A Section 508–conformant HTML version of this article is available at https://doi.org/10.1289/EHP10710.

Association of Combined Exposure to Ambient Air Pollutants, Genetic Risk, and Incident Rheumatoid Arthritis: A Prospective Cohort Study in the UK Biobank

Jie Zhang,[1,2] Xin-Yu Fang,[1,2*] Jun Wu,[3] Yin-Guang Fan,[1,2] Rui-Xue Leng,[1,2] Bo Liu,[1,2] Xiao-Jie Lv,[1,2] Yu-Lu Yan,[1,2] Chen Mao,[4] and Dong-Qing Ye[1,2]*

[1]Department of Epidemiology and Biostatistics, School of Public Health, Anhui Medical University, Hefei, Anhui, China
[2]Inflammation and Immune Mediated Diseases Laboratory of Anhui Province, Hefei, Anhui, China
[3]Department of Epidemiology and Biostatistics, School of Public Health, Guangdong Medical University, Dongguan, Guangdong, China
[4]Department of Epidemiology, School of Public Health, Southern Medical University, Guangzhou, Guangdong, China

Background: Evidence for a potential link between air pollution and rheumatoid arthritis (RA) is inconsistent, and the modified effect of genetic susceptibility on the relationship between air pollution and RA has not been well studied.

Objective: Using a general population cohort from the UK Biobank, this study aimed to investigate the associations between various air pollutants and the risk of incident RA and to further estimate the impact of combined exposure to ambient air pollutants on the risk of developing RA under the modification effect of genetic predisposition.

Methods: A total of 342,973 participants with completed genotyping data and who were free of RA at baseline were included in the study. An air pollution score was constructed by summing the concentrations of each pollutant weighted by the regression coefficients with RA from single-pollutant models to assess the combined effect of air pollutants, including particulate matter (PM) with diameters $\leq$2.5 μm ($PM_{2.5}$), between 2.5 and 10 μm ($PM_{2.5–10}$), and $\leq$10 μm (PM_{10}), as well as nitrogen dioxide (NO_2) and nitrogen oxides (NO_x). In addition, the polygenic risk score (PRS) of RA was calculated to characterize individual genetic risk. The Cox proportional hazard model was used to estimate hazard ratios (HRs) and 95% confidence intervals (95% CIs) of associations of single air pollutant, air pollution score, or PRS with incident RA.

Results: During a median follow-up time of 8.1 y, 2,034 incident events of RA were recorded. The HRs (95% CIs) of incident RA per interquartile range increment in $PM_{2.5}$, $PM_{2.5–10}$, PM_{10}, NO_2, and NO_x were 1.07 (1.01, 1.13), 1.00 (0.96, 1.04), 1.01 (0.96, 1.07), 1.03 (0.98, 1.09), and 1.07 (1.02, 1.12), respectively. We also found a positive exposure–response relationship between air pollution score and RA risk ($p_{Trend} = 0.000053$). The HR (95% CI) of incident RA was 1.14 (1.00, 1.29) in the highest quartile group compared with the lowest quartile group of the air pollution score. Furthermore, the results of the combined effect of air pollution score and PRS on the RA risk showed that the risk of RA incidence in the highest genetic risk and air pollution score group was almost twice that of the lowest genetic risk and air pollution score group [incidence rate (IR) per 100,000 person-years: 98.46 vs. 51.19, and HR = 1.73 (95% CI: 1.39, 2.17) vs. 1 (reference)], although no statistically significant interaction between the air pollution and genetic risk for incident RA was found ($p_{Interaction} > 0.05$).

Discussion: The results revealed that long-term combined exposure to ambient air pollutants might increase the risk of RA, particularly in those with high genetic risk. https://doi.org/10.1289/EHP10710

Introduction

Rheumatoid arthritis (RA) is a chronic systematic autoimmune disorder characterized by progressive joint erosion that leads to severe disability.[1] As one of the most prevalent chronic inflammatory diseases, it affects ~1% of the world's adult population.[2,3] Further, RA disease burden experienced an unexpected steep rise from 2012 to 2017, reaching an age-standardized disability-adjusted life years (DALY) rate of 43.3 (95% Uncertainty Interval: 33.0 to 54.5) per 100,000 population.[4] Despite extensive studies on the exact etiology of RA, it remains unknown but is assumed to be multifactorial, involving both genetic and environmental factors.[5,6]

Currently, air pollution is nominated by the World Health Organization as one of the most significant health threats. It is well established that exposure to smoking[7] or silica exposure,[8] which causes an inflammatory and oxidative stress response, can increase the risk of RA. In addition, a previous study[9] also revealed that the lung may be the site of early related autoimmune injury in RA. Exposure to air pollution has been demonstrated to disrupt oxidation–reduction homeostasis in respiratory mucosal and triggers pro-inflammatory immune responses across multiple immune cells,[10,11] indicating that air pollution may be a potential risk factor for RA.[12] Several studies have focused on the relationship between air pollution and RA[12–17]; however, results were conflicting. The Nurses' Health Study (NHS) examined the association between distance to the nearest major road, a proxy marker of traffic pollution exposure, and the incidence of RA in 90,297 females, and the findings indicated that higher exposure to traffic pollution may be associated with RA risk.[13] Similarly, a retrospective cohort study in Taiwan, China,[14] found that newly diagnosed RA was significantly associated with NO_2 exposure and a study in South Korea[15] also reported that the incidence rate (IR) of RA was positively correlated with the concentration of particulate matter (PM) with an aerodynamic diameter of $\leq$2.5 μm ($PM_{2.5}$). In contrast, a case–control study based on the Swedish Epidemiological Investigation of Rheumatoid Arthritis (EIRA) has shown that after adjusting for the confounding factors of education and smoking status, the associations between particulate pollutants [PM with an aerodynamic diameter of $\leq$10 μm (PM_{10})], gaseous pollutants [nitrogen dioxide (NO_2) and sulfur dioxide (SO_2)], and incident RA were not statistically significant.[16] In addition, air pollutants, including $PM_{2.5}$, PM_{10}, NO_2, and SO_2, were associated with RA were also not observed, neither in the NHS study[17] nor in the British Colombian study.[18] The inconsistency among these findings implies that air pollution needs to be further investigated to evaluate whether it is a potential determinant of RA.

Genetic determinants provide initial insights into the presence of systemic autoimmunity and the identification of potentially at-

*These authors contributed equally to this work.

Address correspondence to Dong-Qing Ye, Department of Epidemiology and Biostatistics, School of Public Health, Anhui Medical University, Hefei, China. Email: ydq@ahmu.edu.cn. And, Chen Mao, Department of Epidemiology, School of Public Health, Southern Medical University, Guangzhou, China. Email: maochen9@smu.edu.cn

Supplemental Material is available online (https://doi.org/10.1289/EHP10710).

All authors declare they have no actual or potential competing financial interests.

Received 1 December 2021; Revised 3 February 2023; Accepted 6 February 2023; Published 13 March 2023.

risk individuals in the pre-RA stage.[19] In recent years, genome-wide association studies (GWAS) have identified several nonhuman leukocyte antigen (non-HLA) risk loci, and >100 genetic loci have been confirmed to be associated with the risk of RA.[20] Although each variant accounts for only a small-to-moderate proportion of the genetic risk of RA, using polygenic risk scores (PRSs) has proven to be an effective method for measuring the cumulative effects of multiple risk-related variants.[21,22] Moreover, quantifying the joint effects of genetic risk and epidemiological factors can greatly improve risk stratification or explore potential environment–gene interactions, thereby providing new insight for precise prediction and targeted interventions in RA. For instance, a previous study used a combination of 39 independent RA risk alleles to establish a PRS, which, combined with family history and epidemiological risk factors, was used to develop a well-performed risk prediction model for seropositive and seronegative RA.[23]

Previous studies have explored only the relationship between ambient air pollutants and RA risk and have largely ignored the modification effects of genetic susceptibility. Therefore, based on a general population cohort from the UK Biobank, the present study aimed to assess whether combined exposure to air pollution and genetic factors contributes to the incidence of RA.

Methods

Study Population

The UK Biobank resource includes ~0.5 million UK participants 39–73 years of age during the period of recruitment between 2006 and 2010. Participants attended one of the 22 assessment centers across England, Scotland, and Wales, where they completed touchscreen and nurse-led questionnaires, underwent physical measurements, and provided biological samples. More details of the study design have been described elsewhere.[24] The UK Biobank study was conducted under the approval of the North West Multi-center Research Ethical Committee (11/NW/0382). All participants provided written informed consent.

In the present study, the baseline time was defined as the time when participants first attended the assessment center, between 2006 and 2010. As genetic quality control (QC) measures, we excluded participants with sex mismatch, heterozygosity rate outliers, missing genotypes, excess relatives, and non-White race/ethnicity ($n=109{,}499$). The White participants included those whose self-reported ancestry was White British (based on UK Biobank Data-Field 1657 "self-reported ethnic group") and was also further confirmed as of Caucasian ancestry (based on UK Biobank Data-Field 22006 "genetic ethnic group"). Non-White participants were excluded under the consideration that the RA GWAS is mainly of European ancestry,[20] and the proportion of non-White participants in the UK Biobank is <5%[25]; thus, there may not be enough incident RA cases for the statistical analyses for the non-White population. Further, participants with prevalent RA ($n=2{,}936$) and those with incomplete information on residential air pollution ($n=32{,}001$) at baseline were also excluded. A total of 342,973 participants who had complete data for the concentration of five air pollutants at baseline and genotyping were included in the final analysis. A flowchart of the study participants selection is shown in Figure S1.

Ascertainment of Outcomes

At the baseline (2006–2010), we combined the self-reported and related therapeutic drugs, including steroids, synthetic disease-modifying anti-rheumatic drugs (DMARDs), and biologic DMARDS, the use of which represents outpatient visits, as well as the hospital inpatient records [using the *International Classification of Diseases, Manual of the International Statistical Classification of Diseases, Injuries and Causes of Death* (ICD-9,[26] codes 71400, 71401, 71403, 71404, 71405, 71406, and 71409) or the *International Statistical Classification of Diseases and Related Health Problems, Tenth Revision* (ICD-10,[27] codes M05 and M06) to identify prevalent cases of RA.] The self-reported RA and records of those who had used related therapeutic drugs for RA were collected through verbal interviews with well-trained nurses, the specific information about RA in verbal interviews was entered as free text and subsequently used as unique coded data. Hospital inpatient records are directly linked to the Health Episode Statistics in England and Wales and the Scottish Morbidity Records in Scotland, allowing for accurate identification of the first recorded date of each diagnosis. After removing cases of RA at baseline, incident RA events were identified during the subsequent following-up from the admission data using ICD-9 or ICD-10 codes. Detailed information on codes used to identify RA cases in this study can be found in Table S1.

