
37. 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**, 087002 (2021).
38. Dirgawati, M. et al. Long-term exposure to low air pollutant concentrations and the relationship with all-cause mortality and stroke in older men. *Epidemiology* **30**, S82–S89 (2019).
39. Hart, J. E., Puett, R. C., Rexrode, K. M., Albert, C. M. & Laden, F. Effect modification of long-term air pollution exposures and the risk of incident cardiovascular disease in US women. *JAHA* **4**, e002301 (2015).
40. Kwon, O. K. et al. Association of short- and long-term exposure to air pollution with atrial fibrillation. *Eur. J. Prev. Cardiol.* **26**, 1208–1216 (2019).
41. Dominici, F., Peng, R. D., Barr, C. D. & Bell, M. L. Protecting human health from air pollution: Shifting from a single-pollutant to a multi-pollutant approach. *Epidemiology* **21**, 187–194 (2010).
42. Franchini, M. & Mannucci, P. M. Thrombogenicity and cardiovascular effects of ambient air pollution. *Blood* **118**, 2405–2412 (2011).
43. Bourdrel, T., Bind, M.-A., Béjot, Y., Morel, O. & Argacha, J.-F. Cardiovascular effects of air pollution. *Arch. Cardiovasc Dis.* **110**, 634–642 (2017).
44. Du, Y., Xu, X., Chu, M., Guo, Y. & Wang, J. Air particulate matter and cardiovascular disease: the epidemiological, biomedical and clinical evidence. *J. Thorac. Dis.* **8**, E8–E19 (2016).
45. Rowland, D. & Lyons, B. Medicare, medicaid, and the elderly poor. *Health Care Financ Rev.* **18**, 61–85 (1996).
46. Chi, G. et al. Individual and neighborhood socioeconomic status and the association between air pollution and cardiovascular disease. *Environ. Health Perspect.* **124**, 1840–1847 (2016).
47. Medina-Ramón, M. & Schwartz, J. Who is more vulnerable to die from ozone air pollution? *Epidemiology* **19**, 672–679 (2008).
48. Liu, G. et al. The gender-based differences in vulnerability to ambient air pollution and cerebrovascular disease mortality: evidences based on 26781 deaths. *Glob. Heart* **15**, 46 (2020).
49. Shin, H. H. et al. Sex-difference in air pollution-related acute circulatory and respiratory mortality and hospitalization. *Sci. Total Environ.* **806**, 150515 (2022).
50. Rodgers, J. L. et al. Cardiovascular risks associated with gender and aging. *J. Cardiovascular Dev. Dis.* **6**, 19 (2019).
51. Chen, M., Zhao, J., Zhuo, C. & Zheng, L. The association between ambient air pollution and atrial fibrillation: a systematic review and meta-analysis. *Int. Heart J.* **62**, 290–297 (2021).
52. Gao, P. et al. Acute effects of ambient nitrogen oxides and interactions with temperature on cardiovascular mortality in Shenzhen, China. *Chemosphere* **287**, 132255 (2022).
53. Hoek, G. et al. Long-term air pollution exposure and cardio-respiratory mortality: a review. *Environ. Health* **12**, 43 (2013).

