Environ Health Perspect

DOI: 10.1289/EHP10710

Supplemental Material

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

Jie Zhang, Xin-Yu Fang, Jun Wu, Yin-Guang Fan, Rui-Xue Leng, Bo Liu, Xiao-Jie Lv, Yu-Lu Yan, Chen Mao, and Dong-Qing Ye

Table of Contents

Table S12. Association between air pollution score and incident rheumatoid arthritis among UK Biobank 342,973 participants (2,034 incident RA events), additionally adjusted the latitude of participants' residence in the sensitivity analysis.

Table S13. Association between air pollution score and incident rheumatoid arthritis among UK Biobank 292,987 participants (1,745 incident RA events), after limiting participants living in the current address for at least five years in the sensitivity analysis.

Table S14. Association between air pollution score and incident rheumatoid arthritis among UK Biobank 313,466 participants (1,701 incident RA events), after excluding participants who used NSAIDs previously or at baseline in the sensitivity analysis.

Figure S1. Overview of design and population selection in the UK Biobank study of the association of air pollution and genetic risk with rheumatoid arthritis incidence.

Figure S2. Exposure-response relationships of each air pollutant with incident rheumatoid arthritis among UK Biobank 342,973 participants (2,034 incident RA events).

Figure S3. Joint distribution of air pollution scores and polygenic risk scores between RA and non-RA among UK Biobank 342,973 participants (2,034 incident RA events).

Figure S4. Validation of association between air pollution score and incident rheumatoid arthritis by 10-fold cross-validation.

Figure S5. Exposure-response relationship of air pollution score with incident rheumatoid arthritis according to age.

Figure S6. Exposure-response relationship of air pollution score with incident rheumatoid arthritis according to sex.

Figure S7. Exposure-response relationship of air pollution score with incident rheumatoid arthritis according to smoking status.

Figure S8. Exposure-response relationship of air pollution score with incident rheumatoid arthritis according to education level.

Figure S9. Exposure-response relationship of air pollution score with rheumatoid arthritis incidence according to rheumatoid factor (RF) status.

Additional File- Excel Document

Supplemental Table S1 Source and corresponding code to identify rheumatoid arthritis cases and relevant prescription medication in the UK Biobank study**.**

Source	UK Biobank data fields	Corresponding code
Hospital inpatient data RA diagnoses of ICD-10, ICD-9	41720: "Diagnoses - ICD10" 41271: "Diagnoses - ICD9"	ICD-9: 71400, 71401, 71403, 71404, 71405, 71406, 71409 ICD-10: M05.0, M05.1, M05.2, M05.3, M05.8, M05.9, M06.0, M06.1, M06.2, M06.3, M06.4, M06.9
Medical conditions self-report past and current RA through a verbal interview by a trained nurse.	20002: "Non-cancer illness code, self-reported"	1464
Prescription medication self-report past and current medicine are taken through a interview by a trained nurse, including the following medicine: (1) Steroids: corticosteroids, depomedrone, triamcinolone, methylprednisolone, prednisolone, prednisone (also listed in UK Biobank as Deltacortril enteric; Deltastab; Precortisyl; Prednesola) (2) Synthetic disease-modifying anti-rheumatic drugs (DMARDs): Auranofin (also listed in UK Biobank as Ridaura), Azathioprine (also listed in UK Biobank as Imuran), hydroxychloroquine (also listed in UK Biobank as Plaquenil), leflunomide (also listed in UK Biobank as Arava), methotrexate, methotrexate injections, myocrisin, penicillamine, sulfasalazine (also listed in UK Biobank as Sulazine; salazopyrin; sulphasalazine) (3) Biologic DMARDS: abatacept, adalimumab (also listed in UK Biobank as Humira injection solution), certolizumab, etanercept, golimumab, infliximab, rituximab, tocilizumab (4) Nonsteroidal anti-inflammatory drugs (NSAIDs): Indomethacin, Sulindac, Diclofenac, Etodolac, Acemetacin, Ketorolac, Aceclofenac, Diclofenac|Misoprostol, Piroxicam, Tenoxicam, Meloxicam, Ibuprofen, Naproxen, Ketoprofen, Flurbiprofen, Tiaprofenic Acid, Ibuprofen |Menthol, Naproxen|Misoprostol, Mefenamic Acid, Tolfenamic Acid, Celecoxib, Etoricoxib, Nabumetone, Glucosamine, Chondroitin Sulfate	20003: "Treatment/medication code"	Steroids, Synthetic DMARDs, and Biologic DMARDS: 1140874936,1140874940,1140874944,1140874950, 1140874954,1140874956,1140874978,1140868426, 1140883058,1140883060,1140883062,1140883064, 1140874976,1140800000,1140874930,1140868364, 1140868370,1140874936,1140874940,1140875316, 1140874944,1140874950,1140875400,1140875404, 1140909864,1140869930,1141145996,1140884308, 1140875392,1141166294,1141166302,1141166304, 1141166306,1140910036,1140869848,1140853054, 1140853056,1140875304,1140875306,1140875308, 1140909702,1141188900, 1140865670,1140865668, 1141188588,1141188594 NSAIDs: 1140871354,1140871604,1140921828,11408711681 140917394,1140871180,1140923920,1140909354, 1140877872,1141182674,1140871174,11411674261, 140877874,1140871188,1140871196,1141193170, 1140875278,1140884558,1140925806,1140871266, 1140927086,1140878036,1140871672,1141169530, 1140871666,1140875346,1140926732,1140926796, 1140926794,1141187776,1140871388,1141153134, 1140871472,1140871522,1140871506,1140871238, 1140871236,1140871616,1140911748,1140871638, 1140881612,1140871542,1140871546,1140928840, 1141176668,1141176670,1141180140,1141180148, 1141180150,1141180152,1140875336,1140875338, 1141188442,1187,1140871666,1140875346, 1140926732,1140926796,1140926794

Supplemental Table S2 Components of healthy diet score and alcohol consumption in the UK Biobank study.

