bmjmedicine

Associations between long term air pollution exposure and first hospital admission for kidney and total urinary system diseases in the US Medicare population: nationwide longitudinal cohort study

Whanhee Lee ,[1,2,3] Xiao Wu,[4] Seulkee Heo,[1] Kelvin C Fong,[1] Ji-Young Son,[1] M Benjamin Sabath,[5] Danielle Braun,[6,7] Jae Yoon Park,[8,9] Yong Chul Kim,[10] Jung Pyo Lee ,[11] Joel Schwartz,[12] Ho Kim,[13,14] Francesca Dominici,[4] Michelle Bell[1]

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/bmjmed-2021-000009).

For numbered affiliations see end of article.
Correspondence to: Dr Whanhee Lee, School of the Environment, Yale University, New Haven, CT 06511, USA; whanhee.lee@yale.edu

Cite this as: *BMJMED* 2022;1:e000009. doi:10.1136/bmjmed-2021-000009

Received: 1 September 2021
Accepted: 21 March 2022

ABSTRACT

OBJECTIVE To estimate the associations between long term exposure to air pollution and the first hospital admission related to kidney and total urinary system diseases.

DESIGN Nationwide longitudinal cohort study.

SETTING Data were collected from the Medicare fee-for-service for beneficiaries living in 34 849 zip codes across the continental United States from 2000 to 2016. Exposure variables were annual averages of traffic related pollutants (fine particles ($PM_{2.5}$) and nitrogen dioxide (NO_2)) that were assigned according to the zip code of residence of each beneficiary with the use of validated and published hybrid ensemble prediction models.

PARTICIPANTS All beneficiaries aged 65 years or older who were enrolled in Medicare part A fee-for-service (n=61 097 767).

PRIMARY AND SECONDARY OUTCOME MEASURES First hospital admission with diagnosis codes for total kidney and urinary system disease or chronic kidney disease (CKD), analyzed separately.

RESULTS The average annual concentrations of air pollution were 9.8 µg/m^3 for $PM_{2.5}$ and 18.9 ppb for NO_2. The total number of first admissions related to total kidney and urinary system disease and CKD were around 19.0 million and 5.9 million, respectively (2000-16). For total kidney and urinary system disease, hazard ratios were 1.076 (95% confidence interval 1.071 to 1.081) for a 5 µg/m^3 increase in $PM_{2.5}$ and 1.040 (1.036 to 1.043) for a 10 ppb increase in NO_2. For CKD, hazard ratios were 1.106 (1.097 to 1.115) for a 5 µg/m^3 increase in $PM_{2.5}$ and 1.013 (1.008 to 1.019) for a 10 ppb increase in NO_2. These positive associations between $PM_{2.5}$ and kidney outcomes persisted at concentrations below national health based air quality standards.

CONCLUSIONS The findings suggest that higher annual air pollution levels were associated with increased risk of first hospital admission related to diseases of the kidney and urinary system or CKD in the Medicare population.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ A few studies have reported that long term exposure to air pollution is associated with an increased risk of incident chronic kidney disease (CKD)
⇒ The generalizability of these results is limited because of the non-representative nature of the study populations

WHAT THIS STUDY ADDS

⇒ The findings of this large nationwide study suggest that long term exposures to fine particles ($PM_{2.5}$) and nitrogen dioxide (NO_2) were associated with an increased risk of first hospital admission related to diseases of the kidney and urinary system or CKD
⇒ These associations were found to persist at levels below the current annual National Ambient Air Quality Standards for $PM_{2.5}$ and NO_2

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE, OR POLICY

⇒ The findings of this study suggest an adverse effect of air pollution on diseases of the kidney and urinary system, which could help establish public health interventions

Introduction

Kidney disease has been recognized as a worldwide public health concern owing to the increased risk of early death, cardiovascular disease, and poor quality of life.[1–3] Globally, over 750 million people have kidney diseases,[4 5] annual mortality attributable to kidney disease might be as high as 5 million,[6] and its prevalence has increased during recent decades.[7 8] Among the types of kidney disease, chronic kidney disease (CKD) is an important public health priority worldwide.[4 7 9] In 2016, CKD was the 13th highest cause of death on a global scale and projections suggest that it will become the world's fifth leading cause of death by 2040.[10] In the United States, about 15% of adults (37 million) were estimated to have CKD in 2021 and the disease is more common in people aged 65 years or older (38.1%).[11] Kidney disease results in high medical costs and financial burden for Medicare. According to the 2020 US Renal Data System annual data report, the total Medicare fee-for-service (FFS) spending for beneficiaries with CKD who did not have end stage renal disease exceeded $81bn in 2018 (£61.7bn, €72.7bn; 22.3% of Medicare FFS spending), and annual expenditure for beneficiaries aged 66 years or older was estimated at $23 691 per person.[12] The total Medicare FFS expenditures for beneficiaries with end stage renal disease was $36.6bn in 2018 (7.2% of Medicare spending).

Recently, air pollution has been suggested as a potential risk factor for low glomerular filtration

rate and kidney disease.[13–16] Inhaled air pollution can cause oxidative stress and DNA damage to kidney tissue,[17] and cardiovascular and endocrine toxicities related to air pollution, such as disturbances in the autonomic nervous and circulatory systems and pulmonary inflammatory disease, have also been suggested as biological mechanisms linking air pollution exposure and kidney disease.[14 18 19] In spite of these plausible biological mechanisms, associations between air pollution and kidney disease have been studied less than those for cardiovascular and respiratory diseases.[13] Several recent cohort studies investigated the association between air pollution exposure and the incidence of kidney disease.[13 14 20–22] Although these studies consistently reported the positive association between exposure to air pollution and the risk of kidney disease, these studies had limitations related to the type of participants (eg, US veterans who were mostly white men),[13 14] locations,[22] and sample size (less than 200 000)[20 21]; therefore, the results were limited in providing generalizable estimates for the association between air pollution and kidney disease. Additionally, air pollution concentrations in the US have been reduced in the past few decades; therefore, it is crucial to evaluate whether the relation between air pollution and kidney disease persists at lower concentrations.[23]

To resolve these gaps in knowledge, we performed a nationwide longitudinal cohort study of Medicare Part A FFS beneficiaries from 2000 to 2016, a population of over 61 million, with 410 million person years of follow-up (that is, an average 6.7 follow-up years per person). We investigated the association between long term exposure to air pollution and the first hospital admission related to total kidney and urinary system disease or CKD. A nationwide air pollution dataset integrating validated zip code level fine particulate matter ($PM_{2.5}$) and nitrogen dioxide (NO_2) prediction models was used. These pollutants are key components of pollution from traffic, although they have other sources such as energy production and industry. Based on the large sample size and geographical scope, the study aimed to provide generalizable and statistically reliable results on the association between traffic related air pollution and the first hospital admission related to total kidney and urinary system disease or CKD with state-of-the-science air pollution prediction models.

Methods

The findings are reported in accordance with the strengthening the reporting of observational studies in epidemiology (STROBE) guidelines (see online supplemental material A).

Study design and participants

Our longitudinal study cohort consisted of beneficiaries who were enrolled in Medicare part A FFS (age 65 years or older) in the US from 1 January 2000 to 31 December 2016. We used the Medicare inpatient hospital claims from the Medicare Provider and Analysis Review files that contain one record per hospital admission. People become eligible to enter Medicare at age 65 years. Follow-up for each beneficiary started on 1 January 2000 or 1 January of the year after entry into the cohort and continued until the first admission with diagnosis codes related to each kidney outcome (total kidney and urinary system disease or CKD, analyzed separately), death, or end of the study period (31 December 2016), whichever came first. For each beneficiary, we extracted the date of the first admission, age, sex, self-reported race, zip code of residence, and Medicaid eligibility (as a proxy for low socioeconomic status[24] in each follow-up year). Additionally, if the residential address changed during the follow-up period, exposure assessment incorporated any address changes in the analysis. For each zip code, the mean population was around 7500 and the median land area is about 92 km^2.

We used international classification of diseases (ICD) codes to identify diseases of the kidney and the urinary system (ICD-9: 580-599; ICD-10: N00-N39) and CKD (ICD-9: 585; ICD-10: N18) as the primary or secondary diagnosis during the study period. Analysis is based on each participant's first hospital admission related to a given cause. The kidney and bladder are linked by ureters and are closely related. For example, neurogenic bladder and bacterial infections in the bladder and urinary tract are major causes of kidney diseases.[25 26] Additionally, kidney stones are an important risk factor for urinary diseases, such as hematuria and urinary tract infections.[27] This study did not subdivide kidney and urinary disease, and previous studies also applied the combined ICD codes for total kidney and urinary system disease (N00-N39) as total renal disease.[28–34] Because kidney disease has causal associations with hypertension, diabetes, and cardiac diseases, and is also closely associated with various complications including heart disease, stroke, and other metabolic diseases,[35] we considered the primary and secondary diagnosis ICD codes in the main analysis. Therefore, the outcomes of this study should be interpreted as the first hospital admission related to total kidney and urinary system disease or CKD.