Estimation of Air Pollutants and Air Pollution Score

Estimates of ambient air pollutants, including $PM_{2.5}$, $PM_{2.5-10}$, PM_{10}, NO_2, and NO_x, were collected by the UK Biobank between 2005–2007 and 2010. Data on ambient air pollution for the 2005–2007 period were derived from EU-wide air pollution maps (resolution 100×100 m), and the UK Biobank overlaid the coordinates of each subject's residential address onto these maps to obtain the corresponding air pollution concentrations of 100×100-m-grid cells.[28] Meanwhile, the 2010 annual average air pollution concentration was calculated by using a land use regression (LUR) model that was combined with the participants' residential addresses given at the baseline visit and the monitoring data from the European Study of Cohorts for Air Pollution Effects (ESCAPE) from 26 January 2010 to 18 January 2011. The LUR model was developed as part of the ESCAPE project, and the validation of the models has been described elsewhere (http://www.escapeproject.eu/).[29]

Annual concentration data for $PM_{2.5}$, $PM_{2.5-10}$, and NO_x were only available for 2010, whereas the concentration data for NO_2 was available for 2005–2007 and 2010. Finally, data for PM_{10} were available for 2007 and 2010. Pollutants for which only the annual concentration data in 2010 was available were directly defined as their respective exposure variable at baseline. Meanwhile, for pollutants for which several years of annual concentration data within the baseline time range were available, we took the average value of multiple annual concentrations as their exposure variable. For example, for PM_{10}, we considered the average value of the annual concentrations in 2007 and 2010 as its exposure level. Furthermore, the air pollution score was constructed by weighted summing concentrations of the five air pollutants, and the weighted method was based on the multivariable-adjusted risk estimates (β coefficients) of RA. The use of the air pollution score has been well accepted in previous UK Biobank studies to assess the association of combined exposure to multiple air pollutants with the risk of chronic diseases, such as type 2 diabetes[30] and heart failure.[31] The air pollution score was obtained using the following formula:

$$\text{Air pollution score} = [\beta_{(PM_{2.5})}\times PM_{2.5} + \beta_{(PM_{2.5-10})}\times PM_{2.5-10} + \beta_{(PM_{10})}\times PM_{10} + \beta_{(NO_2)}\times NO_2 + \beta_{(NO_x)}\times NO_x]\times\left(5/\sum\beta\right)$$

Genotype Data, QC, and PRS

Genotype calling, QC, phasing, and imputation were performed centrally and have been previously described in detail.[25] In brief, ~50,000 and 450,000 participants were separately genotyped on

two closely related purpose-designed arrays (UK BiLEVE Axiom and UK Biobank Axiom), and 805,426 markers were identified in the finally released genotype data. In addition, the data set was staged using computationally efficient methods and combined with the Haplotype Reference Consortium and UK10K Haplotype Resources to impute a total of 96 million genotypes. Based on the above genotyping imputed data, we carried out downstream QC measures; specifically, we removed single nucleotide polymorphisms (SNPs) with imputation information (INFO) score of <0.3, with minor allele frequency of <0.5%, or those that failed Hardy-Weinberg tests with a $p < 1 \times 10^{-6}$ using QCTOOL (version 2; https://www.well.ox.ac.uk/~gav/qctool/index.html).

A PRS, which captures an individual's load of common genetic variants associated with RA risk, was constructed. The score was based on RA summary statistics from a meta-analysis of GWAS data from individuals of European ancestry (http://plaza.umin.ac.jp/~yokada/datasource/software.htm).[32] We applied the clumping and threshold (C+T) method for calculating the PRS of RA, which involved computing PRSs based on a subset of partially independent (clumped) SNPs exceeding a specific GWAS association p-value threshold.[21] To select the SNPs that would be included in the calculation of the genetic risk score, we first filtered the GWAS statistics results to exclude variants with a $p \geq 5 \times 10^{-8}$ (genome-wide significant). Linkage-disequilibrium clumping was performed to identify independently associated variants ($r^2 < 0.01$). If the selected variants were not present in the UK Biobank genotyping data, proxy variants were sought ($r^2 > 0.8$). Moreover, SNPs were also filtered based on the INFO score (INFO score >0.95) to ensure imputed genotyping quality.

In this study, the additive genetic model,[33] including 154 SNPs (see Excel Table S1 for details), was used for PRS calculation, and the final PRS was standardized. The calculation formula is as follows:

$$PRS_j = \left[\sum_{ij} S_i \times G_{ij} - \text{mean}(PRS)\right] / \text{SD}(PRS),$$

where S is the summary statistic for the effective allele; G is the number of the effective allele (0, 1, 2) observed; i is the ith SNP; j is the jth individual; and SD is the standard deviation. The above procedure was performed in PRSice-2[34] with the RA GWAS summary statistics and the genotyping imputed data from the UK Biobank. According to the PRS distribution, individuals were categorized into low (tertile 1)-, intermediate (tertile 2)-, and high (tertile 3)-grade RA genetic risks.

Covariate Measurements

Covariates in this study included age (years, continuous), sex (female, male), assessment center (22 centers), average total household income before tax (<£18,000, £18,000–£29,999, £30,000–£51,999, £52,000–£100,000, >£100,000), educational level [college or university degree, others (including advanced levels (A levels)/(advanced subsidiary levels (AS levels) or equivalent, ordinary levels (O levels)/general certificate of secondary education (GCSEs) or equivalent, CSEs or equivalent, national vocational qualifications (NVQ) or higher national diplomas (HND) or higher national certificates (HNC) or equivalent, and other professional qualifications, such as nursing and teaching)], smoking status (current, previous, never), alcohol consumption (standard-drinks per day, continuous), sedentary activity time (hours per day), physical activity duration (minutes per day), body mass index (BMI, in kilograms per meter squared), and healthy diet score (0–5 points). Some general characteristics (including date of birth, sex, and the name of the recruitment center) of participants were known before arrival at the assessment center, and all were obtained from local NHS Primary Care Trust registries. Sociodemographics (average total household income before tax, educational level) and lifestyle and behaviors data (smoking status, alcohol consumption, sedentary activity time, and physical activity duration) were collected from touchscreen questionnaires at the baseline assessment center visit. The participants' height and weight were measured by trained nurses, and BMI was determined by dividing weight (in kilograms) by height (in meters) squared. Sedentary activity time was obtained from the sum of hours per day spent driving, watching TV, and using a computer. Physical activity duration was defined as the sum of minutes per day spent walking and engaging in moderate and vigorous activity. The healthy diet score was based on the American Heart Association Guidelines[35] and included five dietary components: vegetables, fruit, fish, processed meat, and unprocessed red meat. The total diet score ranged from 0 to 5 points; each time a dietary component intake goal was achieved, it was given 1 point, with a higher score representing a healthier diet. Alcohol consumption was calculated by conversion of alcohol intake collected from a touchscreen questionnaire to standard-drink (8 g pure alcohol) of alcohol per day.[36] More information is shown in Table S2.

Statistical Analysis

The follow-up period was defined as from the date of initial recruitment to the onset of RA or competitive events (death), the date of loss to follow-up, or the date of the current end of follow-up (31 March 2017 for England, 31 October 2016 for Scotland, and 31 January 2017 for Wales). Participants who were lost to follow-up or died before RA occurred were censored at the time of the respective event.

Hazard ratios (HRs) and 95% confidence intervals (CIs) for incident RA associated with the single air pollutant, the air pollution score, or PRS were estimated with Cox proportional hazard models. We evaluated the assumption of the proportionality of hazards by examining the association between standardized Schoenfeld residuals and time.[37] Missing values of covariates were imputed by a multiple chain equation,[38] and the missing value patterns were assumed to be randomly missing. The missing data in the continuous variable were imputed using predictive mean matching, whereas data in the categorical variable were imputed using logistic regression factor (2 levels) and multinomial or ordered logit models (>2 levels).

All the models for the association between a single pollutant and incident RA were adjusted for the covariates mentioned above. Then, the three nested Cox models, which included a basic model adjusted for age and sex, a multivariable adjustment model further adjusted for all listed confounding factors, and a fully adjusted model further adjusted for PRS, were used to estimate the impact of air pollution score on the risk of RA. For analysis models including PRS, we further adjusted for the genotyping batches (11 batches in the UK BiLEVE Axiom array, 95 batches in the Biobank Axiom array) and the first 10 genetic principal components (PC1–PC10). To evaluate the joint associations of PRS and air pollution score with RA risk, we classified participants into 12 groups according to genetic risk and quartiles of the air pollution score. The HRs of incident RA in different groups were estimated compared with those with low genetic risk and the lowest quartile of air pollution score. We performed the Cochran–Armitage test for trends in binomial RA status across the levels of the variable of interest. To quantify the interactions on additive and multiplicative scales, we added a product term that combined high genetic risk and fourth quartile air pollution score in the model. The HR for the product term and the relative excess risk due to interaction (RERI) were used as the measures of interaction on the multiplicative

and additive scales, respectively. The exposure–response relationship of the air pollution score with RA risk was assessed using restricted cubic spline analysis with 3 knots. Spearman's correlation coefficients were also calculated to assess correlations among air pollutants.

A 10-fold cross-validation analysis was performed to further validate the results.[39] The overall data were randomly divided into 10 equal parts, with 9 of them taken as the training data set and the remaining 1 as the testing data set. In the training data set, the single air pollutant was refitted in the Cox model to obtain a new air pollution score–weighting coefficient. Accordingly, a new air pollutant score was constructed and its association strength with incident RA was evaluated in the testing data set. All process steps were repeated 10 times until each of the 10 parts was used once as the testing data set. A fixed-effects meta-analysis was performed to calculate the pooled HR. To further confirm the robustness of the weighted air pollution score, we also performed common mixture pollutants exposure estimation methods, quantile-based g-computation,[40] and Bayesian kernel machine regression (BKMR),[41] to recalculate the combined effect of the mixture air pollution.