Factors	Source and definition	UK Biobank data fields
Healthy diet score	UK Biobank Food Frequency Questionnaire at baseline: A healthy diet score was adapted from the American Heart Association Guidelines.[1] Each healthy dietary component counts as one point of a healthy diet score: 1. Vegetable intake ≥ 4 tablespoons/day 2. Fruit intake ≥ 3 pieces/day 3. Fish intake ≥ twice/week 4. Processed meats ≤ twice/week 5. Unprocessed red meats ≤ twice/week	1289: "Cooked vegetable intake" 1299: "Salad / raw vegetable intake" 1309: "Fresh fruit intake" 1319: "Dried fruit intake" 1329: "Oily fish intake" 1339: "Non-oily fish intake" 1349: "Processed meat intake" 1359: "Poultry intake" 1369: "Beef intake" 1379: "Lamb/mutton intake" 1389: "Pork intake"
Alcohol consumption	UK Biobank Touchscreen questionnaire at baseline; Conversion of self-reported alcohol intake to standard drink of alcohol per day, [1] standard drink=8 g. To compute the standard drink of alcohol per day: for each source of alcohol, multiply by the given factor and divide by 7 (if the input is weekly intake) or 30 (if the input is monthly intake) to get the standard drink/day.[2] Given the factor of each source of alcohol: Red wine:1.5 Champagne/White wine:1.5 Beer/Cider:2.5 Spirits:1 Fortified wine:1 Other alcoholic drinks:1.5	220117: "Alcohol drinker status" 1558: "Alcohol intake frequency" 1568: "Average weekly red wine intake" 1578: "Average weekly champagne plus white wine intake" 1588: "Average weekly beer plus cider intake" 1608: "Average weekly fortified wine intake" 5364: "Average weekly intake of other alcoholic drinks" 4407: "Average monthly red wine intake" 4418: "Average monthly champagne plus white wine intake" 4429: "Average monthly beer plus cider intake" 4440: "Average monthly spirits intake" 4451: "Average monthly fortified wine intake" 4462: "Average monthly intake of other alcoholic drinks"

1. Lloyd-Jones DM, Hong Y, Labarthe D, Mozaffarian D, Appel LJ, Van Horn L, Greenlund K, Daniels S, Nichol G, Tomaselli GF, Arnett DK, Fonarow GC, Ho PM, Lauer MS, Masoudi FA, Robertson RM, Roger V, Schwamm LH, Sorlie P, Yancy CW, Rosamond WD; American Heart Association Strategic Planning Task Force and Statistics Committee. 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. 2010 Feb 2;121(4):586-613. doi: 10.1161/CIRCULATIONAHA.109.192703.

2. Kalinowski A, Humphreys K. Governmental standard drink definitions and low-risk alcohol consumption guidelines in 37 countries. Addiction. 2016 Jul;111(7):1293-8. doi: 10.1111/add.13341. Epub 2016 Apr 13.

Supplemental Table S3 A comparison of baseline characteristics between the current study population (342,973 participants) and the full UK Biobank cohort (502,505 participants).

Characteristics [a]	Study population (*n* = 342,973)	UK Biobank population [b] (*n* = 502,505)
Age (years)	56.99 (7.93)	56.53 (8.10)
Follow-up time (years), (median, IQR)	8.1 (1.2)	
Sex (*n*, %)		
female	183,218 (53.42)	273,382 (54.40)
male	159,755 (46.58)	229,123 (45.60)
household income (£)		
less than 18,000	76,921 (22.43)	97,198 (19.34)
18,000 to 29,999	89,325 (26.04)	108,177 (21.53)
30,000 to 51,999	90,327 (26.34)	110,772 (22.04)
52,000 to 100,000	68,873 (20.08)	86,266 (17.16)
greater than 100,000	17,527(5.11)	22,929(4.56)
Education (*n*, %)		
college or university degree	122,157 (35.62)	161,163 (32.07)
other	220,816 (64.38)	341,342 (48.94)
Smoking status (*n*, %)		
never smoking	187,258 (54.60)	273,522 (54.43)
previous smoking	121,670 (35.48)	173,056 (34.44)
current smoking	34,045 (9.93)	52,978 (10.54)
Alcohol consumption (standard-drink/day [c])	2.03 (2.44)	1.93 (2.44)
Healthy diet score (*n*, %)		
0-1	77,980 (22.74)	94,431 (18.79)
2-3	185,559 (54.10)	223,542 (44.49)
4-5	79,434 (23.16)	98,408 (19.58)
Sedentary time (hours/day)	4.87 (2.38)	4.85 (2.45)
Physical activity (mins/day)	128.80 (102.82)	127.09 (101.90)
Body mass index (kg/m^2)	27.42 (4.75)	27.42 (4.75)
Rheumatoid factor (RF) status (*n*, %)		
RF-positive	12,220 (3.56)	18,713 (3.72)
RF-negative	314,100 (91.58)	450,847 (89.71)
$PM_{2.5}$ (μg/m^3), (median, IQR)	9.88 (1.26)	9.93 (1.28)
$PM_{2.5-10}$ (μg/m^3), (median, IQR)	6.08 (0.76)	6.11 (0.80)
PM_{10} (μg/m^3), (median, IQR)	38.01 (4.44)	38.31 (4.67)
NO_2 (μg/m^3), (median, IQR)	27.30 (10.22)	27.93 (10.73)
NO_x (μg/m^3), (median, IQR)	41.27 (16.10)	42.43 (16.38)
Air pollution score	70.63 (11.13)	