Air pollution data

We obtained daily concentration predictions of ambient levels of $PM_{2.5}$ and NO_2 at 1×1 km^2 spatial resolution across the contiguous US from validated, published models.[36 37] We used these data to calculate the daily mean concentration of each air pollutant for each zip code by averaging the predictions at the grid cells with centroid points inside the boundary of that zip code.[23 24 38] For each calendar

Table 1 | Descriptive cohort characteristics for US Medicare beneficiaries from 2000 to 2016. Values are number (%) or mean (SD)

	Participants in full cohort (n=61 097 767)	Participants in low exposure cohort† (n=7 022 642)	Participants not in low exposure cohort‡ (n=54 075 125)
Age at entry, years			
65-74	46 886 857 (76.7)	5 399 807 (76.9)	41 487 050 (76.7)
75-84	10 434 150 (17.1)	1 188 202 (16.9)	9 245 948 (17.1)
≥85	3 776 760 (6.2)	434 633 (6.2)	3 342 127 (6.2)
Sex			
Men	27 413 250 (44.9)	3 258 875 (46.2)	24 154 385 (44.7)
Women	33 684 507 (55.1)	3 763 767 (53.8)	29 920 740 (55.3)
Race			
White	51 577 679 (84.4)	6 610 683 (94.1)	44 966 996 (83.2)
Black	5 372 559 (8.8)	122 114 (1.7)	5 250 445 (9.7)
Other*	3 480 763 (5.7)	223 421 (3.2)	3 257 342 (6.0)
Medicaid eligibility			
Not eligible	53 561 222 (87.7)	6 355 624 (90.5)	47 205 598 (87.3)
Eligible	7 536 545 (12.3)	667 018 (9.5)	6 869 527 (12.7)
Air pollution concentration			
$PM_{2.5}$ (μg/m^3)	9.8 (3.1)	7.2 (2.2)	10.2 (3.0)
NO_2 (ppb)	19.0 (10.2)	9.9 (3.6)	20.3 (10.2)
Potential confounder (community level characteristics)			
Below poverty level (%)	10.5 (8.1)	9.3 (7.0)	10.7 (8.2)
Population density (people per km^2)	3300.5 (8590.2)	266.3 (702.8)	3744.6 (9109.1)
Median home value ($1000)	203.8 (167.6)	161.8 (127.2)	209.3 (171.9)
Black (%)	11.9 (18.4)	2.8 (7.1)	13.2 (19.1)
Hispanic (%)	7.2 (8.7)	4.6 (5.9)	7.5 (9.0)
Median annual household income ($1000)	52.6 (23.1)	47.7 (16.0)	53.3 (23.9)
Owner occupied housing (%)	67.4 (17.8)	75.2 (11.6)	66.3 (18.2)
Below high school education (%)	27.0 (16.1)	23.5 (14.5)	27.6 (16.3)
Ever smoked (%)	46.0 (7.2)	48.8 (8.0)	45.6 (6.9)
Body mass index	27.5 (1.0)	27.6 (1.0)	27.5 (1.0)

$1.00=£0.76, €0.90. SD=standard deviation.

*Other included Asian, Hispanic, American Indian or Alaskan Native, and unknown. Data on race were obtained from Medicare beneficiary files.

†Fine particles ($PM_{2.5}$)<12 μg/m^3 and nitrogen dioxide (NO_2)<20 ppb.

‡$PM_{2.5}$≥12 μg/m^3 and NO_2≥20 ppb.

year, we assigned the annual average concentrations for each zip code to Medicare enrollees according to the zip code of residence and calendar year to calculate long term exposures to air pollutants.[23 24 38] The coefficients of determination (cross validated R^2) for annual $PM_{2.5}$ and NO_2 were 0.89 and 0.84, respectively, across the study area.[36–38] Detailed information on the air pollution data is given in online supplemental material B.

Statistical analysis

We applied a two pollutant Cox equivalent reparameterized Poisson approach to manage the computational challenges of the conventional Cox proportional hazard model (eg, insufficient memory size and lengthy computation time).[23] This Poisson model was mathematically identical to a time varying Cox proportional hazard model under an Anderson-Gill representation.[23 39] Specifically, for each of two kidney outcomes (first hospital admission related to total kidney and urinary system disease or CKD), a stratified Poisson model with a log link function was fit to estimate the association between time varying annual air pollutant and the first hospital admissions related to total kidney and urinary system disease or CKD. The dependent variable was the number of first hospital admissions in each follow-up year, calendar year, and zip code within stratums specified by individual variables: age at study entry in two year categories (65-66, 67-68, ..., ≥85), sex, race (white, black, and other—Asian, Hispanic, American Indian, Alaskan Native, and unknown), and Medicaid eligibility. Data

Table 2 | Admissions related to total kidney and urinary system disease and chronic kidney disease (CKD), and association between air pollution and first hospital admission related to kidney diseases

	$PM_{2.5}$	NO_2
Total kidney and urinary system disease		
Total cohort		
No of admissions	19 195 342	
Total person years	410 303 502	
Median follow-up year	5	
Hazard ratio (95% CI)	1.076 (1.071 to 1.081)	1.040 (1.036 to 1.043)
Low pollution cohort		
No of admissions	5 233 054	
Total person years	113 564 232	
Median follow-up year	6	
Hazard ratio (95% CI)	1.133 (1.121 to 1.145)	1.021 (1.009 to 1.033)
Chronic kidney disease		
Total cohort		
No of admissions	5 902 129	
Total person years	454 497 761	
Median follow-up year	6	
Hazard ratio (95% CI)	1.106 (1.097 to 1.115)	1.013 (1.008 to 1.019)
Low pollution cohort		
No of admissions	1 638 976	
Total person years	126 479 286	
Median follow-up year	6	
Hazard ratio (95% CI)	1.202 (1.165 to 1.240)	0.980 (0.920 to 1.040)

The low pollution cohort includes subset of cohort with annual exposures <12 $\mu g/m^3$ for fine particles ($PM_{2.5}$) and 20 ppb for nitrogen dioxide (NO_2) over the study period. Hazard ratio: for each 5 $\mu g/m^3$ increase in $PM_{2.5}$ and 10 ppb increase in NO_2.

on race were obtained from Medicare beneficiary files and further divisions of race were not possible owing to the structure of the data. The corresponding total person time of beneficiaries within each stratum was used as the offset. Confidence intervals were calculated using m out of n bootstrap.[40] This study did not consider missing data imputation because the data were complete for the considered variables (eg, age, residential zip code, ICD code). To measure the associations between air pollution and kidney outcomes, we estimated a hazard ratio for each 5 $\mu g/m^3$ increase in $PM_{2.5}$ and 10 ppb increase in NO_2.

For potential confounding, we also adjusted for neighborhood level socioeconomic status indicators.[23 24] These indicators were available at zip code or county level and included population density, proportion of the population below the federally defined poverty level, median household income, racial composition, education level, smoking rate, average body mass index, and average air temperature. Additionally, to adjust potential residual confounding by spatial and temporal trends, we included indicator variables for region (Northeast, Southeast, Midwest, Southwest, and West) and calendar years. Detailed information about the indicators is given in online supplemental material C. Finally, to investigate the potential non-linear association between air pollutants and kidney outcomes, we replaced the linear term of air pollution in the main model with a B spine function with three equally distributed internal knots (at the 25th, 50th, and 75th percentiles of the air pollution concentrations) for each pollutant.[41]

Low pollution cohort

To estimate the associations between air pollution and kidney outcomes at low concentrations, we repeated the main analysis but only included a subset of the main cohort with annual exposures lower than 12 $\mu g/m^3$ and 20 ppb for $PM_{2.5}$ and NO_2, respectively, over the study period. The current health based National Ambient Air Quality Standards (NAAQs) are 12 $\mu g/m^3$ for $PM_{2.5}$ and 53 ppb for NO_2.

Subgroup analysis

To identify subpopulations with higher vulnerability, we evaluated potential effect modification by race, age group (65-74 and ≥75 years), sex, and Medicaid eligibility (as a proxy for low socioeconomic status) by including additional interaction terms between these variables and air pollutants in the main analysis. We also conducted region stratified analyses to assess potential heterogeneities among regions, which could relate to differences in chemical compositions, environmental and climatic conditions, and populations.