Moreover, we conducted subgroup analyses in various dichotomous subgroups according to age (≥65 and <65 y), sex (female or male), education levels (with and without university degrees), and smoking status (previous/current and never). Apart from adjusting for the variables aforementioned in the main analysis, menopausal status (yes or no) and hormone replacement therapy use (yes or no) were additionally adjusted among females in the subgroup analyses. Furthermore, the RA cases were stratified according to the rheumatoid factor (RF) level and divided into positive and negative groups in line with a cutoff value of 20 IU/mL,[42,43] and the association of air pollution exposure with different RF-status RA risks was further checked. The statistical methods in subgroup analysis used were consistent with the main analysis; however, when conducting interaction analysis in each subgroup, the product term included in the model were ≥65 years of age, female, no university degree, previous/current smoking, RF-positive, and fourth quartile air pollution score, respectively.

A series of sensitivity analyses were also conducted to demonstrate the robustness and reliability of the results. First, we established a new air pollution score that included only $PM_{2.5}$, NO_2, and NO_x to further explore the association of air pollution score with the risk of incident RA. Second, $PM_{2.5}$ absorbance, as a measurement of the blackness of $PM_{2.5}$ filters and a proxy for elemental carbon, was assessed by the LUR model, as previously described.[29,44] We included the $PM_{2.5}$ absorbance (2010 available) in the air pollution score to serve as a further supplement to the PM exposure information. Third, to evaluate concerns for potential reverse causation, we restricted incident RA cases to >2 y from the baseline time. Fourth, we additionally adjusted the latitude of the participants' residence because latitude is closely related to exposure to ultraviolet radiation, which in turn affects immune regulation or vitamin D synthesis and thus, potentially, the occurrence of RA.[45,46] Fifth, furthermore, to avoid inaccurate assignments of air pollution estimates caused by changes in residence, we included only participants who had lived at their current address for at least 5 y in the analysis. Finally, the diagnosis of RA may be delayed owing to the use of nonsteroidal anti-inflammatory drugs (NSAIDs) to alleviate symptoms, such as arthritis, so we excluded individuals using any dose of NSAIDs (Table S1) within 3 months before the baseline to avoid latent prevalent RA cases masked by NSAID use being included in our analysis.

All analyses were performed using R software (version 4.0.3; R Development Core Team). All p-values for the tests were two sided, and $p < 0.05$ were considered statistically significant. Cox models were constructed using the "survival" package. Exposure–response relationship analyses were performed using the "rms" package. Interaction analyses were realized by using the "interactionR" package. "qgcomp" and "bkmr" packages were used to conduct quantile-based g-computation and BKMR. The location coordinates of participants' residential addresses measured by the Ordnance Survey in Great Britain (OSGB) grid reference were transformed into latitude/longitude through a JavaScript library[47] with the function of coordinate system conversions. The administrative boundary data of the UK were sourced from UK Government Open Data portal.[48] All maps were drawn in R with the "rgdal" and "ggplot2" packages.

Results

The baseline characteristics of the study participants are presented in Table 1. Among 342,973 participants, 2,034 incident cases of RA were recorded during 2,760,119 person-years of follow-up (median follow-up time = 8.1 y). Compared with participants without RA, the individuals with RA had lower income levels (<£18,000: 36.6% vs. 22.3%), and they were more likely to be previous or current smokers (54.7% vs. 45.3%), have a higher BMI (28.9 kg/m^2 vs. 27.4 kg/m^2), and have a more sedentary lifestyle (5.2 h/d vs. 4.9 h/d sedentary time). A comparison of baseline characteristics between the current study population and the full UK Biobank cohort was also reported in Table S3. The median [interquartile range (IQR)] estimates of $PM_{2.5}$, $PM_{2.5-10}$, PM_{10}, NO_2, and NO_x were respectively 9.97 (1.32), 6.10 (0.79), 38.09 (4.38), 27.80 (9.87), and 42.30 (16.15) $\mu g/m^3$ among participants with incident RA at baseline. The corresponding median (IQR) estimates were 9.88 (1.26), 6.08 (0.76), 38.01 (4.44), 27.30 (10.22), and 41.27 (16.10) $\mu g/m^3$ for those without incident RA. The Spearman correlation coefficients among the five air pollutants are shown in Table S4. The dispersed distribution of air pollution levels and the incident RA cases in the areas where participants lived in 2010 are shown in Figure 1, and the distribution patterns were in line with patterns available in the public UK Air Information Resources.[49]

The associations between individual air pollutants and RA are shown in Table 2. We observed that $PM_{2.5}$, NO_2, and NO_x were positively associated with the risk of RA (p_{Trend} for $PM_{2.5} = 0.000043$, p_{Trend} for $NO_2 = 0.00042$, and p_{Trend} for $NO_x = 0.000011$, respectively) after adjusting for age, sex, UK Biobank assessment centers, household income, education, smoking status, BMI, alcohol consumption, sedentary time, physical activity duration, and healthy diet score. The HRs (95% CI) of RA per IQR increase in $PM_{2.5}$ (IQR: 1.26 $\mu g/m^3$), PM_{10} (IQR: 4.44 $\mu g/m^3$), $PM_{2.5-10}$ (IQR: 1.26 $\mu g/m^3$), NO_2 (IQR: 10.22 $\mu g/m^3$), and NO_x (IQR: 16.10 $\mu g/m^3$) were 1.07 (1.01, 1.13), 1.01 (0.96, 1.07), 1.00 (0.96, 1.04), 1.03 (0.98, 1.09), and 1.07 (1.02, 1.12), respectively. The exposure–response relationship of the association of each air pollutant with incident RA was also checked, and $PM_{2.5}$ showed a relatively strong effect (Figure S2 and Excel Table S2).

The mean air pollution score ± standard deviation (SD) was 70.63 ± 11.12, ranging from 48.54 to 202.00, with a higher score indicating a higher combined exposure to air pollutants. The weights of each air pollutant included in the calculation of the air pollution score are shown in Table S5, and the distribution of air pollution scores among participants is shown in Figure S3 and Excel Table S3. As shown in Table 3, we found a positive exposure–response relationship between the air pollution score and RA risk ($p_{Trend} = 0.000053$). After adjusting for age and sex, the risk of incident RA in the highest quartile of the air pollution score was 34% (95% CI: 18%, 51%) higher than in the lowest quartile group. Moreover, the association of air pollution score with incident RA remained significant after adjusting for the UK

Table 1. Baseline characteristics of 342,973 participants in the UK Biobank study of the association of air pollution and genetic risk with rheumatoid arthritis (RA) incidence from 2006 to 2017.

	Incident RA		
Characteristics[a]	Yes ($n = 2,034$)	No ($n = 340,939$)	Total population ($N = 342,973$)
Age [y (mean ± SD)]	59.94 ± 7.02	56.97 ± 7.93	56.99 ± 7.93
Follow-up time [median (IQR)]	5.2 (3.6)	8.1 (1.2)	8.1 (1.2)
Sex [*n* (%)]			
Female	1,368 (67.26)	181,850 (53.33)	183,218 (53.42)
Male	666 (32.74)	159,089 (46.66)	159,755 (46.58)
Household income [*n* (%)]			
< £18,000	745 (36.63)	76,176 (22.34)	76,921 (22.43)
£18,000–29,999	573 (28.17)	88,752 (26.03)	88,325 (26.04)
£30,000–51,999	430 (21.14)	89,897 (26.37)	90,327 (26.34)
£52,000–100,000	240 (11.80)	68,633 (20.13)	68,873 (20.08)
> £100,000	46 (2.26)	17,481 (5.13)	17,527 (5.11)
Education level [*n* (%)]			
College or university degree	542 (26.65)	121,615 (35.67)	122,157 (35.62)
Other	1,492 (73.35)	219,324 (64.33)	220,816 (64.38)
Smoking status [*n* (%)]			
Never smoking	921 (45.28)	186,337 (54.65)	187,258 (54.60)
Previous smoking	832 (40.90)	120,838 (35.44)	121,670 (35.48)
Current smoking	281 (13.81)	33,764 (9.91)	34,045 (9.93)
Alcohol consumption [standard-drink/d (mean ± SD)[b]]	1.63 ± 2.27	2.03 ± 2.45	2.03 ± 2.44
Healthy diet score [*n* (%)]			
0–1	461 (22.66)	77,519 (22.74)	77,980 (22.74)
2–3	1,064 (52.31)	184,495 (54.12)	185,559 (54.10)
4–5	509 (25.02)	78,925 (23.14)	79,434 (23.16)
Sedentary time [h/d (mean ± SD)]	5.18 ± 2.45	4.87 ± 2.38	4.87 ± 2.38
Physical activity [min/d (mean ± SD)]	126.93 ± 107.24	128.81 ± 102.79	128.80 ± 102.82
Body mass index [kg/m^2 (mean ± SD)]	28.89 ± 5.56	27.41 ± 4.74	27.42 ± 4.75
Rheumatoid factor (RF) status [*n* (%)]			
RF-positive	417 (20.50)	11,803 (3.46)	12,220 (3.56)
RF-negative	1,543 (75.86)	312,557 (91.68)	314,100 (91.58)
Latitude of residence [degree (mean ± SD)]	53.06 ± 1.14	52.87 ± 1.19	52.86 ± 1.19
$PM_{2.5}$ [$\mu g/m^3$, median (IQR)]	9.97 (1.32)	9.88 (1.26)	9.88 (1.26)
$PM_{2.5–10}$ [$\mu g/m^3$, median (IQR)]	6.10 (0.79)	6.08 (0.76)	6.08 (0.76)
PM_{10} [$\mu g/m^3$, median (IQR)]	38.09 (4.38)	38.01 (4.44)	38.01 (4.44)
NO_2 [$\mu g/m^3$, median (IQR)]	27.80 (9.87)	27.30 (10.22)	27.30 (10.22)
NO_x [$\mu g/m^3$, median (IQR)]	42.30 (16.15)	41.27 (16.10)	41.27 (16.10)
Air pollution score (mean ± SD)	71.68 ± 11.35	70.63 ± 11.13	70.63 ± 11.12

Note: IQR, interquartile range; NO_2, nitrogen dioxide; NO_x, nitrogen oxides; $PM_{2.5}$, particulate matter with aerodynamic diameter ≤2.5 μm; $PM_{2.5–10}$, particulate matter with an aerodynamic diameter between 2.5 and 10 μm; PM_{10}, particulate matter with an aerodynamic diameter ≤10 μm.