Abbreviation: $PM_{2.5}$, particular matter with aerodynamic diameter < 2.5 mm; PM_{10}, particular matter with an aerodynamic diameter < 10 mm; $PM_{2.5–10}$, particular matter with an aerodynamic diameter between 2.5 and 10 mm; NO_2, nitrogen dioxide; NO_x, nitrogen oxides. IQR, interquartile range.

[a]Data are mean ±standard deviation (SD) unless otherwise indicated. 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), rheumatoid factor (n = 16,653)

[b] Follow-up time and air pollution score were only available in the study population. And the sum of proportion and less than 100% of each characteristic due to missing values were not imputed in UK Biobank population.

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

Supplemental Table S4 Spearman correlations between the individual air pollutant incorporated in the air pollution score.

	NO_2 ($\mu g/m^3$)	NO_x ($\mu g/m^3$)	$PM_{2.5}$ ($\mu g/m^3$)	$PM_{2.5–10}$ ($\mu g/m^3$)	PM_{10} ($\mu g/m^3$)
NO_2 ($\mu g/m^3$)	1				
NO_x ($\mu g/m^3$)	0.80*	1			
$PM_{2.5}$ ($\mu g/m^3$)	0.75*	0.87*	1		
$PM_{2.5–10}$ ($\mu g/m^3$)	0.23*	0.26*	0.25*	1	
PM_{10} ($\mu g/m^3$)	0.72*	0.64*	0.62*	0.54*	1

*Indicates $P < 0.001$

Supplemental Table S5 Weights of each air pollutant included in the calculation of air pollution score in the main analysis.

Air pollutants($\mu g/m^3$)	β-weights in the air pollution score[a]	SD
$PM_{2.5}$	0.0527914	0.0212967
$PM_{2.5-10}$	-0.0009232	0.02468
PM_{10}	0.002401	0.006354
NO_2	0.002948	0.00278
NO_x	0.0040157	0.001425

NO_x, nitric oxide; NO_2, nitrogen dioxide; $PM_{2.5}$, 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. [a]Adjusted for age, sex, UK Biobank assessment center, household income education, smoking status, body mass index, alcohol consumption, sedentary time, physical activity duration, and healthy diet score.

Supplemental Table S6 Comparison of contribution proportion of individual air pollutant to total effect among weighted air pollution score, quantile g-computation, and Bayesian kernel machine regression in the sensitivity analysis.

Individual air pollutants ($\mu g/m^3$)	Weighted air pollution score			Quantile g-computation	Bayesian kernel machine regression	
	β [a]	*P*-values	Scaled effect [b]	Weight percent and Direction	Conditional PIP	Group PIP [c]
$PM_{2.5}$	0.05	0.013	0.026852	19.6% in the sum of positive coefficients (0.137)	0.77	1
PM_{10}	0.0024	0.7	-0.0585	100% in the sum of negative coefficients (-0.0585)	0.05	
$PM_{2.5–10}$	-0.0009	0.97	0.019865	14.5% in the sum of positive coefficients (0.137)	0.12	
NO_2	0.0029	0.29	0.033428	24.2% in the sum of positive coefficients (0.137)	0.13	
NO_x	0.004	0.0048	0.056855	41.5% in the sum of positive coefficients (0.137)	0.94	

NO_x, nitric oxide; NO_2, nitrogen dioxide; $PM_{2.5}$, 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. PIP, posterior inclusion probabilities

[a] log(HR) (95% CI in incident RA per $\mu g/m^3$ increase in air pollutants

[b] log(HR) (95% CI) in incident RA per quartile increase in air pollutants

[c] Groups included particulate matters and nitrogen oxides.

All models were adjusted for age, sex, UK Biobank assessment center, household income education, smoking status, body mass index, alcohol consumption, sedentary time, physical activity duration, and healthy diet score.

Supplemental Table S7 Comparison of the total effect of the mixture of air pollutants among weighted air pollution score, quantile g-computation, and Bayesian kernel machine regression in the sensitivity analysis.

The mixture of air pollutants	Weighted air pollution score	Quantile g-computation	Bayesian kernel machine regression
	1.06 (1.01, 1.10) [a]	1.08 (1.02, 1.15) [b]	
Quartile 1	Ref	Ref	Ref
Quartile 2	1.00 (0.89, 1.14)	1.08 (1.02, 1.15)	1.05 (1.03, 1.08)
Quartile 3	1.01 (0.89, 1.15)	1.17 (1.04, 1.31)	1.05 (1.03, 1.08)
Quartile 4	1.14 (1.00, 1.29)	1.26 (1.06, 1.51)	1.41 (1.34, 1.49)

[a] Hazard ratio (95% confidence interval) in incident RA per SD increase in air pollution score.