Sensitivity analysis

We performed sensitivity analyses to examine whether our main results are robust to the selection of confounders, inclusion of participants with prevalent disease, exposure time window, and potential outcome misclassification. To exclude those with potentially prevalent disease, we repeated the analyses using data excluding beneficiaries who had their first hospital admission for the outcomes in the first two years of follow-up. We also applied one year lag period exposures as an alternative exposure window. Finally, we performed a sensitivity analysis restricting kidney outcomes to those with primary diagnosis codes only.

Patient and public involvement

Patients and the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research. This study did not include

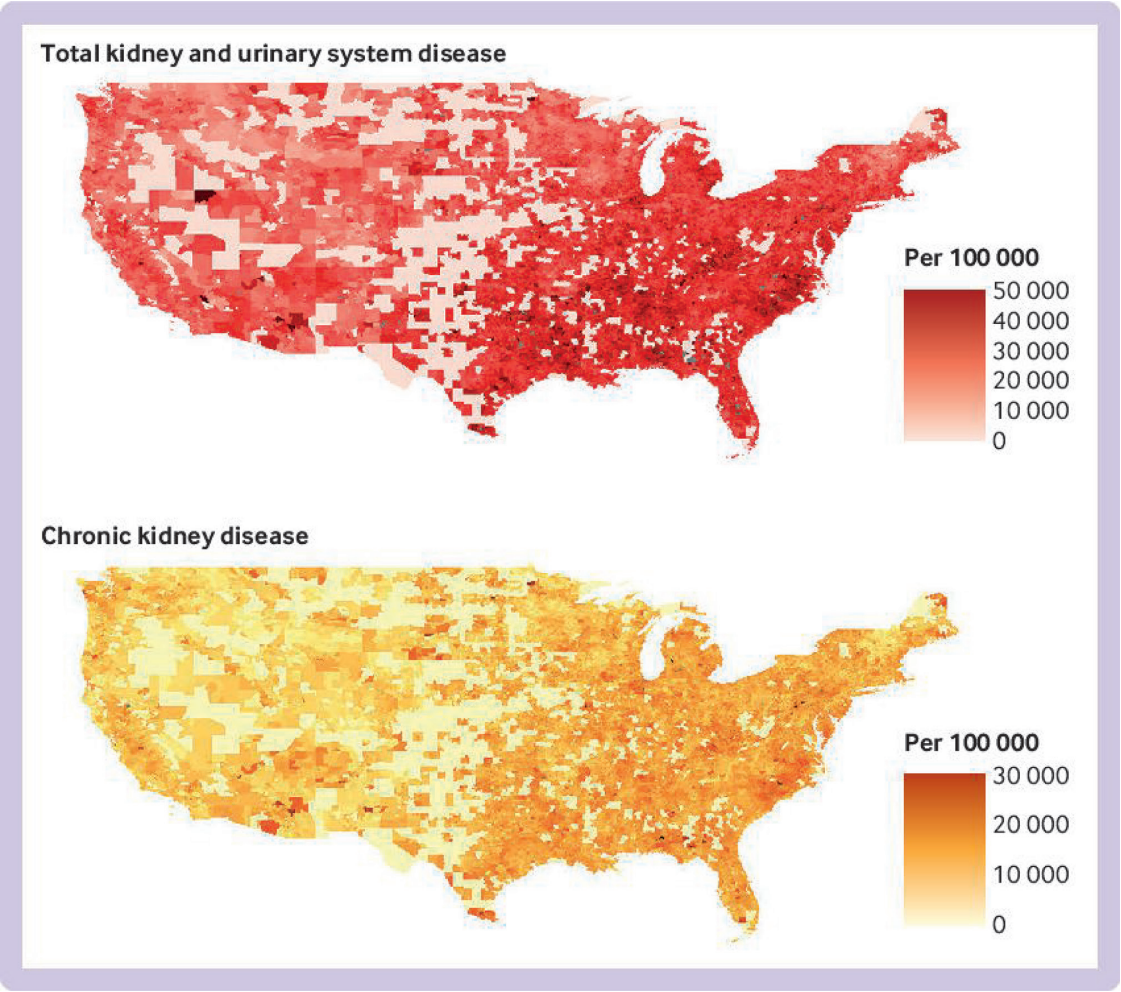

Figure 1 | Nationwide first hospital admissions for total kidney and urinary system disease (upper panel) and chronic kidney disease (lower panel) per 100 000 Medicare fee-for-service beneficiaries across the contiguous United States from 2000 to 2016

plans to recruit participants and only used pre-existing Medicare part A datasets. All data used in this study were prerecorded and deidentified.

Results

The full cohort dataset included 61 097 767 beneficiaries living in 34 849 zip codes. Table 1 presents descriptive information about the beneficiaries. There were 410.3 million person years of follow-up for total kidney and urinary system disease and 454.5 million for CKD (table 2). The median follow-up was five years for total kidney and urinary system disease and six years for CKD. For hospital admissions, 27.6% for total kidney and urinary system disease (5 291 896) and 0.98% for CKD (57 875) were identified as the primary discharge diagnosis code. The total number of first admissions was around 19.0 million for total kidney and urinary system disease and 5.9 million for CKD. Figure 1 shows the geographical distribution of first hospital admissions for total kidney and urinary system disease, and CKD. The low pollution cohort included 113.5 million person years of follow-up for total kidney and urinary system disease and 126.5 million for CKD (table 2).

The average annual concentrations of air pollution over the study period were 9.8 $\mu g/m^3$ for $PM_{2.5}$ and 18.9 ppb for NO_2. The highest $PM_{2.5}$ concentrations were in California and the eastern and southeastern US. The highest NO_2 concentrations were observed in metropolitan areas such as New York, Los Angeles, Denver, and Chicago (figure 2).

In the full cohort, long term exposures to air pollution were positively associated with total kidney and urinary system disease for both pollutants, with hazard ratios of 1.076 (95% confidence interval 1.071 to 1.081) for a 5 $\mu g/m^3$ increase in $PM_{2.5}$ and 1.040 (1.036 to 1.043) for a 10 ppb increase in NO_2 (table 2). In the low pollution cohort, the associations remained significant: 1.133 (1.121 to 1.145) for $PM_{2.5}$ and 1.021 (1.009 to 1.033) for NO_2. CKD also showed positive associations with $PM_{2.5}$ (1.106, 1.097 to 1.115) and NO_2 (1.013, 1.008 to 1.019) in the full cohort. In the low pollution cohort, the risk related to $PM_{2.5}$ was higher in the low pollution cohort than in the full cohort with a hazard ratio of 1.202