[a]Missing values for each characteristic: household income ($n = 48,144$), education ($n = 62,218$), smoking status ($n = 1,182$), alcohol consumption ($n = 2,848$), healthy diet score ($n = 56,104$), sedentary time ($n = 8,334$), physical activity ($n = 65,263$), body mass index ($n = 1,079$), and rheumatoid factor ($n = 16,653$).

[b]A standard-drink of alcohol consumption = 8 g of pure alcohol intake.

Biobank assessment center, household income, education, smoking status, BMI, alcohol consumption, sedentary time, physical activity duration, and healthy diet score, with the HR (95% CI) for the fourth quartile group being 1.14 (1.00, 1.29). Furthermore, the above results remained stable when we further included the PRS of RA, the first 10 genetic principal components, and genotyping batches into the models. The results from cross-validation analysis further confirmed the robustness of the findings, with the value of $HR_{Fixed\text{-}effect}$ (95% CI) being 1.05 (1.01, 1.10) per SD increment in air pollution score (Figure S4). We compared the impacts of the air pollutants mixture using three methods, including BKMR, the quantile g-computation model, and our air pollution score. In the air pollution score we constructed, $PM_{2.5}$ and NO_x were the highest contributors ($\beta_{PM2.5} = 0.05$ and $\beta_{NOx} = 0.004$), which was relatively consistent with the BKMR [conditional posterior inclusion probabilities ($PIP_{PM2.5} = 0.77$ and $PIP_{NOx} = 0.94$); Table S6]. In quantile g-computation models, the magnitude of the overall mixture effects (Ψ); that is, the HR (95% CI) of incident RA was 1.08 (1.02, 1.15) per quartile increase in the concentration of all air pollutants, showing an association pattern similar to our air pollution score (Table S7).

The exposure–response relationships of the air pollution score with incident RA according to stratification of age, sex, smoking status, education level, and RF status are shown in Figures S5–S9. Age is an important modifier in the impact of air pollution on the risk of RA ($p_{Multiplicative\ interaction} = 0.044$), and the HR (95% CI) per SD increased air pollution score of those ≥65 years of age was 1.09 (1.01, 1.18); Table S8, Figure S5, and Excel Table S4]. A strong interaction between the air pollution and sex on the RA risk was also observed ($p_{Multiplicative\ interaction} = 0.005$ and $p_{Additive\ interaction} = 0.0009$); the effect of air pollution on the RA risk in females was more apparent than in males, and the HR (95% CI) per SD increased air pollution score in females was 1.10 (1.05, 1.16); Table S8, Figure S6, and Excel Table S5]. Compared with nonsmokers, the exposure–response curve of smokers' air pollution scores and RA incidence risk seemed to rise more rapidly (Figure S7 and Excel Table S6); however, there was no significant interaction between the air pollution and smoking status on the RA risk ($p_{Multiplicative\ interaction} = 0.99$ and $p_{Additive\ interaction} = 0.98$; Table S8). Although there were significant differences in the strength of association between air pollution and RA risk in different education levels (Figure S8 and Excel Table S7), only marginal significant additive interaction was found [RERI = 0.23 (95% CI: 0.002, 0.46); Table S8]. In addition, the risk of seronegative-RF RA increased significantly with an increase in air pollution score, suggesting there may be a negative interaction between air pollution and RF-positive RA ($p_{Multiplicative\ interaction} = 0.0071$; Table S8, Figure S9, and Excel Table S8).

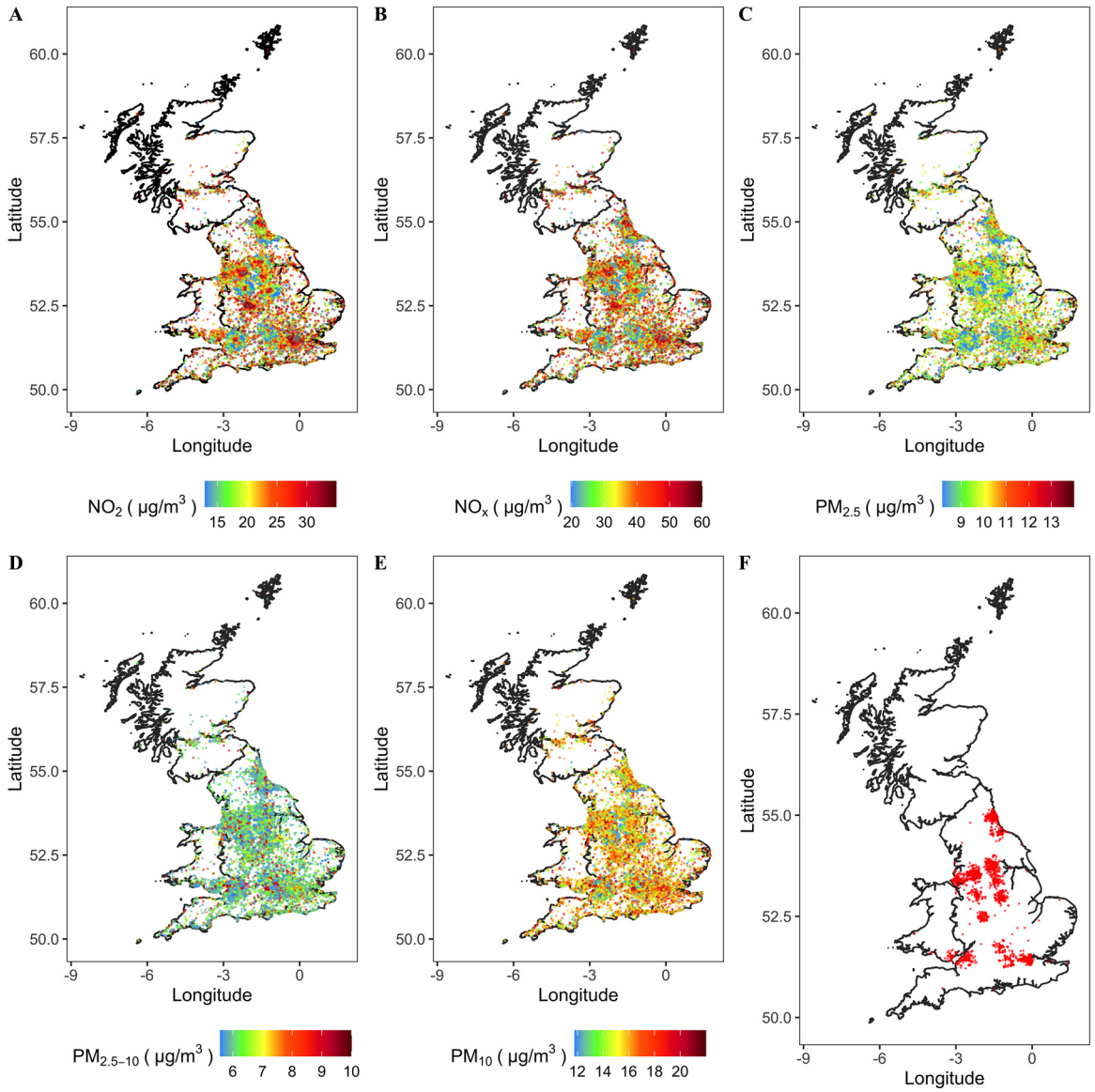

Figure 1. Map of air pollution (NO_2, NO_x, $PM_{2.5}$, $PM_{2.5–10}$, and PM_{10}) of areas where participants lived in 2010 and map of incident RA scatter distribution. The administrative boundary data of the UK are sourced from UK Government Open Data portal (https://data.gov.uk/dataset/3fd8d2d2-b591-42ff-b333-c53a6a513e96/countries-december-2017-full-clipped-boundaries-in-great-britain). These data are UK government–released open data. (A) NO_2, (B) NO_x, (C) $PM_{2.5}$, (D) $PM_{2.5–10}$, (E) PM_{10}, and (F) incident RA. Note: NO_2, nitrogen dioxide; NO_x, nitrogen oxides; $PM_{2.5}$, particulate matter with aerodynamic diameter ≤2.5 μm; $PM_{2.5–10}$, particulate matter with aerodynamic diameter 2.5–10 μm; PM_{10}, particulate matter with aerodynamic diameter ≤10 μm; RA, rheumatoid arthritis.

As shown in Table 4, a significant positive association was observed between the PRS of RA and the risk of incident RA. After adjusting for sex, age, assessment center, first 10 genetic principal components, and genotyping batches, the HR (95% CI) of incident RA per increment in SD in the PRS of RA was 1.22 (1.17, 1.27). Moreover, the HR (95% CI) of incident RA in the high genetic risk group was 1.48 (1.33, 1.65) when compared with the low genetic risk group. All the findings remained stable in the multivariate-adjusted models.

The joint association of the air pollution score and the PRS with the risk of RA incidence was further assessed, and a significant exposure–response relationship between air pollutants and incident RA was found in the low ($p_{Trend} = 0.00064$) and intermediate genetic risk groups ($p_{Trend} = 0.0096$) (Figure 2). Furthermore, the results of the combined effect of air pollution score and PRS on the RA risk showed that the risk of RA incidence in the highest genetic risk and air pollution score group was almost twice that of the lowest genetic risk and air pollution score group [incidence rate (IR) per 100,000 person-years = 98.46 vs. 51.19, and HR = 1.73 (95% CI: 1.39, 2.17) vs. 1 (reference)], although no statistically significant interaction between the air pollution and genetic risk on incident RA was found ($p_{Multiplicative\ interaction} = 0.057$ and $p_{Additive\ interaction} = 0.54$; Figure 2).

In addition, several sensitivity analyses were performed to confirm our findings. We first recalculated air pollution scores that excluded $PM_{2.5–10}$ and PM_{10} (the remaining weights of $PM_{2.5}$, NO_2, and NO_x were the same as those in the main analyses) and examined the relationship between RA incidence and the recalculated air pollution scores, and the results remained unchanged (Table S9). In addition, we found no significant impact of $PM_{2.5}$ absorbance on RA risk in the single-pollutant model ($p_{Trend} = 0.10$), and the magnitude of the association was slightly attenuated in the multivariable-adjusted model analysis after

Table 2. Association between single air pollutant and incident rheumatoid arthritis (RA) among UK Biobank participants ($N = 342{,}973$; $n = 2{,}034$ incident RA cases).