[b]Ψ; HR (95% CI) in incident RA per quartile increase in all air pollutants

All models were adjusted for age, sex, UK Biobank assessment center, household income education, smoking status, body mass index, alcohol consumption, sedentary time, physical activity duration, and healthy diet score.

Supplemental Table S8 Association between air pollution score and RA risk in different subgroups, and interaction analysis, according to age, sex, education level, smoking status, and rheumatoid factor status.

Subgroup	case / *n*	HR (95% CI) [a]	Multiplicative interaction		Additive interaction	
			HR (95% CI) [b]	*P*-value	RERI (95% CI)	*P*-value
Age			1.13 (1.02, 1.27)	0.044	0.04 (-0.20, 0.27)	0.38
≥65	625 / 68,253	1.09 (1.01, 1.18)				
<65	1,409 / 274,720	1.02 (0.97, 1.07)				
Sex			1.43 (1.16, 1.77)	0.005	0.58 (0.28, 0.87)	0.0009
Women	1,368 / 183,218	1.10 (1.05, 1.16)				
Men	666 / 159,755	0.97 (0.90, 1.05)				
Smoking status			0.98 (0.77, 1.13)	0.99	-0.02 (-0.28, 0.23)	0.58
Never	921 / 187,258	1.06 (0.99, 1.13)				
Previous/Current	1,113 / 155,715	1.06 (1.00, 1.12)				
Education			1.21 (0.96, 1.51)	0.14	0.23 (0.002, 0.46)	0.026
University degree	542 / 122,157	1.02 (0.94, 1.11)				
No university degree	1,492 / 220,816	1.07 (1.02, 1.13)				
RF-status			0.77 (0.65, 0.86)	0.007	0.62 (-0.82, 2.08)	0.2
Positive	417 / 341,356	1.07 (0.97, 1.17)				
Negative	1,543 / 342,482	1.08 (1.02, 1.13)				

HR (95% CI): Hazard ratio (95% confidence interval); RERI: relative excess risk due to interaction

[a] HR (95% CI) in incident RA per SD increase in air pollution score.

[b] HR (95% CI) for the product term included in the model were age ≥ 65 years, women, no university degree, previous/current smoking, RF positive and 4th quartile air pollution score.

All subgroups were adjusted for age, sex, UK Biobank assessment center, household income education, smoking status, body mass index, alcohol consumption, sedentary time, physical activity duration and healthy diet score. Menopausal status (yes or no) and hormone replacement therapy use (yes or no) were additionally adjusted in women.

Supplemental Table S9 Association between air pollution score（eliminating $PM_{2.5-10}$ and PM_{10}）and incident rheumatoid arthritis among UK Biobank 342,973 participants (2,034 incident RA events) in the sensitivity analysis

Air pollution score[a]	case / *n*	HR (95% CI) of incident RA			*P* for trend
		Model 1[b]	Model 2[c]	Model 3[d]	
Per SD increment	-	1.12 (1.07, 1.17)	1.06 (1.01, 1.10)	1.06 (1.01, 1.11)	
Quartile 1	455 / 85,743	Ref	Ref	Ref	0.000041
Quartile 2	487 / 85,743	1.08 (0.95, 1.23)	0.99 (0.87, 1.13)	0.99 (0.87, 1.13)	
Quartile 3	514 / 85,743	1.17 (1.03, 1.33)	1.01 (0.89, 1.15)	1.01 (0.89, 1.15)	
Quartile 4	578 / 85,744	1.34 (1.18, 1.51)	1.13 (1.00, 1.28)	1.13 (1.00, 1.28)	

HR, hazard ratio; CI, confidence interval.

[a]Air pollution score ranges: quartile 1: (26.95, 34.83); quartile 2: (34.84, 38.69); quartile 3: (38.70, 42.48); quartile 4: (42.49, 118.04). Mean (SD) of the air pollution score is 39.11 (6.53)

[b]Adjusted for age and sex.

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

Supplemental Table S10 Association between air pollution score（including $PM_{2.5}$ absorbance） and incident rheumatoid arthritis among UK Biobank 342,973 participants (2,034 incident RA events) in the sensitivity analysis

Exposure	case / *n*	HR (95% CI) of incident RA			*P* for trend
		Model 1[a]	Model 2[b]	Model 3[c]	
$PM_{2.5}$ absorbance [d]					
Per IQR increment	-	1.06 (1.01, 1.11)	1.02 (0.97, 1.07)	1.02 (0.97, 1.07)	
Quartile 1	467 / 85,434	Ref	Ref	Ref	0.10
Quartile 2	553 / 90,353	1.12 (0.99, 1.26)	1.04 (0.92, 1.17)	1.04 (0.91, 1.17)	
Quartile 3	524 / 82,066	1.17 (1.03, 1.32)	1.05 (0.92, 1.19)	1.05 (0.92, 1.18)	
Quartile 4	490 / 83,086	1.16 (1.02, 1.31)	1.03 (0.90, 1.17)	1.02 (0.90, 1.17)	
Air pollution score [e]					
Per SD increment	-	1.11 (1.07, 1.16)	1.05 (1.01, 1.10)	1.05 (1.00, 1.10)	
Quartile 1	459 / 85,744	Ref	Ref	Ref	0.00017
Quartile 2	488 / 85,743	1.07 (0.95, 1.22)	0.98 (0.87, 1.12)	1.00 (0.86, 1.12)	
Quartile 3	518 / 85,742	1.17 (1.03, 1.32)	1.01 (0.89, 1.14)	1.00 (0.88, 1.14)	
Quartile 4	569 / 85,744	1.31 (1.16, 1.48)	1.11 (0.98, 1.26)	1.11 (0.98, 1.26)	

HR, hazard ratio; CI, confidence interval.