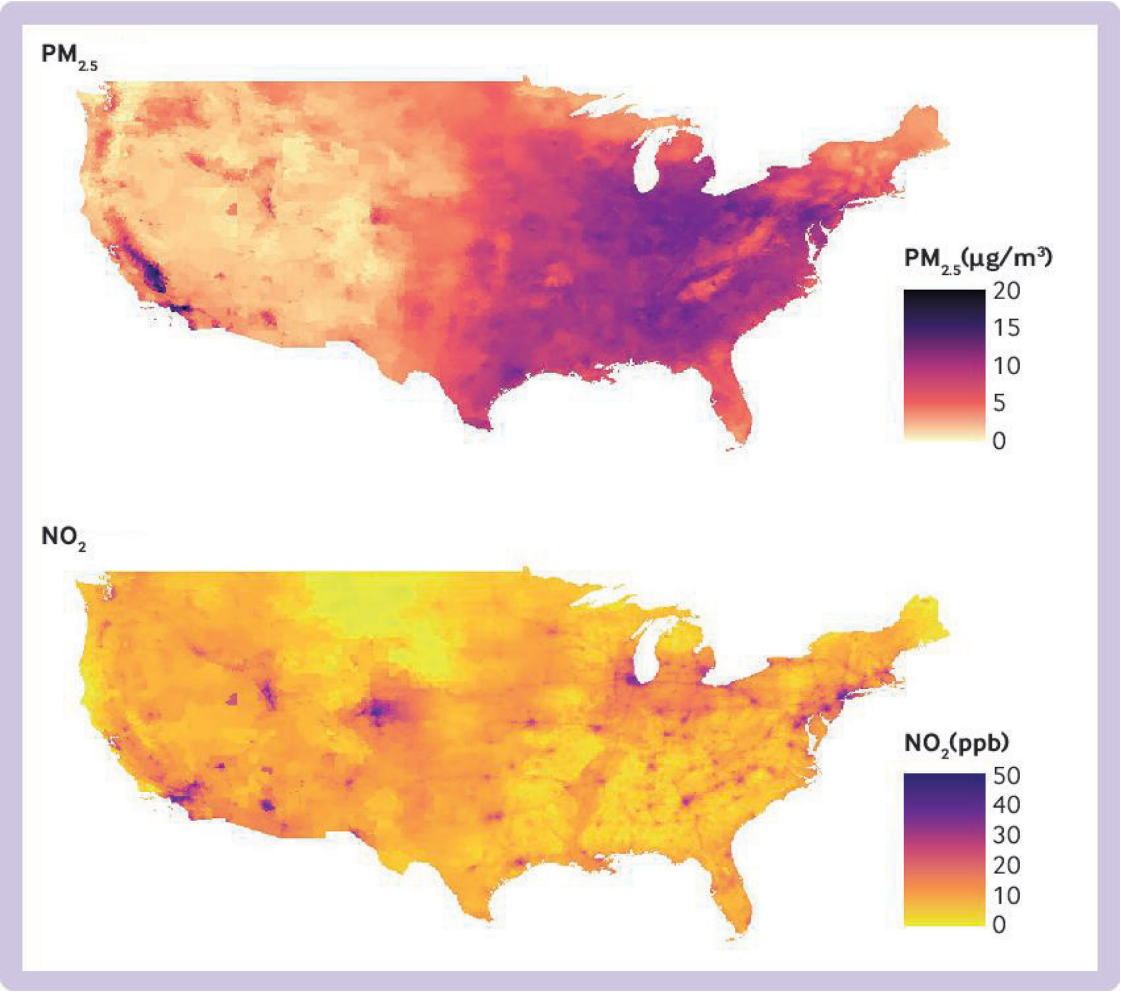

Figure 2 | Nationwide concentrations of air pollution across the contiguous United States from 2000 to 2016. Seventeen year average of annual fine particulate matter ($PM_{2.5}$; upper panel) and nitrogen dioxide (NO_2; lower panel)

(1.165 to 1.240) for the low pollution cohort compared with the full cohort, although there was no evidence of a statistically significant relation with NO_2.

Figure 3 shows the concentration-response associations for each kidney outcome and each pollutant. Table 2 presents results from the linear model. For $PM_{2.5}$, the non-linear association between pollutant level and hazard ratio showed a consistent rise, with an increased risk at higher levels of $PM_{2.5}$ for both kidney outcomes in general, and the slopes for both curves were steeper at concentrations lower than 8 $\mu g/m^3$. NO_2 also showed a consistently increased risk with higher exposure for total kidney and urinary system disease, whereas the association with CKD showed a decreasing pattern with larger confidence intervals for NO_2 concentrations higher than around 20 ppb (approximately 60th percentile of the NO_2 distribution).

Table 3 shows the results from subgroup analyses. For both kidney outcomes, white participants had higher associations with $PM_{2.5}$ compared with the black population, and beneficiaries who were eligible for Medicaid generally had lower estimated associations with both pollutants than those who were not eligible for Medicaid. Women showed lower air pollution related risks on kidney outcomes than men. Additionally, positive associations with air pollution were generally observed in all regions of the US for both kidney outcomes (online supplemental table S1), with the highest $PM_{2.5}$ and NO_2 association estimates observed in the Southwest and Midwest, respectively.

The findings from our sensitivity analysis were generally consistent with the main results. Our findings were robust to adjustment for other confounders, removal of people with prevalent disease, and use of a different exposure window (online supplemental table S2). For total kidney and urinary system disease, exclusion of those identified by secondary diagnostic codes did not change the main results, although all estimates slightly increased; however, CKD showed negative associations with all air pollutants when those identified by secondary diagnostic codes were excluded.

Discussion

Principal findings

This large nationwide prospective cohort study of over 61 million Medicare beneficiaries found that

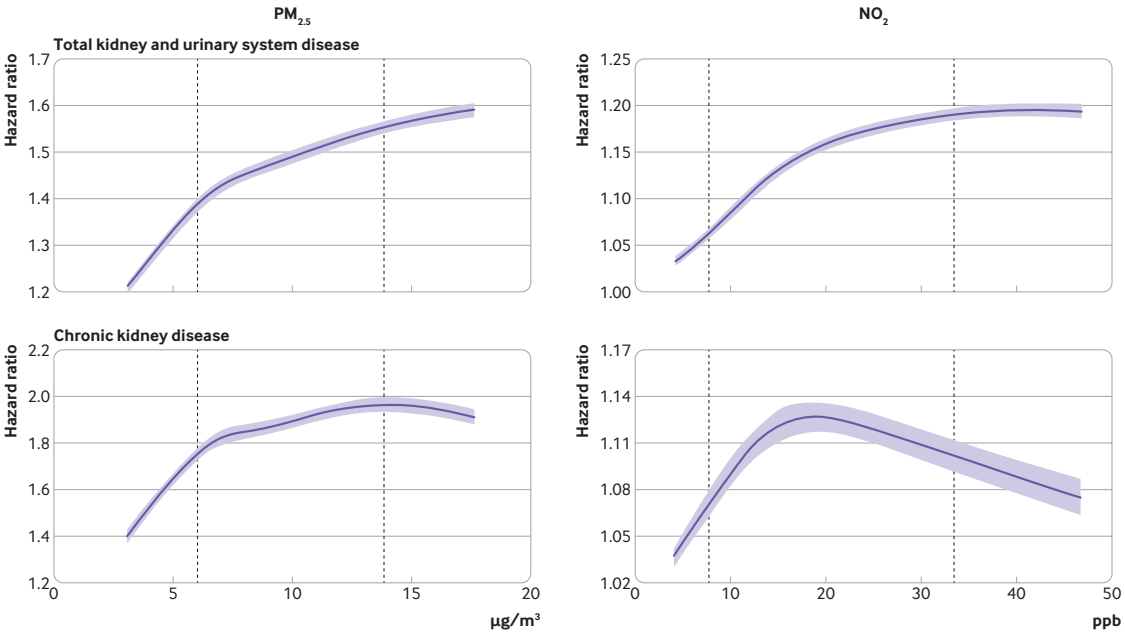

Figure 3 | Concentration-response curves showing association between long term air pollution exposure and kidney diseases. Total kidney and urinary system disease (upper panels) and chronic kidney disease (lower panels). Dashed vertical lines represent 10th and 90th percentiles of each air pollution concentration and shaded areas indicate 95% confidence intervals. Reference exposure points: 0 μg/m^3 for fine particles ($PM_{2.5}$) and 0 ppb for nitrogen dioxide (NO_2)

long term exposures to $PM_{2.5}$ and NO_2 were associated with an increased risk of first hospital admission related to diseases of the kidney and the urinary system or CKD. These associations persisted even at levels below the current annual NAAQS for $PM_{2.5}$ and NO_2. Furthermore, the study found that white people and those not eligible to receive Medicaid are more vulnerable than the general population. These findings suggest that improving air quality can contribute to public health benefits overall and also provide evidence on the benefit of a more stringent air pollution standard for people who potentially have kidney disease.

Table 3 | Association between air pollution and first hospital admission for total kidney and urinary system disease or chronic kidney disease for specific subpopulations

	$PM_{2.5}$	NO_2
Total kidney and urinary system disease		
Age at entry, years		
≥75 *v* 65-74	0.983 (0.979 to 0.987)	1.012 (1.012 to 1.017)
Sex		
Women *v* men	0.985 (0.983 to 0.988)	0.987 (0.985 to 0.989)
Race		
Black *v* white	0.976 (0.969 to 0.982)	1.014 (1.009 to 1.019)
Other *v* white	1.003 (0.991 to 1.015)	0.950 (0.942 to 0.958)
Medicaid eligibility		
Eligible *v* not eligible	0.980 (0.974 to 0.986)	0.962 (0.957 to 0.966)
Chronic kidney disease		
Age at entry, years		
≥75 *v* 65-74	0.932 (0.927 to 0.938)	1.007 (1.004 to 1.010)
Sex		
Women *v* men	0.935 (0.931 to 0.939)	0.955 (0.952 to 0.958)
Race		
Black *v* white	0.988 (0.978 to 0.997)	1.018 (1.013 to 1.024)
Other *v* white	0.993 (0.976 to 1.010)	1.012 (1.006 to 1.019)
Medicaid eligibility		
Eligible *v* not eligible	1.018 (1.011 to 1.026)	0.978 (0.974 to 0.982)

Data are hazard ratio (95% confidence interval) for each 5 μg/m^3 increase in fine particles ($PM_{2.5}$) and 10 ppb increase in nitrogen dioxide (NO_2).
*Other included Asian, Hispanic, American Indian or Alaskan Native, and unknown. Data on race were obtained from Medicare beneficiary files.