Air pollutants[a]	Case/*n*	HR (95% CI) of incident RA[b]	p_{Trend}
$PM_{2.5}$			
Per IQR increment	—	1.07 (1.01, 1.13)	
Q1	467/86,448	Ref	0.000043
Q2	480/86,172	0.97 (0.86, 1.11)	
Q3	504/84,731	1.00 (0.88, 1.14)	
Q4	583/85,622	1.12 (1.00, 1.26)	
PM_{10}			
Per IQR increment	—	1.01 (0.96, 1.07)	
Q1	486/86,006	Ref	0.13
Q2	511/85,457	0.99 (0.88, 1.13)	
Q3	516/85,870	0.98 (0.87, 1.11)	
Q4	521/85,640	1.05 (0.92, 1.19)	
$PM_{2.5–10}$			
Per IQR increment	—	1.00 (0.96, 1.04)	
Q1	498/86,570	Ref	0.12
Q2	504/85,471	1.00 (0.88, 1.13)	
Q3	499/85,500	1.00 (0.88, 1.13)	
Q4	533/85,432	1.05 (0.93, 1.19)	
NO_2			
Per IQR increment	—	1.03 (0.98, 1.09)	
Q1	442/85,746	Ref	0.00042
Q2	512/85,758	1.06 (0.93, 1.20)	
Q3	536/85,722	1.08 (0.95, 1.23)	
Q4	544/85,747	1.14 (1.00, 1.30)	
NO_x			
Per IQR increment	—	1.07 (1.02, 1.12)	
Q1	442/85,750	Ref	0.000011
Q2	498/85,758	1.05 (0.92, 1.19)	
Q3	518/85,740	1.05 (0.93, 1.20)	
Q4	576/85,725	1.17 (1.03, 1.33)	

Note: —, not applicable; CI, confidence interval; HR, hazard ratio; IQR, interquartile range; NO_2, nitrogen dioxide; NO_x, nitrogen oxides; $PM_{2.5}$, particulate matter with aerodynamic diameter ≤2.5 μm; $PM_{2.5–10}$, particulate matter with an aerodynamic diameter between 2.5 and 10 mm; PM_{10}, particulate matter with an aerodynamic diameter ≤10 μm; Q, quartile.

[a]$PM_{2.5}$ ranges: quartile 1: (8.17–9.23) μg/m^3, quartile 2: (9.24–9.88) μg/m^3, quartile 3: (9.89–10.49) μg/m^3, and quartile 4: (10.50–21.31) μg/m^3 and IQR is 1.26 μg/m^3; PM_{10} ranges: quartile 1: (25.73–35.89) μg/m^3, quartile 2: (35.90–38.01) μg/m^3, quartile 3: (38.02–40.33) μg/m^3, and quartile 4: (40.34–60.16) μg/m^3 and IQR is 4.44 μg/m^3; $PM_{2.5–10}$ ranges: quartile 1: (5.57–5.83) μg/m^3, quartile 2: (5.84–6.08) μg/m^3, quartile 3: (6.09–6.59) μg/m^3, and quartile 4: (6.60–12.82) μg/m^3 and IQR is 1.26 μg/m^3; NO_2 ranges: quartile 1: (8.86–22.40) μg/m^3, quartile 2: (22.41–27.30) μg/m^3, quartile 3: (27.31–32.62) μg/m^3, and quartile 4: (32.63–125.12) μg/m^3 and IQR is 10.22 μg/m^3; NO_x ranges: quartile 1: (19.74–33.38) μg/m^3, quartile 2: (33.39–41.27) μg/m^3, quartile 3: (41.28–49.48) μg/m^3, and quartile 4: (49.49–265.94) μg/m^3 and IQR is 16.10 μg/m^3.

[b]Adjusted for age, sex, UK Biobank assessment center, household income, education level, smoking status, body mass index, alcohol consumption, sedentary time, physical activity duration, and healthy diet score.

incorporating it into the air pollution score (Table S10). Participants with a follow-up time of <2 y were removed from the analysis, and this did not appreciably change the results (Table S11). Furthermore, the latitude of participants' residence was further adjusted (Table S12) and participants in the analysis were limited to those who had lived at their current address for at least 5 y (Table S13), and the findings were found to be robust. Finally, after excluding participants who had used NSAIDs, the results were still comparable (Table S14).

Discussion

To the best of our knowledge, this is the first prospective cohort study to investigate the association of combined exposure to multiple air pollutants with the risk of incident RA while considering the modification effect of genetic risk. By weighting the regression coefficient of each air pollutant ($PM_{2.5}$, $PM_{2.5–10}$, PM_{10}, NO_2, and NO_x), we constructed an air pollution score to represent comprehensive air pollution and assessed the association with the risk of RA. The results showed a positive association with RA incidence. Furthermore, we found that the IR of RA almost increased monotonically with increasing air pollution scores across different genetic risk strata, particularly in the low and intermediate genetic risk groups. The nonstatistically significant exposure–response relationship in the high genetic risk group reflected that the health effect of air pollution may play a minor role compared with a high genetic predisposition. However, the highest IR and HR of RA risk in the population that had the highest air pollution and genetic risk was still more concerning.

To date, the effects of single air pollutants, such as NO_2 and $PM_{2.5}$, on health effects have been widely demonstrated.[50] However, given that humans are exposed to a mixture of air pollutants, seeking to use a mixture of pollutant exposure estimation methods is important.[51,52] In this study, we attempted to construct an air pollution score using the weighted regression coefficient method to characterize the mixed exposure to multiple air pollutants and observed a modest positive association between the air pollution score and the risk of RA. Previous UK Biobank studies[30,31] used the same algorithm to deal with the additive linear effects of different air pollutants and developed an air pollution score that proved that joint exposure to air pollutants is significantly associated with the risk of type 2 diabetes and heart failure. Similar integrated scores have been applied not only in the field of environmental health[53,54] but also in other epidemiological studies.[55,56] Moreover, we additionally checked whether the magnitude of our current air pollution score health effect for RA was comparable to that using other common mixture pollutants exposure estimate methods, such as quantile-based g-computation[40] and BKMR,[41] in our further validation (Tables S6 and S7). In the quantile g-computation model, the overall mixture effects (Ψ) were close to our air pollution score. The results from the BKMR models also demonstrated a similar positive association between air pollution and RA risk. In general, our air pollution

Table 3. Association between air pollution score and incident rheumatoid arthritis (RA) among UK Biobank participants ($N = 342{,}973$; $n = 2{,}034$ incident RA cases).

		HR (95% CI) of incident RA			
Air pollution score[a]	Case/*n*	Model 1[b]	Model 2[c]	Model 3[d]	p_{Trend}
Per standard deviation increment	—	1.12 (1.07, 1.16)	1.06 (1.01, 1.10)	1.06 (1.01, 1.10)	
Q1	454/85,743	Ref	Ref	Ref	0.000053
Q2	490/85,743	1.09 (0.96, 1.24)	1.00 (0.89, 1.14)	1.00 (0.88, 1.14)	
Q3	514/85,743	1.17 (1.03, 1.33)	1.01 (0.89, 1.15)	1.01 (0.89, 1.15)	
Q4	576/85,744	1.34 (1.18, 1.51)	1.14 (1.00, 1.29)	1.14 (1.00, 1.29)	

Note: —, not applicable; CI, confidence interval; HR, hazard ratio; Q, quartile; Ref, reference; SD, standard deviation.

[a]Air pollution score ranges: quartile 1: (48.54–63.35); quartile 2: (63.36–69.94); quartile 3: (69.95–76.43); and quartile 4: (76.44–202.00). Mean ± SD of the air pollution score is 70.63 ± 11.12.

[b]Adjusted for age and sex.

[c]Adjusted for age, sex, UK Biobank assessment center, household income, education level, smoking status, body mass index, alcohol consumption, sedentary time, physical activity duration, and healthy diet score.

[d]Adjusted for age, sex, UK Biobank assessment center, household income, education level, smoking status, body mass index, alcohol consumption, sedentary time, physical activity duration, healthy diet score, polygenic risk score, first 10 genetic principal components, and genotyping batch.

Table 4. Association between genetic risk and incident rheumatoid arthritis (RA) among UK Biobank participants ($N = 342{,}973$; $n = 2{,}034$ incident RA cases).

Polygenic risk score[a]	Case/*n*	HR (95% CI) of incident RA		p_{Trend}
		Model 1[b]	Model 2[c]	
Per standard deviation increment	—	1.22 (1.17, 1.27)	1.22 (1.17, 1.27)	
Low genetic risk	554/113,181	Ref	Ref	2.63×10^{-13}
Intermediate genetic risk	640/113,181	1.16 (1.03, 1.29)	1.16 (1.04, 1.30)	
High genetic risk	840/116,611	1.48 (1.33, 1.65)	1.48 (1.33, 1.65)	

Note: —, not applicable; CI, confidence interval; HR, hazard risk; PRS, polygenic risk score; Ref, reference.
[a]PRS ranges: low genetic risk (tertile 1): (−3.524 to −0.497); intermediate genetic risk (tertile 2): (−0.498 to 0.297); high genetic risk (tertile 3): (0.298 to 5.066). Mean ± standard deviation of PRS is 0 ± 1.
[b]Adjusted for age, sex, first 10 genetic principal components, and genotyping batch.
[c]Adjusted for age, sex, UK Biobank assessment center, household income, education level, smoking status, body mass index, alcohol consumption, sedentary time, physical activity duration, healthy diet score, PRS, first 10 genetic principal components, and genotyping batch.

scores were relatively accurate and reliable, and cross-validation methods avoided the problem of overfitting. Moreover, as a continuous variable, the air pollution score can contribute to the determination of the possible risk threshold for disease prevention and also facilitate complex interaction analysis with other risk factors of interest. In the multivariable-adjusted model, only the air pollution score in the highest quartile was significantly associated with incident RA. However, it cannot be ignored that the air pollution scores in this study did not cover all air pollutants and may have resulted in an underestimated relationship between air pollution and the risk of RA.

Environmental agents are thought to interact with genetic factors and jointly trigger the immunologic processes before clinical RA.[57] A classic example supporting the environment–gene interaction is that between human leucocyte antigen-shared epitope gene (*HLA-SE*) alleles and smoking, which strongly increases the risk of seropositive RA.[7,58] Previous studies[13–18] have often ignored the modification effects of genetic susceptibility and have reported conflicting evidence on the relationship between air pollution and RA. In this study, to explore the potential interaction between air pollution and genetic risk, we constructed a PRS that represents the overall genetic risk of RA. We found that the intensity of the association between RA risk and air pollution level was higher with the increase of PRS even though the interaction between the high air pollution exposure and high genetic risk was not significant. However, the RA risk–associated genomic loci identified from population-based genetic association studies collectively account for only ∼15% of the phenotypic variance of RA[59,60] and may partially explain the nonstatistically significant interaction.