[a]Adjusted for age and sex.

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

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

[d]$PM_{2.5}$ absorbance ranges: quartile 1: (0.83, 0.98) $\mu g/m^3$, quartile 2: (0.99, 1.11) $\mu g/m^3$, quartile 3: (1.12, 1.27) $\mu g/m^3$, quartile 4: (1.28, 4.60) $\mu g/m^3$ and IQR is 0.29 $\mu g/m^3$

[e]Air pollution score ranges: quartile 1: (25.27, 32.86); quartile 2: (32.87, 36.24); quartile 3: (36.25, 39.67); quartile 4: (39.68, 103.65). Mean (SD) of the air pollution score is 36.65 (5.77)

Supplemental Table S11 Association between air pollution score and incident rheumatoid arthritis among UK Biobank 342,691 participants (1,752 incident RA events), after excluding participants with a follow-up time of fewer than 2 years in the sensitivity analysis.

Air pollution score[a]	case / *n*	HR (95% CI) of incident RA			*P* for trend
		Model 1[b]	Model 2[c]	Model 3[e]	
Per SD increment	-	1.11 (1.06, 1.16)	1.05 (1.00, 1.10)	1.05 (1.00, 1.10)	
Quartile 1	384 / 85,673	Ref	Ref	Ref	0.000091
Quartile 2	423 / 85,676	1.11 (0.97, 1.28)	1.03 (0.89, 1.18)	1.03 (0.89, 1.18)	
Quartile 3	451 / 85,680	1.22 (1.06, 1.39)	1.06 (0.92, 1.21)	1.06 (0.92, 1.21)	
Quartile 4	494 / 85,662	1.36 (1.19, 1.55)	1.16 (1.01, 1.33)	1.16 (1.01, 1.33)	

HR, hazard ratio; CI, confidence interval.

[a]Air pollution score ranges: quartile 1: (48.54, 63.35); quartile 2: (63.36, 69.95); quartile 3: (69.96, 76.43); 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, 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, smoking status, body mass index, alcohol consumption, sedentary time, physical activity duration, healthy diet score, polygenic risk score, first ten genetic principal components, and genotyping batch.

Supplemental Table S12 Association between air pollution score and incident rheumatoid arthritis among UK Biobank 342,973 participants (2,034 incident RA events), additionally adjusted the latitude of participants' residence in the sensitivity analysis.

Air pollution score[a]	case / *n*	HR (95% CI) of incident RA			*P* for trend
		Model 1[b]	Model 2[c]	Model 3[d]	
Per SD increment	-	1.12 (1.07, 1.16)	1.06 (1.01, 1.10)	1.06 (1.01, 1.10)	
Quartile 1	454 / 85,743	Ref	Ref	Ref	0.000053
Quartile 2	490 / 85,743	1.09 (0.96, 1.24)	1.00 (0.88, 1.13)	1.00 (0.88, 1.13)	
Quartile 3	514 / 85,743	1.17 (1.03, 1.33)	1.01 (0.89, 1.15)	1.01 (0.89, 1.15)	
Quartile 4	576 / 85,744	1.34 (1.18, 1.51)	1.14 (1.00, 1.29)	1.14 (1.01, 1.30)	

HR, hazard ratio; CI, confidence interval.

[a]Air pollution score ranges: quartile 1: (48.54, 63.35); quartile 2: (63.36, 69.94); quartile 3: (69.95, 76.43); 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, smoking status, body mass index, alcohol consumption, sedentary time, physical activity duration, healthy diet score, and latitude of participants' residence.

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

Supplemental Table S13 Association between air pollution score and incident rheumatoid arthritis among UK Biobank 292,987 participants (1,745 incident RA events), after limiting participants living in the current address for at least five years in the sensitivity analysis.

Air pollution score[a]	case / *n*	HR (95% CI) of incident RA			*P* for trend
		Model 1[b]	Model 2[c]	Model 3[e]	
Per SD increment	-	1.11 (1.07, 1.17)	1.05 (1.00, 1.10)	1.05 (1.00, 1.10)	
Quartile 1	385 / 72,413	Ref	Ref	Ref	0.000064
Quartile 2	422 / 73,997	1.08 (0.94, 1.25)	1.00 (0.87, 1.15)	1.00 (0.87, 1.15)	
Quartile 3	443 / 74,033	1.17 (1.02, 1.34)	1.01 (0.88, 1.16)	1.01 (0.88, 1.16)	
Quartile 4	495 / 72,544	1.35 (1.18, 1.54)	1.15 (1.00, 1.32)	1.15 (1.00, 1.32)	

HR, hazard ratio; CI, confidence interval.

[a]Air pollution score ranges: quartile 1: (48.54, 63.35); quartile 2: (63.36, 69.94); quartile 3: (69.95, 76.43); quartile 4: (76.44, 202.00). Mean (SD) of the air pollution score is 70.60 (11.00).

[b]Adjusted for age and sex.