Comparision with other studies

The findings of this study are consistent with previous research studying the association between long term exposure to air pollution and kidney diseases. Two recent US Veteran cohort studies (about 2–2.5 million cohort participants who were mostly white men) showed that exposures to higher concentrations of particulate matter no larger than 10 μg/m^3 (PM_{10}), $PM_{2.5}$, carbon monoxide, and NO_2 are associated with increased risk of incident CKD and end stage renal disease.[13 14] Another cohort study conducted in four US counties (10 997 participants) reported that higher annual average $PM_{2.5}$ was associated with increased urinary albumin-creatinine ratio and higher risk of incident CKD.[22] Cohort studies in China (71 151 participants) and Taiwan (around 100 000-160 000 participants) found that exposure to long term air pollution was associated with an increased risk of CKD, end stage renal disease,[20 21] and membranous nephropathy.[42] Nevertheless, owing to the limited cohort participants and different study regions, the findings from these previous studies have limited generalizability. This

study provides more generalizable estimates for the risks of air pollution relating to total kidney and urinary system disease and CKD for the US Medicare part A FFS population by using a cohort of almost 61 million beneficiaries across the contiguous US with high spatial resolution air pollution data.

Strengths and limitations

We observed a significant association between $PM_{2.5}$ and the first hospital admission related to diseases of the kidney and the urinary system at low concentrations of air pollution, suggesting that the health benefit per unit decrease in the concentration of $PM_{2.5}$ might be consistent across concentrations below the current NAAQS. This potential saturation effect (lower air pollution risk at high ranges of annual air pollution) can be explained by smaller risk changes in annual mean air pollution concentration in regions with higher baseline air pollution levels.[41] Furthermore, from the non-linear concentration-response curves, we found no evidence of a threshold value (a concentration below which air pollution exposure does not affect kidney outcomes) for harmful pollution.

The large sample from the Medicare part A FFS cohort enabled us to estimate air pollution risks relating to kidney diseases among subpopulations.[24] However, the results should be carefully interpreted because of small sample sizes for some subpopulations, such as racial minority groups and people eligible for Medicaid, compared with others. Further research is needed that investigates the association between air pollution and kidney disease in potentially vulnerable subpopulations, including potential differences in diagnosis and access to healthcare, especially given the high number of people with undiagnosed disease. The estimated associations between air pollution and kidney outcomes were generally higher in white participants and those not eligible for Medicaid. This finding conflicts with results from previous US studies reporting higher air pollution mortality risk in non-white populations and those eligible for Medicaid,[24 43] although it is consistent with some other study results. A few plausible explanations exist; one possibility is competing risks. A previous study in Taiwan reported that people with comorbidities showed a lower risk related to $PM_{2.5}$ and CKD development.[20] Another study in Korea reported that patients with CKD and healthy lifestyles (normal weight, non-smokers, and non-drinkers) showed a higher air pollution mortality risk than patients with CKD and non-healthy lifestyles.[44] These findings suggest the possibility of competing risk between air pollution and biological or behavioral factors that affects the incidence of kidney disease; the results by race and Medicaid eligibility could be affected by the potential competing risks of underlying health conditions and other risk factors. Another possibility is that the longer life expectancy of high income and white populations could affect the results of this study. The life expectancy of those with high income is longer than that of low income populations,[45] and white populations also have a longer life expectancy than black populations.[46] Because the older population is more vulnerable to the incidence of kidney disease,[11 47] the relatively longer life expectancy of Medicaid non-eligible and white populations could be related to the higher air pollution related risks on kidney outcomes in this study. Finally, because low income and racial or ethnic minority populations generally have less access to medical facilities,[48] underdiagnosis could potentially affect the results. Subgroup results should be examined in future studies with more detailed clinical and socioeconomic indicators.

This study has several limitations. Firstly, the outcome of first hospital admission with diagnosis codes for kidney diseases has limitations when interpreted as the onset of kidney disease. According to the Centers for Disease Control and Prevention, most adults with CKD do not recognize they have CKD,[11] and hospital admission can occur at more advanced stages of the disease or for treating complications attributable to kidney disease. Moreover, kidney disease is generally diagnosed through laboratory tests (eg, the estimated glomerular filtration rate, creatinine, and the accumulation of end products of nitrogen metabolism),[35 49–51] and previous studies have reported that the incidence of kidney disease was substantially underestimated when ICD codes were used to define kidney disease.[52 53] Therefore, the hospital admission records do not fully represent disease incidence, and our cohort probably undercounted the total kidney disease incidence. Additionally, because the use of ICD codes limits the classification of individual disease and the small sample size of specific disease, we did not investigate specific kidney diseases except for CKD and instead used the total kidney and urinary system diseases based on earlier published studies.[28–34] Therefore, we were not able to investigate kidney and urinary disease separately, although an association exists between some kidney and urinary diseases.[30–33] Therefore, our study did not investigate the potential heterogeneity of biological mechanisms among specific kidney and urinary diseases related to air pollution.

Secondly, the potential competing events (eg, hypertension, diabetes, and cardiovascular diseases; not death) that might affect the first hospital admission related to kidney disease[35 49 54] were not considered in this study because of limited data availability. Thirdly, patients with CKD defined in this study represent a subset of those with total kidney disease; these patients were not independent of the total number of patients with kidney and urinary system disease. Therefore, the apparent air pollution related risks for total kidney and urinary system disease could be due to confounding with CKD,[55] and bias could exist in the results of total kidney and urinary system disease resulting from estimating the effects on CKD. Furthermore, additional research is needed on the association between air pollution and CKD in relation to diagnosis, especially given the large number

of participants with undetected disease. In this study CKD was identified as the primary diagnosis in less than 1%; most participants with CKD were identified through hospital admission for comorbidities.[54] Other areas of interest are the impacts of air pollution on recurrent events. This study investigated first hospital admissions only, but subsequent hospital admissions might also be of interest. Importantly, the association between air pollution and CKD was present when primary and secondary ICD codes were considered, but not when the primary code only was considered, which warrants further investigation.

Fourthly, our results only represent the Medicare part A FFS population that does not include all Medicare beneficiaries (the Medicare FFS population covers up to around 75% of the Medicare population).[23] We had no information on Medicare Health Maintenance Organization claims, therefore, we could not cover the entire Medicare population. Fifthly, although we adjusted for several neighborhood level indicators, Medicare claims do not include extensive individual level data on behavioral and socioeconomic risk factors, which could be crucial confounders. Additionally, owing to limitations of the study data, we were not able to consider the potential effect modifications by comorbidities and CKD stages. These unmeasured individual level confounders might help to elucidate the biological mechanisms underlying the observed association between air pollution and kidney disease, and they should be looked at in future studies. Finally, future research could also investigate complex air pollution mixtures. We focused on traffic related air pollutants of NO_2 and $PM_{2.5}$, although each of these pollutants also has non-traffic sources, and actual air pollution mixtures are complexes with other pollutants and different chemical structures of $PM_{2.5}$. We also used state-of-the-science air pollution prediction data, which are only available until 2016; therefore, future studies should investigate the association between air pollution and kidney disease with more recent data.

Conclusion

We found that long term exposures to higher levels of air pollution were associated with increased risk of a first hospital admission related to total kidney and urinary system disease or CKD, even at low pollution concentrations. Our findings suggest that improvement of air quality might have some benefits in reducing kidney and urinary system disease.

AUTHOR AFFILIATIONS

[1]School of the Environment, Yale University, New Haven, CT, USA
[2]Department of Environmental Medicine, College of Medicine, Ewha Womans University, Seoul, Republic of Korea
[3]Institute of Ewha-SCL for Environmental Health (IESEH), Seoul, Republic of Korea
[4]Department of Biostatistics, Harvard University T H Chan School of Public Health, Boston, MA, USA
[5]Harvard University Faculty of Arts and Sciences, Cambridge, MA, USA
[6]Harvard University T H Chan School of Public Health, Boston, MA, USA
[7]Department of Data Science, Dana-Farber Cancer Institute, Boston, MA, USA
[8]Department of Internal Medicine, Dongguk University Ilsan Hospital, Goyang, Gyeonggi-do, Republic of Korea
[9]Department of Internal Medicine, Dongguk University College of Medicine, Goyang, Gyeonggi-do, Republic of Korea
[10]Department of Internal Medicine, Seoul National University Hospital, Jongno-gu, Seoul, Republic of Korea
[11]Department of Internal Medicine, Seoul National University Boramae Medical Center, Seoul, Republic of Korea
[12]Environmental Health, Harvard University T H Chan School of Public Health, Boston, MA, USA
[13]Graduate School of Public Health, Seoul National University, Gwanak-gu, Seoul, Republic of Korea
[14]Institute for Sustainable Development, Graduate School of Public Health, Seoul, Republic of Korea

Acknowledgements This study was supported by the project titled "Institute of Ewha-SCL for Environmental Health (IESEH) and Research of Environmental Examination Model for Children and Women (No 1-2022-0205-001-1)."