The pathogenic mechanism of air pollution in RA development has been comprehensively explored, with several hypotheses being proposed about potentially involved factors, including T cell imbalance, production of pro-inflammatory cytokines, local pulmonary inflammation, oxidative stress, and methylation changes.[61–65] The association of air pollution with RA was not limited to the risk of incidence, the two case-crossover studies[66,67] further revealed that exposure to high levels of air pollutants will also affect disease activity and drug response for RA. In addition, age, sex, socioeconomic status, and lifestyle are all potential modifiers of the magnitude of the effects of air pollution on RA.[68–73] Therefore, to confirm the robustness of the results, we further performed stratified analyses of

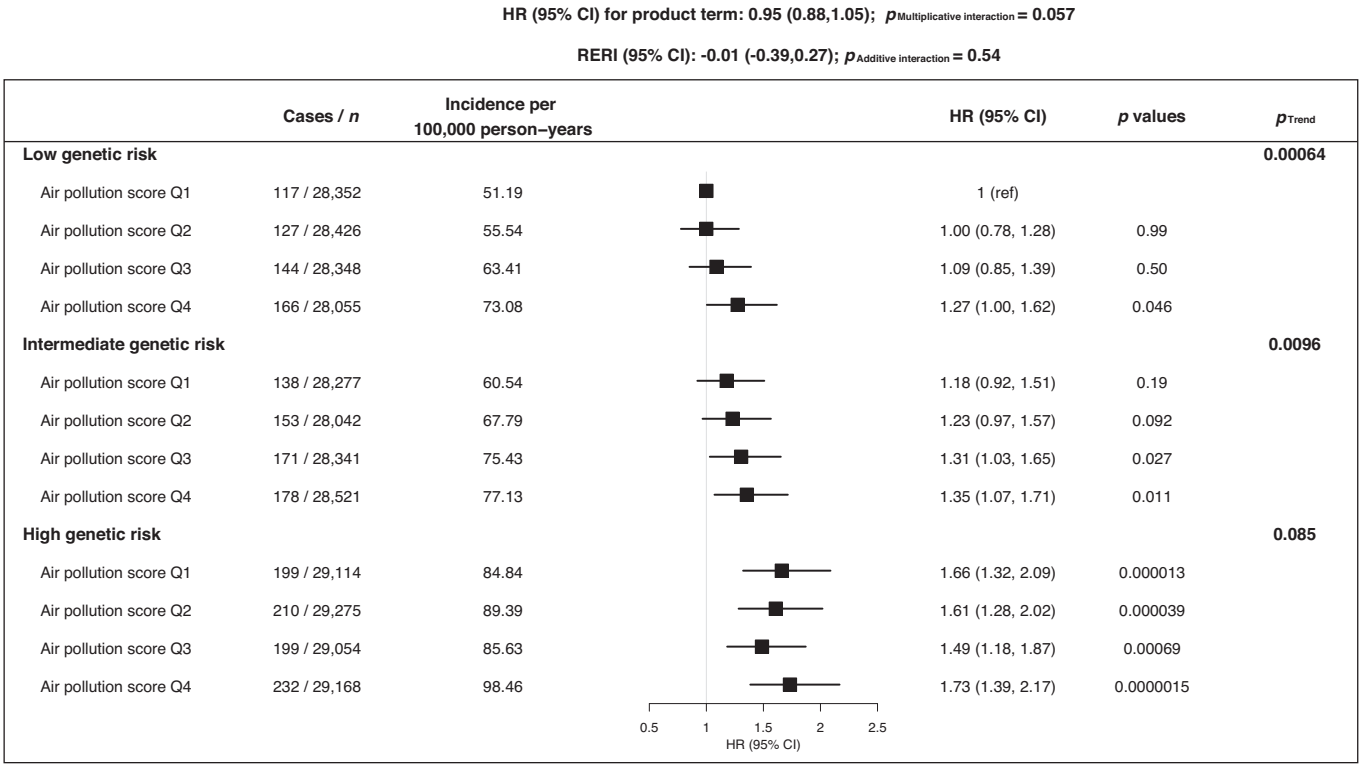

Figure 2. Association of combined air pollution and genetic risk with incident RA among UK Biobank [$N = 342{,}973$ participants ($n = 2{,}034$ incident RA)]. Adjusted for age, sex, UK Biobank assessment center, household income, education level, smoking status, body mass index, alcohol consumption, sedentary time, physical activity duration, healthy diet score, PRS, first 10 genetic principal components and genotyping batch. Genetic risk was categorized into three levels by tertiles of PRS: low (tertile 1): (−3.524 to −0.497); intermediate (tertile 2): (−0.498 to 0.297); and high (tertile 3): (0.298 to 5.066). Air pollution score ranges: quartile 1: (48.54–63.35); quartile 2: (63.36–69.94); quartile 3: (69.95–76.43); and quartile 4: (76.44–202.00). Note: CI, confidence interval; HR, hazard risk; PRS, polygenic risk score; Q, quartile; RA, rheumatoid arthritis; RERI, relative excess risk due to interaction.

these factors. The results revealed that larger effect estimates of air pollution exposure on the risk of RA were among participants who were ≥65 years of age, female, and lacking a university degree. These results could be related to several potential mechanisms. For example, the immune system and hormonal functions are different in people of different ages and genders[68,69]; for instance, the immune function and lung compensation ability of the elderly are often weakened or even impaired, making them particularly susceptible to air pollution.[70] A larger health impact of air pollution among females be may be partly explained by lifestyle factors such that more time spent at home results in better accuracy of residential air pollution exposure assessment, as well as biological factors, such as greater airway reactivity and hormonal action.[71] The increasing trend of the risk of RF-negative RA with higher air pollution scores seemed to be more apparent than that of RF-positive RA, which is consistent with the evidence from a study in the Studies of the Etiology of Rheumatoid Arthritis (SERA) that reported that ambient annual PM levels are not associated with the early development of RA-related autoimmunity prior to the development of articular RA.[74] However, the study was limited to PM and did not include other air pollutants, so the linkage of overall air pollution and early RA-related autoimmunity may need further confirmation in the future.

The results of the present study should be interpreted in the context of its strengths and limitations. The main strength of this study is that it is based on the UK Biobank prospective cohort design, with a large sample size. Our study strictly controlled confounding factors, including socioeconomic status and lifestyle, and used cross-validation to ensure the stability of the results. This study has some limitations. First, this observational study could not fully control for all unknown or unmeasured confounding factors and was unable to determine a causal relationship between air pollution and RA. Then, the measurement of air pollution was only within the baseline time range, making it impossible to further explore the important window period of the impact of air pollution exposure on the risk of RA. Further studies with repeated air pollution measurements are required to confirm our findings. Moreover, the coverage of air pollutants was not comprehensive and the exposure to air pollutants, such as SO_2, carbon monoxide, and ozone, was unavailable. In the future, it may be necessary to integrate more air pollutants to establish a more powerful and explainable air pollution score. Furthermore, although the number of incident RA cases was sufficient for the main analysis, in the specific stratified analysis the statistical power may have been limited by the decreased number of cases. In addition, we lacked information on exposure to these pollutants in locations other than participants' residential addresses, such as outdoor or work sites. This prevented us from exploring the impact of the total air pollution exposure of each participant on the incidence of RA. Moreover, UK Biobank participants are not representative of the UK general population owing to evidence of a healthy volunteer selection bias,[75] and the coverage of only the 40- to 70-y-old population may have limited us from finding more incident RA events in younger individuals. Furthermore, we determined incident RA cases based on hospital inpatient records, and latent incident RA cases with less severe clinical symptoms may not have been admitted and recorded in the hospital inpatient records. Finally, our study participants were all White, so it may be necessary to validate the generalizability of the conclusions to other ethnic populations.

Conclusions

Our findings showed that long-term combined exposure to ambient air pollutants was associated with an elevated risk of RA, and this association was more pronounced in populations with high genetic risk. We highlight the importance of comprehensive assessment for air pollution and genetic predisposition in the prevention of RA.

Acknowledgments

J.Z., X.-Y.F. and D.-Q.Y. conceived the idea for the paper. J.Z. conducted the analysis. J.Z. and X.-Y.F. are joint first authors, and had primary responsibility for drafting the manuscript. J.W., Y.-G.F., R.-X.L., B.L., X.-J.L., Y.-L.Y. contributed to the data cleaning. D.-Q.Y., C.M., J.W., Y.-G.F., R.-X.L. contributed to the analysis or interpretation of the data. All authors critically reviewed the manuscript for important intellectual content. D.-Q.Y. directed the study. D.-Q.Y. is the study guarantor and has full access to data.

The authors thank the UK Biobank participants. This research has been conducted using the UK Biobank resource under application no. 62663.

Funding was provided by the Chinese national high level personnel special support plan.

The funders played no role in the study design or implementation; data collection, management, analysis or interpretation; manuscript preparation, review or approval; or the decision to submit the manuscript for publication.