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

Supplemental Table S14 Association between air pollution score and incident rheumatoid arthritis among UK Biobank 313,466 participants (1,701 incident RA events), after excluding participants who used NSAIDs previously or at baseline in the sensitivity analysis.

Air pollution score[a]	case / *n*	HR (95% CI) of incident RA			*P* for trend
		Model 1[b]	Model 2[c]	Model 3[e]	
Per SD increment	-	1.11 (1.06, 1.16)	1.05 (1.01, 1.10)	1.05 (1.00, 1.10)	
Quartile 1	390 / 77,733	Ref	Ref	Ref	0.024
Quartile 2	406 / 78,046	1.09 (0.96, 1.23)	1.00 (0.88, 1.14)	1.00 (0.88, 1.14)	
Quartile 3	425 / 78,410	1.16 (1.02, 1.32)	1.01 (0.89, 1.15)	1.01 (0.89, 1.15)	
Quartile 4	480 / 79,157	1.32 (1.16, 1.49)	1.13 (1.00, 1.28)	1.13 (1.00, 1.28)	

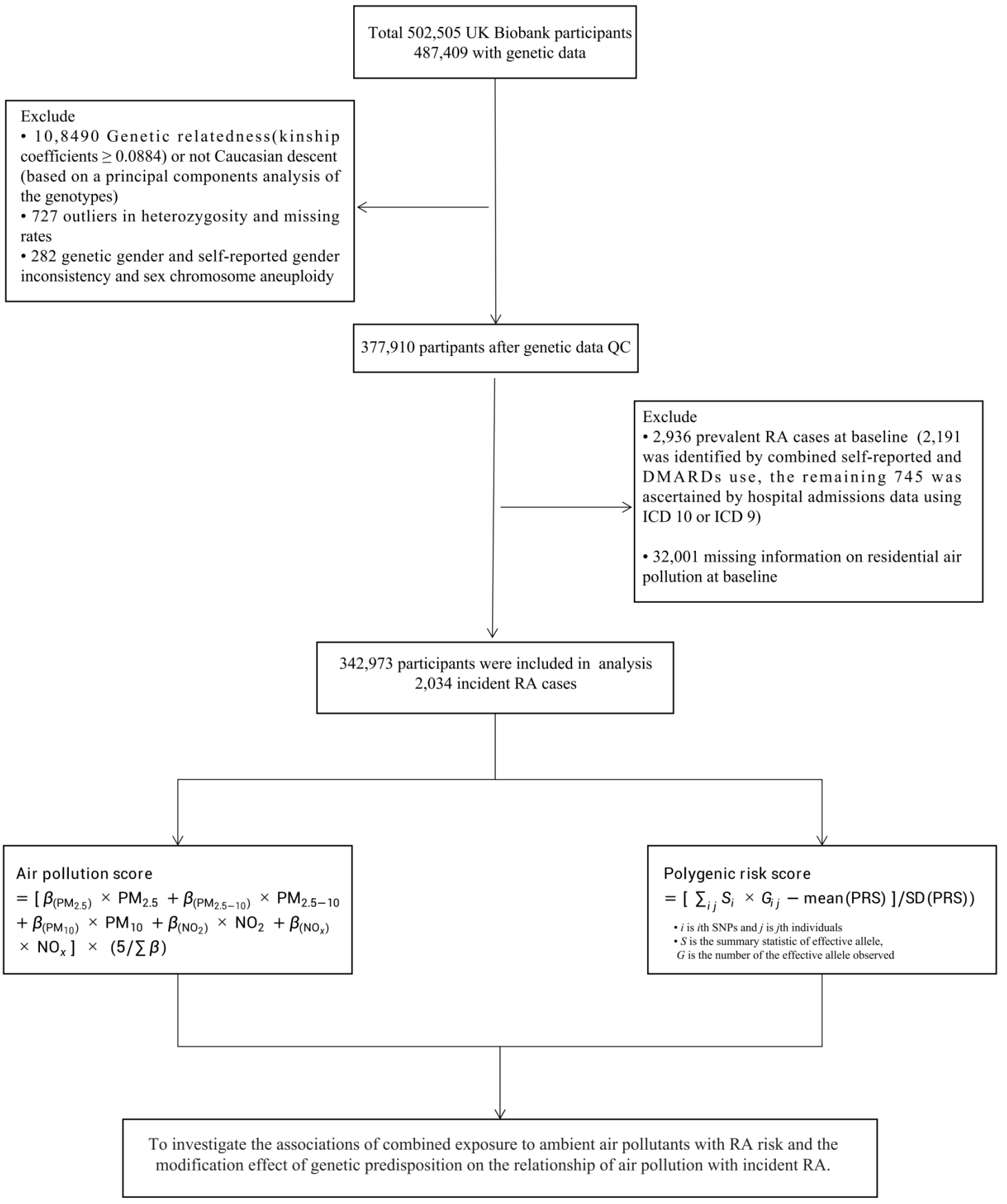
HR, hazard ratio; CI, confidence interval.

[a]Air pollution score ranges: quartile 1: (48.62, 63.35); quartile 2: (63.36, 69.95); quartile 3: (69.96, 76.43); quartile 4: (76.44, 202.00). Mean (SD) of the air pollution score is 70.71 (11.15).