54. Vallée, A. Sex associations between air pollution and estimated atherosclerotic cardiovascular disease risk determination. *Int J. Public Health* **68**, 1606328 (2023).
55. Hicken, M. T. et al. Air pollution, cardiovascular outcomes, and social disadvantage: the multi-ethnic study of atherosclerosis. *Epidemiology* **27**, 42–50 (2016).
56. Johnson, D. & Parker, J. D. Air pollution exposure and self-reported cardiovascular disease. *Environ. Res.* **109**, 582–589 (2009).
57. Gouveia, N. & Fletcher, T. Time series analysis of air pollution and mortality: effects by cause, age and socioeconomic status. *J. Epidemiol. Community Health* **54**, 750–755 (2000).
58. Zanobetti, A. & Schwartz, J. Race, gender, and social status as modifiers of the effects of PM10 on mortality. *J. Occup. Environ. Med.* **42**, 469 (2000).
59. Dragano, N. et al. Traffic exposure and subclinical cardiovascular disease: is the association modified by socioeconomic characteristics of individuals and neighbourhoods? Results from a multilevel study in an urban region. *Occup. Environ. Med.* **66**, 628–635 (2009).
60. Wang, Y. et al. Spatial decomposition analysis of NO₂ and PM_{2.5} air pollution in the United States. *Atmos. Environ.* **241**, 117470 (2020).
61. Ji, J. S. et al. NO₂ and PM_{2.5} air pollution co-exposure and temperature effect modification on pre-mature mortality in advanced age: a longitudinal cohort study in China. *Environ. Health* **21**, 97 (2022).
62. Lamsal, L. N. et al. U.S. NO₂ trends (2005–2013): EPA Air Quality System (AQS) data versus improved observations from the Ozone Monitoring Instrument (OMI). *Atmos. Environ.* **110**, 130–143 (2015).
63. Danesh Yazdi, M. et al. The effect of long-term exposure to air pollution and seasonal temperature on hospital admissions with cardiovascular and respiratory disease in the United States: A difference-in-differences analysis. *Sci. Total Environ.* **843**, 156855 (2022).
64. Klepeis, N. E. et al. The National Human Activity Pattern Survey (NHAPS): a resource for assessing exposure to environmental pollutants. *J. Expo. Sci. Environ. Epidemiol.* **11**, 231–252 (2001).
65. Wei, Y. et al. The impact of exposure measurement error on the estimated concentration–response relationship between long-term exposure to PM_{2.5} and mortality. *Environ. Health Perspect.* **130**, 077006 (2022).
66. Weisskopf, M. G. & Webster, T. F. Trade-offs of personal vs. more proxy exposure measures in environmental epidemiology. *Epidemiology* **28**, 635–643 (2017).
67. Bell, M. L., Peng, R. D. & Dominici, F. The exposure–response curve for ozone and risk of mortality and the adequacy of current ozone regulations. *Environ. Health Perspect.* **114**, 532–536 (2006).
68. Di, Q. et al. Assessing NO₂ concentration and model uncertainty with high spatiotemporal resolution across the contiguous United States using ensemble model averaging. *Environ. Sci. Technol.* **54**, 1372–1384 (2020).
69. Di, Q. et al. An ensemble-based model of PM_{2.5} concentration across the contiguous United States with high spatiotemporal resolution. *Environ. Int.* **130**, 104909 (2019).
70. Requia, W. J. et al. An ensemble learning approach for estimating high spatiotemporal resolution of ground-level ozone in the contiguous United States. *Environ. Sci. Technol.* **54**, 11037–11047 (2020).
71. Rhee, J. et al. Impact of long-term exposures to ambient PM_{2.5} and ozone on ards risk for older adults in the United States. *Chest* **156**, 71–79 (2019).
72. Davari, M., Maracy, M. R. & Khorasani, E. Socioeconomic status, cardiac risk factors, and cardiovascular disease: A novel approach to determination of this association. *ARYA Atheroscler.* **15**, 260–266 (2019).
73. Bureau, U. S. C. Behavioral risk factor surveillance system (2005–2019). (2020).
74. Wennberg, J. E. & Cooper, M. M. *The Dartmouth Atlas of Health Care in the United States: The Center for the Evaluative Clinical Sciences*. (American Hospital Publishing, Inc., Chicago (IL), 1996).
75. ESRI Data and Maps & United States Geological Survey. USA Hospitals. (2010).
76. Shi, L., Kloog, I., Zanobetti, A., Liu, P. & Schwartz, J. D. Impacts of temperature and its variability on mortality in New England. *Nat. Clim. Chang* **5**, 988–991 (2015).
77. Stewart, S., Keates, A. K., Redfern, A. & McMurray, J. J. V. Seasonal variations in cardiovascular disease. *Nat. Rev. Cardiol.* **14**, 654–664 (2017).
78. Abatzoglou, J. T. Development of gridded surface meteorological data for ecological applications and modelling. *Int. J. Climatol.* **33**, 121–131 (2013).
79. Sofer, T., Richardson, D. B., Colicino, E., Schwartz, J. & Tchetgen Tchetgen, E. J. On negative outcome control of unobserved confounding as a generalization of difference-in-differences. *Stat. Sci.* **31**, 348–361 (2016).

Acknowledgements

This work was made possible by J.D.S.'s US Environmental Protection Agency grant RD-835872. Its contents are solely the responsibility of the grantee and do not necessarily represent the official views of the US Environmental Protection Agency. Furthermore, the US Environmental Protection Agency does not endorse the purchase of any commercial products or services mentioned in the publication. This work was also supported by J.D.S.'s National Institutes of Health grant R01 ES032418-01 and National Institute of Environmental Health Sciences grant P30 ES000002.

Author contributions

Y.W. (Yichen Wang) contributed to methodology, data analysis, visualization, original draft writing, and review & editing of the manuscript. M.D.Y. and Y.W. (Yaguang Wei) contributed to data curation, and review & editing of the manuscript. J.D.S. contributed to funding acquisition, conceptualization, methodology, supervision, and review & editing of the manuscript.

Competing interests

J.D.S. reports having been an expert witness for the US Department of Justice on cases involving violations of the Clean Air Act. The remaining authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-52117-8>.

Correspondence and requests for materials should be addressed to Yichen Wang.

Peer review information *Nature Communications* thanks the anonymous, reviewers for their contribution to the peer review of this work. A peer review file is available.

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

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

Open Access This article is licensed under a Creative Commons Attribution-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) 2024