References

1. Smolen JS, Aletaha D, McInnes IB. 2016. Rheumatoid arthritis. Lancet 388(10055):2023–2038, PMID: 27156434, https://doi.org/10.1016/S0140-6736(16)30173-8.
2. van der Woude D, van der Helm-van Mil AHM. 2018. Update on the epidemiology, risk factors, and disease outcomes of rheumatoid arthritis. Best Pract Res Clin Rheumatol 32(2):174–187, PMID: 30527425, https://doi.org/10.1016/j.berh.2018.10.005.
3. Alamanos Y, Drosos AA. 2005. Epidemiology of adult rheumatoid arthritis. Autoimmun Rev 4(3):130–136, PMID: 15823498, https://doi.org/10.1016/j.autrev.2004.09.002.
4. Safiri S, Kolahi AA, Hoy D, Smith E, Bettampadi D, Mansournia MA, et al. 2019. Global, regional and national burden of rheumatoid arthritis 1990–2017: a systematic analysis of the Global Burden of Disease study 2017. Ann Rheum Dis 78(11):1463–1471, PMID: 31511227, https://doi.org/10.1136/annrheumdis-2019-215920.
5. Arleevskaya M, Takha E, Petrov S, Kazarian G, Novikov A, Larionova R, et al. 2021. Causal risk and protective factors in rheumatoid arthritis: a genetic update. J Transl Autoimmun 4:100119, PMID: 34522877, https://doi.org/10.1016/j.jtauto.2021.100119.
6. Deane KD, Demoruelle MK, Kelmenson LB, Kuhn KA, Norris JM, Holers VM. 2017. Genetic and environmental risk factors for rheumatoid arthritis. Best Pract Res Clin Rheumatol 31(1):3–18, PMID: 29221595, https://doi.org/10.1016/j.berh.2017.08.003.
7. Karlson EW, Chang SC, Cui J, Chibnik LB, Fraser PA, De Vivo I, et al. 2010. Gene–environment interaction between *HLA-DRB1* shared epitope and heavy cigarette smoking in predicting incident rheumatoid arthritis. Ann Rheum Dis 69(1):54–60, PMID: 19151010, https://doi.org/10.1136/ard.2008.102962.
8. Stolt P, Yahya A, Bengtsson C, Källberg H, Rönnelid J, Lundberg I, et al. 2010. Silica exposure among male current smokers is associated with a high risk of developing ACPA-positive rheumatoid arthritis. Ann Rheum Dis 69(6):1072–1076, PMID: 19966090, https://doi.org/10.1136/ard.2009.114694.
9. Demoruelle MK, Weisman MH, Simonian PL, Lynch DA, Sachs PB, Pedraza IF, et al. 2012. Brief report: airways abnormalities and rheumatoid arthritis-related autoantibodies in subjects without arthritis: early injury or initiating site of autoimmunity? Arthritis Rheum 64(6):1756–1761, PMID: 22183986, https://doi.org/10.1002/art.34344.
10. Glencross DA, Ho TR, Camiña N, Hawrylowicz CM, Pfeffer PE. 2020. Air pollution and its effects on the immune system. Free Radic Biol Med 151:56–68, PMID: 32007522, https://doi.org/10.1016/j.freeradbiomed.2020.01.179.
11. Huff RD, Carlsten C, Hirota JA. 2019. An update on immunologic mechanisms in the respiratory mucosa in response to air pollutants. J Allergy Clin Immunol 143(6):1989–2001, PMID: 31176381, https://doi.org/10.1016/j.jaci.2019.04.012.
12. Sigaux J, Biton J, André E, Semerano L, Boissier MC. 2019. Air pollution as a determinant of rheumatoid arthritis. Joint Bone Spine 86(1):37–42, PMID: 29524589, https://doi.org/10.1016/j.jbspin.2018.03.001.
13. Hart JE, Laden F, Puett RC, Costenbader KH, Karlson EW. 2009. Exposure to traffic pollution and increased risk of rheumatoid arthritis. Environ Health Perspect 117(7):1065–1069, PMID: 19654914, https://doi.org/10.1289/ehp.0800503.

14. Jung CR, Hsieh HY, Hwang BF. 2017. Air pollution as a potential determinant of rheumatoid arthritis: a population-based cohort study in Taiwan. Epidemiology 28(suppl 1):S54–S59, PMID: 29028676, https://doi.org/10.1097/EDE.0000000000000732.
15. Park JS, Choi S, Kim K, Chang J, Kim SM, Kim SR, et al. 2021. Association of particulate matter with autoimmune rheumatic diseases among adults in South Korea. Rheumatology (Oxford) 60(11):5117–5126, PMID: 33560298, https://doi.org/10.1093/rheumatology/keab127.
16. Hart JE, Källberg H, Laden F, Bellander T, Costenbader KH, Holmqvist M, et al. 2013. Ambient air pollution exposures and risk of rheumatoid arthritis: results from the Swedish EIRA case–control study. Ann Rheum Dis 72(6):888–894, PMID: 22833374, https://doi.org/10.1136/annrheumdis-2012-201587.
17. Hart JE, Källberg H, Laden F, Costenbader KH, Yanosky JD, Klareskog L, et al. 2013. Ambient air pollution exposures and risk of rheumatoid arthritis. Arthritis Care Res (Hoboken) 65(7):1190–1196, PMID: 23401426, https://doi.org/10.1002/acr.21975.
18. De Roos AJ, Koehoorn M, Tamburic L, Davies HW, Brauer M. 2014. Proximity to traffic, ambient air pollution, and community noise in relation to incident rheumatoid arthritis. Environ Health Perspect 122(10):1075–1080, PMID: 24905961, https://doi.org/10.1289/ehp.1307413.
19. Hunt L, Emery P. 2014. Defining populations at risk of rheumatoid arthritis: the first steps to prevention. Nat Rev Rheumatol 10(9):521–530, PMID: 24934191, https://doi.org/10.1038/nrrheum.2014.82.
20. Okada Y, Eyre S, Suzuki A, Kochi Y, Yamamoto K. 2019. Genetics of rheumatoid arthritis: 2018 status. Ann Rheum Dis 78(4):446–453, PMID: 30530827, https://doi.org/10.1136/annrheumdis-2018-213678.
21. Choi SW, Mak TS, O'Reilly PF. 2020. Tutorial: a guide to performing polygenic risk score analyses. Nat Protoc 15(9):2759–2772, PMID: 32709988, https://doi.org/10.1038/s41596-020-0353-1.
22. Khera AV, Chaffin M, Aragam KG, Haas ME, Roselli C, Choi SH, et al. 2018. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. Nat Genet 50(9):1219–1224, PMID: 30104762, https://doi.org/10.1038/s41588-018-0183-z.
23. Sparks JA, Chen CY, Jiang X, Askling J, Hiraki LT, Malspeis S, et al. 2015. Improved performance of epidemiologic and genetic risk models for rheumatoid arthritis serologic phenotypes using family history. Ann Rheum Dis 74(8):1522–1529, PMID: 24685909, https://doi.org/10.1136/annrheumdis-2013-205009.
24. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. 2015. UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med 12(3):e1001779, PMID: 25826379, https://doi.org/10.1371/journal.pmed.1001779.
25. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. 2018. The UK Biobank resource with deep phenotyping and genomic data. Nature 562(7726):203–209, PMID: 30305743, https://doi.org/10.1038/s41586-018-0579-z.
26. CDC (Centers for Disease Control and Prevention). 2013. *International Classification of Diseases, Ninth Revision, Clinical Modification* (ICD-9-CM). www.cdc.gov/nchs/icd/icd9cm.htm [accessed 2 April 2021].
27. WHO. 2016. *International Statistical Classification of Diseases and Related Health Problems, 10th Revision*. http://apps.who.int/classifications/icd10/browse/2016/en [accessed 2 April 2021].
28. Vienneau D, de Hoogh K, Bechle MJ, Beelen R, van Donkelaar A, Martin RV, et al. 2013. Western European land use regression incorporating satellite- and ground-based measurements of NO_2 and PM_{10}. Environ Sci Technol 47(23):13555–13564, PMID: 24156783, https://doi.org/10.1021/es403089q.
29. Eeftens M, Beelen R, de Hoogh K, Bellander T, Cesaroni G, Cirach M, et al. 2012. Development of land use regression models for $PM_{2.5}$, $PM_{2.5}$ absorbance, PM_{10} and PM_{coarse} in 20 European study areas; results of the ESCAPE project. Environ Sci Technol 46(20):11195–11205, PMID: 22963366, https://doi.org/10.1021/es301948k.
30. Li X, Wang M, Song Y, Ma H, Zhou T, Liang Z, et al. 2021. Obesity and the relation between joint exposure to ambient air pollutants and incident type 2 diabetes: a cohort study in UK Biobank. PLoS Med 18(8):e1003767, PMID: 34460827, https://doi.org/10.1371/journal.pmed.1003767.
31. Wang M, Zhou T, Song Y, Li X, Ma H, Hu Y, et al. 2021. Joint exposure to various ambient air pollutants and incident heart failure: a prospective analysis in UK Biobank. Eur Heart J 42(16):1582–1591, PMID: 33527989, https://doi.org/10.1093/eurheartj/ehaa1031.
32. Okada Y, Wu D, Trynka G, Raj T, Terao C, Ikari K, et al. 2014. Genetics of rheumatoid arthritis contributes to biology and drug discovery. Nature 506(7488):376–381, PMID: 24390342, https://doi.org/10.1038/nature12873.
33. Babb de Villiers C, Kroese M, Moorthie S. 2020. Understanding polygenic models, their development and the potential application of polygenic scores in healthcare. J Med Genet 57(11):725–732, PMID: 32376789, https://doi.org/10.1136/jmedgenet-2019-106763.
34. Choi SW, O'Reilly PF. 2019. PRSice-2: polygenic risk score software for biobank-scale data. Gigascience 8(7):giz082, PMID: 31307061, https://doi.org/10.1093/gigascience/giz082.
35. Lloyd-Jones DM, Hong Y, Labarthe D, Mozaffarian D, Appel LJ, Van Horn L, et al. 2010. Defining and setting national goals for cardiovascular health promotion and disease reduction: the American Heart Association's Strategic Impact Goal through 2020 and beyond. Circulation 121(4):586–613, PMID: 20089546, https://doi.org/10.1161/CIRCULATIONAHA.109.192703.
36. Kalinowski A, Humphreys K. 2016. Governmental standard drink definitions and low-risk alcohol consumption guidelines in 37 countries. Addiction 111(7):1293–1298, PMID: 27073140, https://doi.org/10.1111/add.13341.
37. Schoenfeld D. 1982. Partial residuals for the proportional hazards regression model. Biometrika 69(1):239–241, https://doi.org/10.1093/biomet/69.1.239.
38. White IR, Royston P, Wood AM. 2011. Multiple imputation using chained equations: issues and guidance for practice. Stat Med 30(4):377–399, PMID: 21225900, https://doi.org/10.1002/sim.4067.
39. Simon RM, Subramanian J, Li MC, Menezes S. 2011. Using cross-validation to evaluate predictive accuracy of survival risk classifiers based on high-dimensional data. Brief Bioinform 12(3):203–214, PMID: 21324971, https://doi.org/10.1093/bib/bbr001.
40. Keil AP, Buckley JP, O'Brien KM, Ferguson KK, Zhao S, White AJ. 2020. A quantile-based g-computation approach to addressing the effects of exposure mixtures. Environ Health Perspect 128(4):47004, PMID: 32255670, https://doi.org/10.1289/EHP5838.
41. Bobb JF, Valeri L, Claus Henn B, Christiani DC, Wright RO, Mazumdar M, et al. 2015. Bayesian kernel machine regression for estimating the health effects of multi-pollutant mixtures. Biostatistics 16(3):493–508, PMID: 25532525, https://doi.org/10.1093/biostatistics/kxu058.
42. McQueenie R, Nicholl BI, Jani BD, Canning J, Macdonald S, McCowan C, et al. 2020. Patterns of multimorbidity and their effects on adverse outcomes in rheumatoid arthritis: a study of 5658 UK Biobank participants. BMJ Open 10(11):e038829, PMID: 33234629, https://doi.org/10.1136/bmjopen-2020-038829.
43. Beckman Coulter, Inc. 2020. Chemistry Information Sheet on Rheumatoid Factor (REF44707). https://www.beckmancoulter.com/download/file/phx988646AP-EN_US/988646AP?type=pdf [accessed 10 August 2022].
44. Zhang Z, Chen L, Wang X, Wang C, Yang Y, Li H, et al. 2023. Associations of air pollution and genetic risk with incident dementia: a prospective cohort study. Am J Epidemiol 192(2):182–194, PMID: 36269005, https://doi.org/10.1093/aje/kwac188.
45. Vieira VM, Hart JE, Webster TF, Weinberg J, Puett R, Laden F, et al. 2010. Association between residences in U.S. northern latitudes and rheumatoid arthritis: a spatial analysis of the Nurses' Health Study. Environ Health Perspect 118(7):957–961, PMID: 20338859, https://doi.org/10.1289/ehp.0901861.
46. Arkema EV, Hart JE, Bertrand KA, Laden F, Grodstein F, Rosner BA, et al. 2013. Exposure to ultraviolet-B and risk of developing rheumatoid arthritis among women in the Nurses' Health Study. Ann Rheum Dis 72(4):506–511, PMID: 23380431, https://doi.org/10.1136/annrheumdis-2012-202302.
47. Veness C. 2022. Geodesy Library: Movable Type scripts. https://www.movable-type.co.uk/scripts/geodesy-library.html [accessed 10 August 2022].
48. UK Office for National Statistics. 2018. Countries (December 2017) Full Clipped Boundaries in Great Britain. Last updated December 2019. https://www.data.gov.uk/dataset/3fd8d2d2-b591-42ff-b333-c53a6a513e96/countries-december-2017-full-clipped-boundaries-in-great-britain [accessed 23 November 2022].
49. UK Department for Environment Food & Rural Affairs. 2019. Modelled background pollution data. https://uk-air.defra.gov.uk/data/pcm-data [accessed 6 April 2022].
50. Costa S, Ferreira J, Silveira C, Costa C, Lopes D, Relvas H, et al. 2014. Integrating health on air quality assessment—review report on health risks of two major European outdoor air pollutants: PM and NO_2. J Toxicol Environ Health B Crit Rev 17(6):307–340, PMID: 25333993, https://doi.org/10.1080/10937404.2014.946164.
51. Dominici F, Peng RD, Barr CD, Bell ML. 2010. Protecting human health from air pollution: shifting from a single-pollutant to a multipollutant approach. Epidemiology 21(2):187–194, PMID: 20160561, https://doi.org/10.1097/EDE.0b013e3181cc86e8.
52. Johns DO, Stanek LW, Walker K, Benromdhane S, Hubbell B, Ross M, et al. 2012. Practical advancement of multipollutant scientific and risk assessment approaches for ambient air pollution. Environ Health Perspect 120(9):1238–1242, PMID: 22645280, https://doi.org/10.1289/ehp.1204939.
53. Hong YC, Leem JH, Ha EH, Christiani DC. 1999. PM_{10} exposure, gaseous pollutants, and daily mortality in Inchon, South Korea. Environ Health Perspect 107(11):873–878, PMID: 10544154, https://doi.org/10.1289/ehp.99107873.
54. Park SK, Tao Y, Meeker JD, Harlow SD, Mukherjee B. 2014. Environmental risk score as a new tool to examine multi-pollutants in epidemiologic research: an example from the NHANES study using serum lipid levels. PLoS One 9(6):e98632, PMID: 24901996, https://doi.org/10.1371/journal.pone.0098632.
55. Lourida I, Hannon E, Littlejohns TJ, Langa KM, Hyppönen E, Kuzma E, et al. 2019. Association of lifestyle and genetic risk with incidence of dementia. JAMA 322(5):430–437, PMID: 31302669, https://doi.org/10.1001/jama.2019.9879.