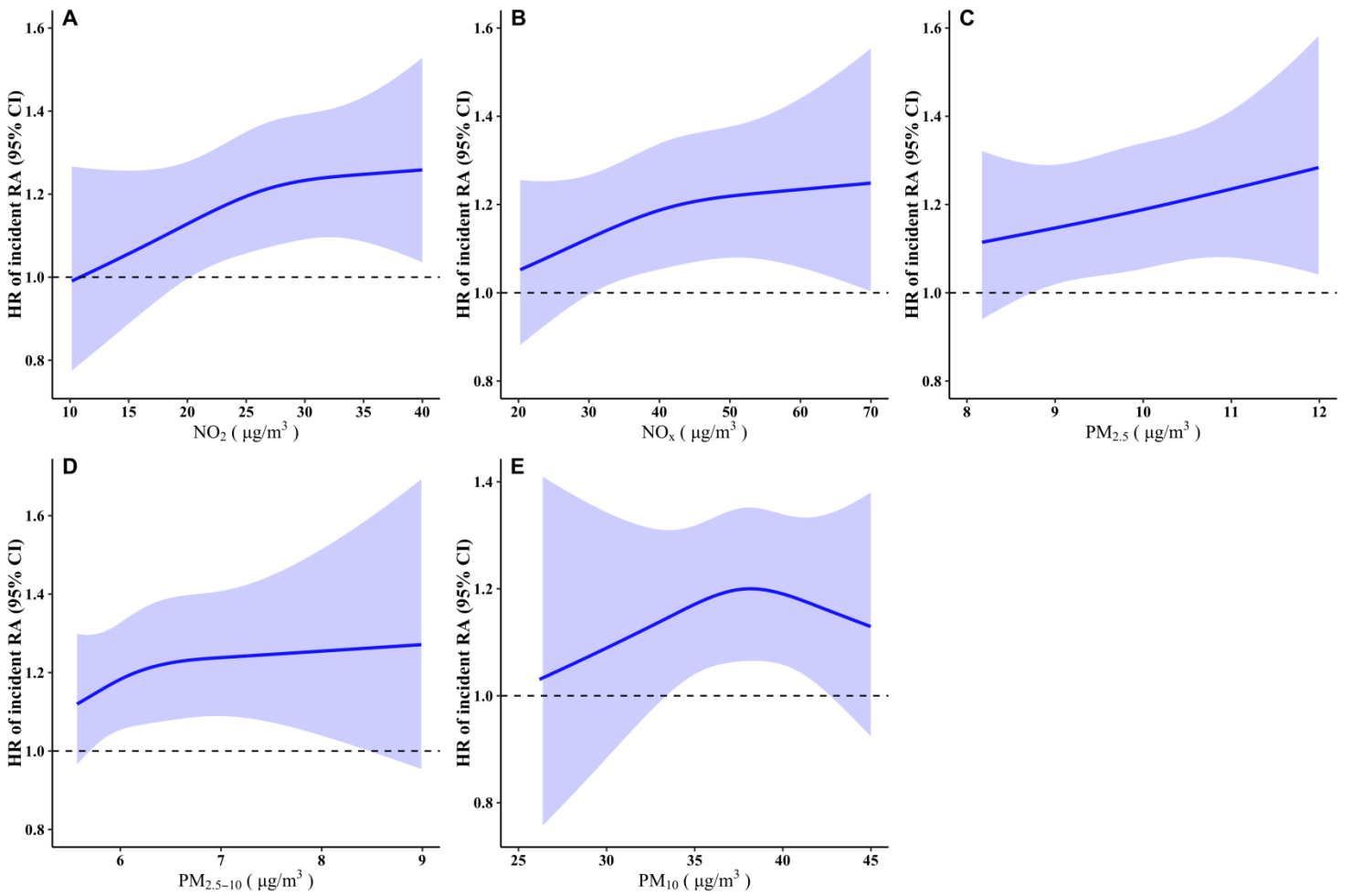
[b]Adjusted for age and sex.