Contributors WL is the guarantor; had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. WL was responsible for the study concept and design, drafting of the manuscript, statistical analysis, and final approval of the article to be published. MBS, DB, and JS were responsible for the acquisition of data, analysis, or interpretation of data. WL, XW, SH, KCF, J-YS, DB, JYP, YCK, JPL, FD, HK and MLB contributed to the critical revision of the manuscript for important intellectual content. MB obtained funding, and provided administrative, technical, or material support. WL and MLB supervised the study. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted. Transparency: The lead author (the manuscript's guarantor) affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

Funding This paper was developed under Assistance Agreement No RD835871 awarded by the US Environmental Protection Agency to Yale University. It has not been formally reviewed by EPA. WL was supported by the 2020 Science and Technology Subsequent Generation Support Project (NRF-2021R1A6A3A03038675), implemented by the National Research Foundation of Korea.

Disclaimer The views expressed in this document are solely those of MB and other co-authors and do not necessarily reflect those of the Agency. EPA does not endorse any products or commercial services mentioned in this publication.

Competing interests Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/disclosure-of-interest/ and declare: support from the US Environmental Protection Agency and National Research Foundation of Korea for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Ethics approval This study was approved by the Yale institutional review board (protocol No 1608018335).

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iDs
Whanhee Lee http://orcid.org/0000-0001-5723-9061
Jung Pyo Lee http://orcid.org/0000-0002-4714-1260

REFERENCES

1 Gansevoort RT, Correa-Rotter R, Hemmelgarn BR, *et al*. Chronic kidney disease and cardiovascular risk: epidemiology, mechanisms, and prevention. *Lancet* 2013;382:339–52. doi:10.1016/S0140-6736(13)60595-4
2 Luyckx VA, Al-Aly Z, Bello AK. Sustainable development goals relevant to kidney health: an update on progress. *Nature Reviews Nephrology*2020:1–18. doi:10.1038/s41581-021-00473-9
3 Perlman RL, Finkelstein FO, Liu L, *et al*. Quality of life in chronic kidney disease (CKD): a cross-sectional analysis in the renal research Institute-CKD study. *Am J Kidney Dis* 2005;45:658–66. doi:10.1053/j.ajkd.2004.12.021
4 Jager KJ, Kovesdy C, Langham R, *et al*. A single number for advocacy and communication—worldwide more than 850 million individuals have kidney diseases. *Nephrology Dialysis Transplantation* 2019;34:1803–5. doi:10.1093/ndt/gfz174
5 GBD 2015 DALYs and HALE Collaborators. Global, regional, and national disability-adjusted life-years (DALYs) for 315 diseases and injuries and healthy life expectancy (HALE), 1990-2015: a systematic analysis for the global burden of disease study 2015. *Lancet* 2016;388:1603–58. doi:10.1016/S0140-6736(16)31460-X
6 Luyckx VA, Tonelli M, Stanifer JW. The global burden of kidney disease and the sustainable development goals. *Bull World Health Organ* 2018;96:414–22. doi:10.2471/BLT.17.206441
7 Bikbov B, Purcell CA, Levey AS, *et al*. Global, regional, and national burden of chronic kidney disease, 1990–2017: a systematic analysis for the global burden of disease study 2017. *The Lancet* 2020;395:709–33. doi:10.1016/S0140-6736(20)30045-3
8 Lameire NH, Bagga A, Cruz D, *et al*. Acute kidney injury: an increasing global concern. *Lancet* 2013;382:170–9. doi:10.1016/S0140-6736(13)60647-9
9 Levin A, Tonelli M, Bonventre J, *et al*. Global kidney health 2017 and beyond: a roadmap for closing gaps in care, research, and policy. *Lancet* 2017;390:1888–917. doi:10.1016/S0140-6736(17)30788-2
10 Foreman KJ, Marquez N, Dolgert A, *et al*. Forecasting life expectancy, years of life lost, and all-cause and cause-specific mortality for 250 causes of death: reference and alternative scenarios for 2016–40 for 195 countries and territories. *The Lancet* 2018;392:2052–90. doi:10.1016/S0140-6736(18)31694-5
11 Centers for Disease Control and Prevention. *Chronic kidney disease in the United States, 2021*. Atlanta, GA: US Department of Health and Human Services, Centers for Disease Control and Prevention, 2021.
12 United States Renal Data System. *2020 USRDS annual data report: epidemiology of kidney disease in the United States*. Bethesda, MD: National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases.
13 Bowe B, Xie Y, Li T, *et al*. Associations of ambient coarse particulate matter, nitrogen dioxide, and carbon monoxide with the risk of kidney disease: a cohort study. *Lancet Planet Health* 2017;1:e267–76. doi:10.1016/S2542-5196(17)30117-1
14 Bowe B, Xie Y, Li T, *et al*. Particulate matter air pollution and the risk of incident CKD and progression to ESRD. *J Am Soc Nephrol* 2018;29:218–30. doi:10.1681/ASN.2017030253
15 Ye J-J, Wang S-S, Fang Y, *et al*. Ambient air pollution exposure and risk of chronic kidney disease: a systematic review of the literature and meta-analysis. *Environ Res* 2021;195:110867. doi:10.1016/j.envres.2021.110867
16 Xu X, Nie S, Ding H, *et al*. Environmental pollution and kidney diseases. *Nat Rev Nephrol* 2018;14:313–24. doi:10.1038/nrneph.2018.11
17 Mehta AJ, Zanobetti A, Bind M-AC, *et al*. Long-Term exposure to ambient fine particulate matter and renal function in older men: the Veterans administration normative aging study. *Environ Health Perspect* 2016;124:1353–60. doi:10.1289/ehp.1510269
18 Chin MT. Basic mechanisms for adverse cardiovascular events associated with air pollution. *Heart* 2015;101:253–6. doi:10.1136/heartjnl-2014-306379
19 Rajagopalan S, Brook RD. Air pollution and type 2 diabetes: mechanistic insights. *Diabetes* 2012;61:3037–45. doi:10.2337/db12-0190
20 Chan T-C, Zhang Z, Lin B-C, *et al*. Long-Term exposure to ambient fine particulate matter and chronic kidney disease: a cohort study. *Environ Health Perspect* 2018;126:107002. doi:10.1289/EHP3304
21 Lin S-Y, Ju S-W, Lin CL, *et al*. Air pollutants and subsequent risk of chronic kidney disease and end-stage renal disease: a population-based cohort study. *Environ Pollut* 2020;261:114154. doi:10.1016/j.envpol.2020.114154
22 Blum MF, Surapaneni A, Stewart JD, *et al*. Particulate matter and albuminuria, glomerular filtration rate, and incident CKD. *Clin J Am Soc Nephrol* 2020;15:311–9. doi:10.2215/CJN.08350719
23 Shi L, Wu X, Danesh Yazdi M, Yazdi MD, *et al*. Long-term effects of $PM_{2.5}$ on neurological disorders in the American Medicare population: a longitudinal cohort study. *Lancet Planet Health* 2020;4:e557–65. doi:10.1016/S2542-5196(20)30227-8
24 Di Q, Wang Y, Zanobetti A, *et al*. Air pollution and mortality in the Medicare population. *N Engl J Med Overseas Ed* 2017;376:2513–22. doi:10.1056/NEJMoa1702747
25 Kemec Z, Tüzün C, Gürel A. Neurogenic bladder and acute kidney injury in leukodystrophy. *Cureus* 2020;12:e8707. doi:10.7759/cureus.8707
26 Salo J, Ikäheimo R, Tapiainen T, *et al*. Childhood urinary tract infections as a cause of chronic kidney disease. *Pediatrics* 2011;128:840–7. doi:10.1542/peds.2010-3520
27 Alelign T, Petros B. Kidney stone disease: an update on current concepts. *Adv Urol* 2018;2018:1–12. doi:10.1155/2018/3068365
28 Hansen AL, Bi P, Ryan P, *et al*. The effect of heat waves on hospital admissions for renal disease in a temperate city of Australia. *Int J Epidemiol* 2008;37:1359–65. doi:10.1093/ije/dyn165
29 Kim E, Kim H, Kim YC, *et al*. Association between extreme temperature and kidney disease in South Korea, 2003–2013: stratified by sex and age groups. *Sci Total Environ* 2018;642:800–8. doi:10.1016/j.scitotenv.2018.06.055
30 Hansen A, Bi P, Ryan P. The effect of Heatwaves on hospital admissions for renal disease in Adelaide, South Australia. *Epidemiology* 2008;19:S105–6.
31 Borg M, Bi P, Nitschke M, *et al*. The impact of daily temperature on renal disease incidence: an ecological study. *Environmental Health* 2017;16:1–30. doi:10.1186/s12940-017-0331-4
32 Liu Y, Hoppe BO, Convertino M. Threshold evaluation of emergency risk communication for health risks related to hazardous ambient temperature. *Risk Anal* 2018;38:2208–21. doi:10.1111/risa.12998
33 Kovats RS, Hajat S, Wilkinson P. Contrasting patterns of mortality and hospital admissions during hot weather and heat waves in greater London, UK. *Occup Environ Med* 2004;61:893–8. doi:10.1136/oem.2003.012047
34 Ballotari P, Ranieri SC, Luberto F, *et al*. Sex differences in cardiovascular mortality in diabetics and nondiabetic subjects: a population-based study (Italy). *Int J Endocrinol* 2015;2015:1–10. doi:10.1155/2015/914057
35 Webster AC, Nagler EV, Morton RL, *et al*. Chronic kidney disease. *The Lancet* 2017;389:1238–52. doi:10.1016/S0140-6736(16)32064-5
36 Di Q, Kloog I, Koutrakis P, *et al*. Assessing PM2.5 exposures with high spatiotemporal resolution across the continental United States. *Environ Sci Technol* 2016;50:4712–21. doi:10.1021/acs.est.5b06121
37 Di Q, Amini H, Shi L, *et al*. Assessing NO_2 concentration and model uncertainty with high spatiotemporal resolution across the contiguous United States using ensemble model averaging. *Environ Sci Technol* 2020;54:1372–84. doi:10.1021/acs.est.9b03358
38 Wei Y, Yazdi MD, Di Q, *et al*. Emulating causal dose-response relations between air pollutants and mortality in the Medicare population. *Environmental Health* 2021;20:1–10. doi:10.1186/s12940-021-00742-x
39 Andersen PK, Gill RD. Cox's regression model for counting processes: a large sample study. *Ann Stat* 1982:1100–20. doi:10.1214/aos/1176345976
40 Bickel PJ, Götze F, van Zwet WR. *Resampling fewer than N observations: gains, losses, and remedies for losses*. Selected works of Willem van Zwet: Springer, 2012: 267–97.
41 Liu C, Chen R, Sera F, *et al*. Ambient particulate air pollution and daily mortality in 652 cities. *N Engl J Med* 2019;381:705–15. doi:10.1056/NEJMoa1817364
42 Xu X, Wang G, Chen N, *et al*. Long-Term exposure to air pollution and increased risk of membranous nephropathy in China. *J Am Soc Nephrol* 2016;27:3739–46. doi:10.1681/ASN.2016010093
43 Pope CA, Lefler JS, Ezzati M, *et al*. Mortality risk and fine particulate air pollution in a large, representative cohort of U.S. adults. *Environ Health Perspect* 2019;127:077007. doi:10.1289/EHP4438
44 Jung J, Park JY, Kim YC, *et al*. Effects of air pollution on mortality of patients with chronic kidney disease: a large observational cohort study. *Sci Total Environ* 2021;786:147471. doi:10.1016/j.scitotenv.2021.147471
45 National Academies of Sciences E, Medicine Committee on Population. *The growing gap in life expectancy by income: implications for federal programs and policy responses*. National Academies Press, 2015.