56. Ahola-Olli AV, Mustelin L, Kalimeri M, Kettunen J, Jokelainen J, Auvinen J, et al. 2019. Circulating metabolites and the risk of type 2 diabetes: a prospective study of 11,896 young adults from four Finnish cohorts. Diabetologia 62(12):2298–2309, PMID: 31584131, https://doi.org/10.1007/s00125-019-05001-w.
57. Klareskog L, Padyukov L, Lorentzen J, Alfredsson L. 2006. Mechanisms of disease: genetic susceptibility and environmental triggers in the development of rheumatoid arthritis. Nat Clin Pract Rheumatol 2(8):425–433, PMID: 16932734, https://doi.org/10.1038/ncprheum0249.
58. Linn-Rasker SP, van der Helm-van Mil AHM, van Gaalen FA, Kloppenburg M, de Vries RRP, le Cessie S, et al. 2006. Smoking is a risk factor for anti-CCP antibodies only in rheumatoid arthritis patients who carry HLA-DRB1 shared epitope alleles. Ann Rheum Dis 65(3):366–371, PMID: 16014670, https://doi.org/10.1136/ard.2005.041079.
59. Kim K, Bang SY, Lee HS, Bae SC. 2017. Update on the genetic architecture of rheumatoid arthritis. Nat Rev Rheumatol 13(1):13–24, PMID: 27811914, https://doi.org/10.1038/nrrheum.2016.176.
60. Lenz TL, Deutsch AJ, Han B, Hu X, Okada Y, Eyre S, et al. 2015. Widespread non-additive and interaction effects within HLA loci modulate the risk of autoimmune diseases. Nat Genet 47(9):1085–1090, PMID: 26258845, https://doi.org/10.1038/ng.3379.
61. Zhao CN, Xu Z, Wu GC, Mao YM, Liu LN, Dan YL, et al. 2019. Emerging role of air pollution in autoimmune diseases. Autoimmun Rev 18(6):607–614, PMID: 30959217, https://doi.org/10.1016/j.autrev.2018.12.010.
62. Farhat SCL, Silva CA, Orione MAM, Campos LMA, Sallum AME, Braga ALF. 2011. Air pollution in autoimmune rheumatic diseases: a review. Autoimmun Rev 11(1):14–21, PMID: 21763467, https://doi.org/10.1016/j.autrev.2011.06.008.
63. Essouma M, Noubiap JJN. 2015. Is air pollution a risk factor for rheumatoid arthritis? J Inflamm (Lond) 12:48, PMID: 26225124, https://doi.org/10.1186/s12950-015-0092-1.
64. Chen R, Li H, Cai J, Wang C, Lin Z, Liu C, et al. 2018. Fine particulate air pollution and the expression of microRNAs and circulating cytokines relevant to inflammation, coagulation, and vasoconstriction. Environ Health Perspect 126(1):017007, PMID: 29342453, https://doi.org/10.1289/EHP1447.
65. Fiorito G, Vlaanderen J, Polidoro S, Gulliver J, Galassi C, Ranzi A, et al. 2018. Oxidative stress and inflammation mediate the effect of air pollution on cardio- and cerebrovascular disease: a prospective study in nonsmokers. Environ Mol Mutagen 59(3):234–246, PMID: 29114965, https://doi.org/10.1002/em.22153.
66. Adami G, Viapiana O, Rossini M, Orsolini G, Bertoldo E, Giollo A, et al. 2021. Association between environmental air pollution and rheumatoid arthritis flares. Rheumatology (Oxford) 60(10):4591–4597, PMID: 33470401, https://doi.org/10.1093/rheumatology/keab049.
67. Adami G, Rossini M, Viapiana O, Orsolini G, Bertoldo E, Pontalti M, et al. 2021. Environmental air pollution is a predictor of poor response to biological drugs in chronic inflammatory arthritides. ACR Open Rheumatol 3(7):451–456, PMID: 34060251, https://doi.org/10.1002/acr2.11270.
68. van Onna M, Boonen A. 2016. The challenging interplay between rheumatoid arthritis, ageing and comorbidities. BMC Musculoskelet Disord 17(1):184, PMID: 27118031, https://doi.org/10.1186/s12891-016-1038-3.
69. Favalli EG, Biggioggero M, Crotti C, Becciolini A, Raimondo MG, Meroni PL. 2019. Sex and management of rheumatoid arthritis. Clinic Rev Allerg Immunol 56(3):333–345, PMID: 29372537, https://doi.org/10.1007/s12016-018-8672-5.
70. Kelly FJ, Dunster C, Mudway I. 2003. Air pollution and the elderly: oxidant/antioxidant issues worth consideration. Eur Respir J Suppl 40:70s–75s, PMID: 12762578, https://doi.org/10.1183/09031936.03.00402903.
71. Doiron D, de Hoogh K, Probst-Hensch N, Fortier I, Cai Y, De Matteis S, et al. 2019. Air pollution, lung function and COPD: results from the population-based UK Biobank study. Eur Respir J 54(1):1802140, PMID: 31285306, https://doi.org/10.1183/13993003.02140-2018.
72. Kan H, Heiss G, Rose KM, Whitsel E, Lurmann F, London SJ. 2007. Traffic exposure and lung function in adults: the Atherosclerosis Risk in Communities study. Thorax 62(10):873–879, PMID: 17442705, https://doi.org/10.1136/thx.2006.073015.
73. O'Neill MS, Jerrett M, Kawachi I, Levy JI, Cohen AJ, Gouveia N, et al. 2003. Health, wealth, and air pollution: advancing theory and methods. Environ Health Perspect 111(16):1861–1870, PMID: 14644658, https://doi.org/10.1289/ehp.6334.
74. Gan RW, Deane KD, Zerbe GO, Demoruelle MK, Weisman MH, Buckner JH, et al. 2013. Relationship between air pollution and positivity of RA-related autoantibodies in individuals without established RA: a report on SERA. Ann Rheum Dis 72(12):2002–2005, PMID: 23572338, https://doi.org/10.1136/annrheumdis-2012-202949.
75. Fry A, Littlejohns TJ, Sudlow C, Doherty N, Adamska L, Sprosen T, et al. 2017. Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. Am J Epidemiol 186(9):1026–1034, PMID: 28641372, https://doi.org/10.1093/aje/kwx246.