46 Arias E, JQ X. United States life tables, 2018. National vital statistics reports. *National Center for Health Statistics* 2020;69.
47 Ravani P, Quinn R, Fiocco M, *et al*. Association of age with risk of kidney failure in adults with stage IV chronic kidney disease in Canada. *JAMA Netw Open* 2020;3:e2017150–e50. doi:10.1001/jamanetworkopen.2020.17150
48 Dubay LC, Lebrun LA. Health, behavior, and health care disparities: disentangling the effects of income and race in the United States. *Int J Health Serv* 2012;42:607–25. doi:10.2190/HS.42.4.c
49 Bellomo R, Kellum JA, Ronco C. Acute kidney injury. *The Lancet* 2012;380:756–66. doi:10.1016/S0140-6736(11)61454-2
50 Levey AS, Coresh J. Chronic kidney disease. *The Lancet* 2012;379:165–80. doi:10.1016/S0140-6736(11)60178-5
51 Gaitonde DY, Cook DL, Rivera IM. Chronic kidney disease: detection and evaluation. *Am Fam Physician* 2017;96:776–83.
52 Logan R, Davey P, De Souza N, *et al*. Assessing the accuracy of ICD-10 coding for measuring rates of and mortality from acute kidney injury and the impact of electronic alerts: an observational cohort study. *Clin Kidney J* 2020;13:1083-1090. doi:10.1093/ckj/sfz117
53 Vlasschaert MEO, Bejaimal SAD, Hackam DG, *et al*. Validity of administrative database coding for kidney disease: a systematic review. *Am J Kidney Dis* 2011;57:29–43. doi:10.1053/j.ajkd.2010.08.031
54 Chen TK, Knicely DH, Grams ME. Chronic kidney disease diagnosis and management. *JAMA* 2019;322:1294. doi:10.1001/jama.2019.14745
55 Sun X, Briel M, Walter SD, *et al*. Is a subgroup effect believable? updating criteria to evaluate the credibility of subgroup analyses. *BMJ* 2010;340:c117. doi:10.1136/bmj.c117

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/bmjmed-2021-000009).

Supplementary Appendix

Associations between long-term air pollution and kidney diseases in the US Medicare population

Whanhee Lee[1], Xiao Wu[2], Seulkee Heo[1], Kelvin C. Fong[1], Ji-Young Son[1], M. Benjamin Sabath[3], Danielle Braun[2,4], Jae Yoon Park[5,6], Yong Chul Kim[7], Jung Pyo Lee[7,8], Joel Schwartz[9], Ho Kim[10,11], Francesca Dominici[2], Michelle L. Bell[1]

Correspondence to: Whanhee Lee, PhD. Yale School of the Environment, Yale University, (Address) 195 Prospect Street, New Haven, CT 06511, USA. Telephone: (1) 203 432-9869. Fax: (1) 203 466-9158. E-mail: whanhee.lee@yale.edu

[1]Yale School of the Environment, Yale University, New Haven, CT, USA.

[2]Department of Biostatistics, Harvard T H Chan School of Public Health, Boston, MA, USA.

[3]Faculty of Arts and Sciences Research Computing Department, Harvard University, Boston, MA, USA.

[4]Department of Data Science, Dana-Farber Cancer Institute, Boston, MA, USA.

[5]Department of Internal Medicine, Dongguk University Ilsan Hospital, Goyang, Republic of Korea.

[6]Department of Internal Medicine, Dongguk University College of Medicine, Goyang, Republic of Korea.

[7]Department of Internal Medicine, Seoul National University Hospital, Seoul, Seoul, Republic of Korea.

[8]Department of Internal Medicine, Seoul National University Boramae Medical Center, Seoul, Republic of Korea.

[9]Department of Environmental Health, Harvard T H Chan School of Public Health, Boston, MA, USA.

[10]Department of Public Health Science, Graduate School of Public Health, Seoul National University, Seoul, Republic of Korea.

[11]Institute for Sustainable Development, Graduate School of Public Health, Seoul National University, Seoul, Republic of Korea.

Lee W, *et al. BMJMED* 2022; 1:e000009. doi: 10.1136/bmjmed-2021-000009

A. STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	page 1 (Title)
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	page 3 (Abstract)
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	page 4 (Introduction)
Objectives	3	State specific objectives, including any prespecified hypotheses	page 4 (Introduction)
Methods			
Study design	4	Present key elements of study design early in the paper	page 4-5 (Methods – Study design and participants)
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	page 4-5 (Methods – Study design and participants)
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	page 4-5 (Methods – Study design and participants)
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case	NA.
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	page 4-5 (Methods)
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	page 4-6 (Methods)
Bias	9	Describe any efforts to address potential sources of bias	page 6 (Sensitivity analysis)
Study size	10	Explain how the study size was arrived at	page 4-5 (Methods – Study design and participants)
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were	page 4-6 (Methods)

		chosen and why	
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	page 5 (Methods – Statistical Analysis)
		(*b*) Describe any methods used to examine subgroups and interactions	page 6 (Methods – Subgroup analysis)
		(*c*) Explain how missing data were addressed	page 5 (Methods – Statistical Analysis)
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	NA.
		(*e*) Describe any sensitivity analyses	page 6 (Methods – Sensitivity analysis)

Continued on next page

Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	page 6 (Results)
		(b) Give reasons for non-participation at each stage	NA.
		(c) Consider use of a flow diagram	-
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	page 6-7 (Results), Figure 1-2, Table 1
		(b) Indicate number of participants with missing data for each variable of interest	NA.
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	page 6 (Results), Table 2
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	page 6 (Results), Table 2, Figure 1
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	-
		Cross-sectional study—Report numbers of outcome events or summary measures	-
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	page 6-7 (Results), Table 2-3
		(*b*) Report category boundaries when continuous variables were categorized	page 5 (Statistical Analysis) Table 1, 3
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	-
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	page 6-7 (Results), Table 3, Table S3-4
Discussion			
Key results	18	Summarise key results with reference to study objectives	page 7 (Discussion)
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	page 7 (Discussion)
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	page 7 (Discussion)
Generalisability	21	Discuss the generalisability (external validity) of the study results	page 7 (Discussion)
Other information			

Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	page 12 (Funding)

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

B. Details on air pollution data

The predictions for $PM_{2.5}$ and NO_2 were estimated by hybrid ensemble models incorporating random forest, gradient boosting, and neural network. Multiple predictor variables including monitoring data, satellite data, meteorological conditions, land-use variables, and chemical transport model simulation values were included in these models. The technical details of the prediction models have been previously published with excellent performance.[1-3]

$PM_{2.5}$. We collected daily $PM_{2.5}$ predictions at a lkm^2 spatial resolution across the 48 contiguous US states and Washington DC from a well-validated ensemble model.[1] The model included more than 100 predictors from monitoring data, satellite-based aerosol optical depth (AOD) data, chemical transport model outputs, meteorological data, aerosol index, and land-use data. The model was calibrated with daily $PM_{2.5}$ concentrations measured at 2156 monitors operated by the US Environmental Protection Agency's Air Quality System database and IMPROVE monitoring network.[4] For AOD data, the algorithm called MAIAC that retrieves AOD with a spatial resolution of $1km^2$ from the Moderate Resolution Imaging Spectroradiometer (MODIS) was used. The GEOS-Chem chemical transport model, a global 3-dimensional chemical transport model, was used to simulate ground-level $PM_{2.5}$ concentrations. Meteorological variables were collected from NCEP North American Regional Reanalysis data, which includes daily estimated meteorological variables (at 0.3º grid cells), such as air temperature, accumulated total precipitation, downward shortwave radiation flux, wind speed, and humidity. Absorbing aerosol index (AAI) was also considered in the prediction model to consider absorbing aerosols, such as organic carbon and soil dust. In addition, land-use data were used to consider elevation, road density, emission inventory, population density, % of urban, and residential greenness. R^2 between cross-validated predicted $PM_{2.5}$ and monitored $PM_{2.5}$ was calculated to quantify model performance, and 10-fold cross-validated R^2 of 0.89 for annual $PM_{2.5}$ predictions across the US.

NO_2. We collected daily NO2 predictions at a lkm^2 spatial resolution across the 48 contiguous US states and Washington DC from a well-validated ensemble model.[2] Total 912 No2 monitoring sites operated by the US Environmental Protection Agency were included in this prediction model. First, 16 meteorological variables from the National Centers for Environmental Prediction (NCEP) and National Center for Atmospheric Research (NCAR), such as surface air temperature, accumulated total precipitation, specific humidity at 2m, and medium cloud area fraction, were used with spatial resolution of approximately 32 km. NO_2 column density (from the Aura satellite) and chemical transport models (the global-scale GEOS-Chem and the regional-scale Community Multiscale Air Quality model) were used to estimate surface-level NO_2 as a predictor variable in the modeling. In addition, seven categories of land-cover variables (land-cover types, truck traffic, road density, restaurant density, elevation, normalized difference vegetation index, and nighttime light) were considered as predictor variables in the prediction model. Finally, other ancillary variables, such as variables related to aerosol concentration and aerosol type, cloud coverage, surface albedo/reflectance, were also considered in this model. R^2 between cross-validated predicted NO_2 and monitored NO_2 was calculated to quantify model performance, and 10-fold cross-validated R^2 of 0.84 for annual NO_2 predictions across the US.

C. Details on neighborhood-level indicators

We adjusted for a total of 12 neighborhood-level indicators in the main model to consider potential confounding: eight ZIP code-level indicators, two county-level indicators, average temperature for each ZIP code, and indicator variables indicating geographical regions.

ZIP code-level indicators: Eight indicators available at ZIP Code Tabulation Areas (ZCTA) level were derived from the 2000 U.S. Census, the 2010 U.S. Census, and the American Community Survey (ACS) from 2005-2016. If indicators were missing for a year, we linearly interpolated or extrapolated their values using available data. The ZCTA indicators included the % of the population below the poverty level, population density (persons per km^2), median home value (US $), % of the population that is Black, % of the population that is Hispanic, % of the population with other race (not Black or White), median household income (US $), % of homes with owner-occupied housing and % of the population without a high school education. These ZCTA data were matched to ZIP code.

County-level indicators: Two county-level indicators (average body mass index (BMI) and % of the population that had ever smoked) were collected from the Behavioral Risk Factor Surveillance System (BRFSS) for the period of 2000-2016. These county-level indicators were matched to ZIP code if the ZIP code centroids fell within the country boundary.

ZIP code-level average temperature: We collected annual average air temperature (2m) across the continental US (2000-2016), which was provided from the North American Regional Reanalysis (NARR) with grids that were approximately 32km*32km, and assigned the annual average temperature for each ZIP code based on the nearest grid cell for each ZIP code centroid.

Region indicator: We used indicator variables for regions in the US (Northeast, Southeast, Midwest, Southwest, and West).

D. Supplementary Tables

Table S1. Region-specific association (HR and 95% CI) between air pollution and the first hospital admission due to total renal system disease or chronic kidney disease (CKD) in Medicare Part A FFS 2000-2016. Hazard ratio: $PM_{2.5}$ (per 5 $\mu g/m^3$) and NO_2 (per 10 ppb).

	$PM_{2.5}$	NO_2
Total renal system disease		
Midwest	1.080 (1.066, 1.094)	1.113 (1.085, 1.141)
Northeast	1.070 (1.057, 1.083)	0.998 (0.992, 1.004)
Southeast	1.049 (1.037, 1.062)	1.038 (1.031, 1.046)
Southwest	1.132 (1.113, 1.151)	1.001 (0.991, 1.010)
West	0.990 (0.980, 0.999)	1.055 (1.048, 1.061)
Chronic kidney disease (CKD)		
Midwest	1.129 (1.105, 1.154)	1.093 (1.046, 1.141)
Northeast	1.069 (1.046, 1.093)	0.978 (0.968, 0.988)
Southeast	1.043 (1.023, 1.063)	1.030 (1.016, 1.043)
Southwest	1.193 (1.161, 1.226)	0.992 (0.978, 1.006)
West	0.979 (0.962, 0.996)	1.029 (1.019, 1.040)

Table S2. Sensitivity analysis of the association (HR and 95% CI) between air pollution and risk of first hospital admissions for total renal system disease or chronic kidney disease (CKD) based on: selection of confounders, inclusion of prevalent cases, exposure time window, and potential outcome misclassification in Medicare Part A FFS 2000-2016. Hazard ratio: $PM_{2.5}$ (per 5 $\mu g/m^3$) and NO_2 (per 10 ppb).

	$PM_{2.5}$	**NO_2**
Total renal system disease		
Main model	1.076 (1.071, 1.081)	1.040 (1.036, 1.043)
Excluding neighborhood-level confounders	1.105 (1.099, 1.110)	1.002 (0.998, 1.006)
Excluding region indicator variables	1.089 (1.097, 1.112)	1.037 (1.033, 1.040)
Primary diagnosis code only	1.104 (1.097, 1.112)	1.062 (1.047, 1.076)
Exposure with 1-year lag period	1.074 (1.069, 1.079)	1.041 (1.038, 1.045)
Excluding potential prevalent cases (exclusion of the first 2 years of follow-up)	1.083 (1.077, 1.088)	1.044 (1.040, 1.048)
Chronic kidney disease (CKD)		
Main model	1.106 (1.097, 1.115)	1.013 (1.008, 1.019)
Excluding neighborhood-level confounders	1.145 (1.136, 1.154)	0.953 (0.948, 0.958)
Excluding region indicator variables	1.128 (1.119, 1.138)	1.017 (1.012, 1.022)
Primary diagnosis code only	0.987 (0.963, 1.012)	0.940 (0.895, 0.987)
Exposure with 1-year lag period	1.103 (1.094, 1.112)	1.013 (1.008, 1.018)
Excluding potential prevalent cases (exclusion of the first 2 years of follow-up)	1.116 (1.106, 1.125)	1.015 (1.009, 1.020)

References for the Appendix

1. Di Q, Kloog I, Koutrakis P, et al. Assessing PM2. 5 exposures with high spatiotemporal resolution across the continental United States. *Environmental science technology* 2016;50(9):4712-21.
2. Di Q, Amini H, Shi L, et al. Assessing NO2 concentration and model uncertainty with high spatiotemporal resolution across the contiguous United States using ensemble model averaging. *Environmental science technology* 2019;54(3):1372-84.
3. Wei Y, Yazdi MD, Di Q, et al. Emulating causal dose-response relations between air pollutants and mortality in the Medicare population. *Environmental Health* 2021;20(1):1-10.
4. Shi L, Wu X, Yazdi MD, et al. Long-term effects of PM2· 5 on neurological disorders in the American Medicare population: a longitudinal cohort study. 2020