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



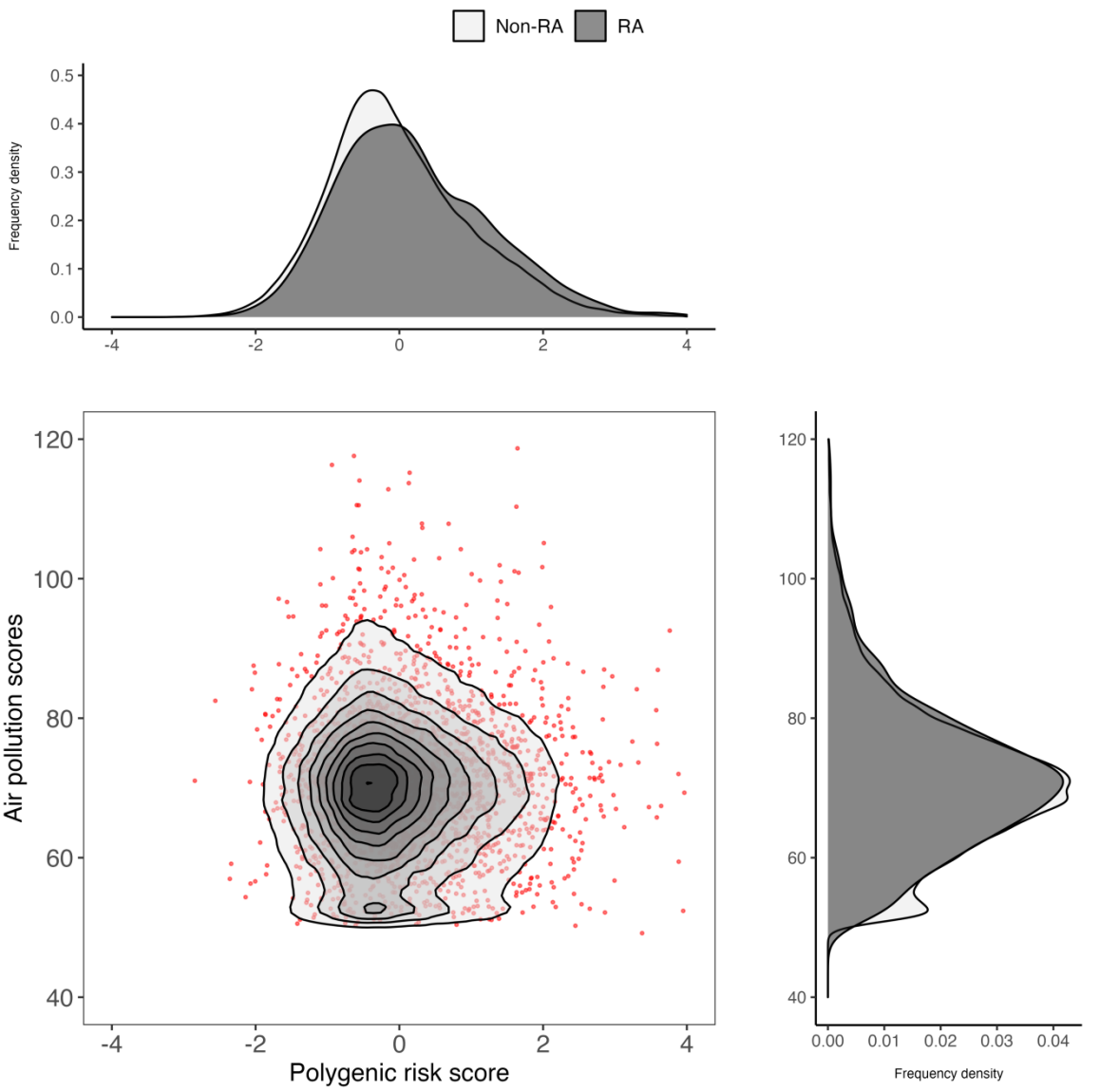
Supplemental Figure S1 Overview of design and population selection in the UK Biobank study of the association of air pollution and genetic risk with rheumatoid arthritis incidence.



Supplemental Figure S2 Exposure-response relationships of each air pollutant with incident rheumatoid arthritis among UK Biobank 342,973 participants (2,034 incident RA events).

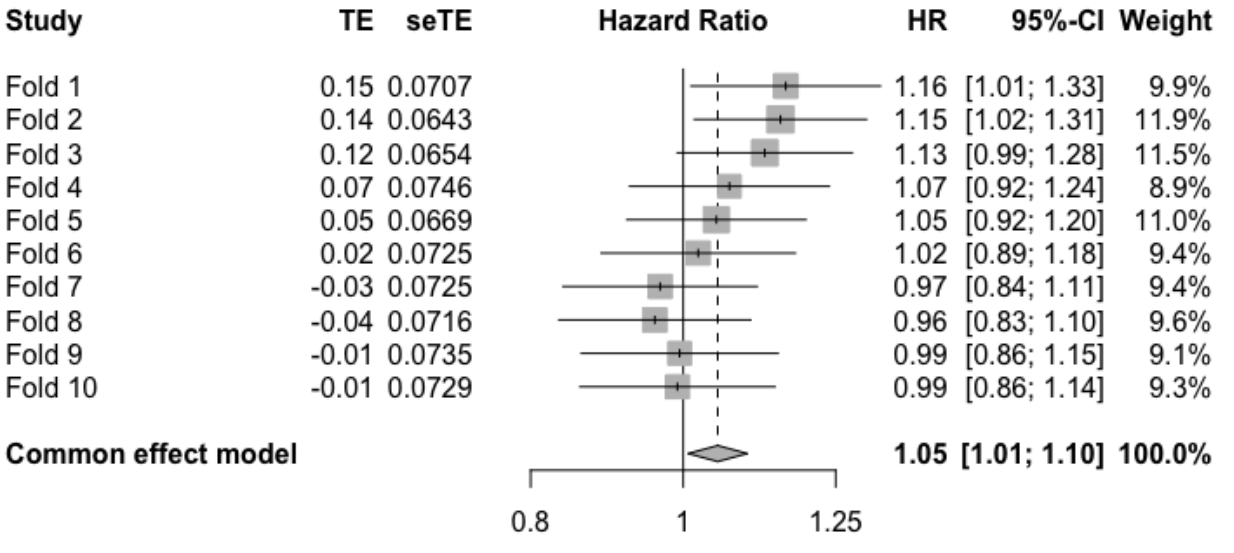
The exposure-response relationship of concentrations for each air pollutants with hazard ratios for incident RA was assessed using restricted cubic spline analysis with 3 knots. The source data used to generate the figure was in Excel Table S2

(A)NO2 (B)NOx (C)PM2.5 (D)PM2.5-10 (E)PM10 ; Both were adjusted for age, sex, UK Biobank assessment center, household income education, smoking status, body mass index, alcohol consumption, sedentary time, physical activity duration, healthy diet score, polygenic risk score, first ten genetic principal components and genotyping batch.

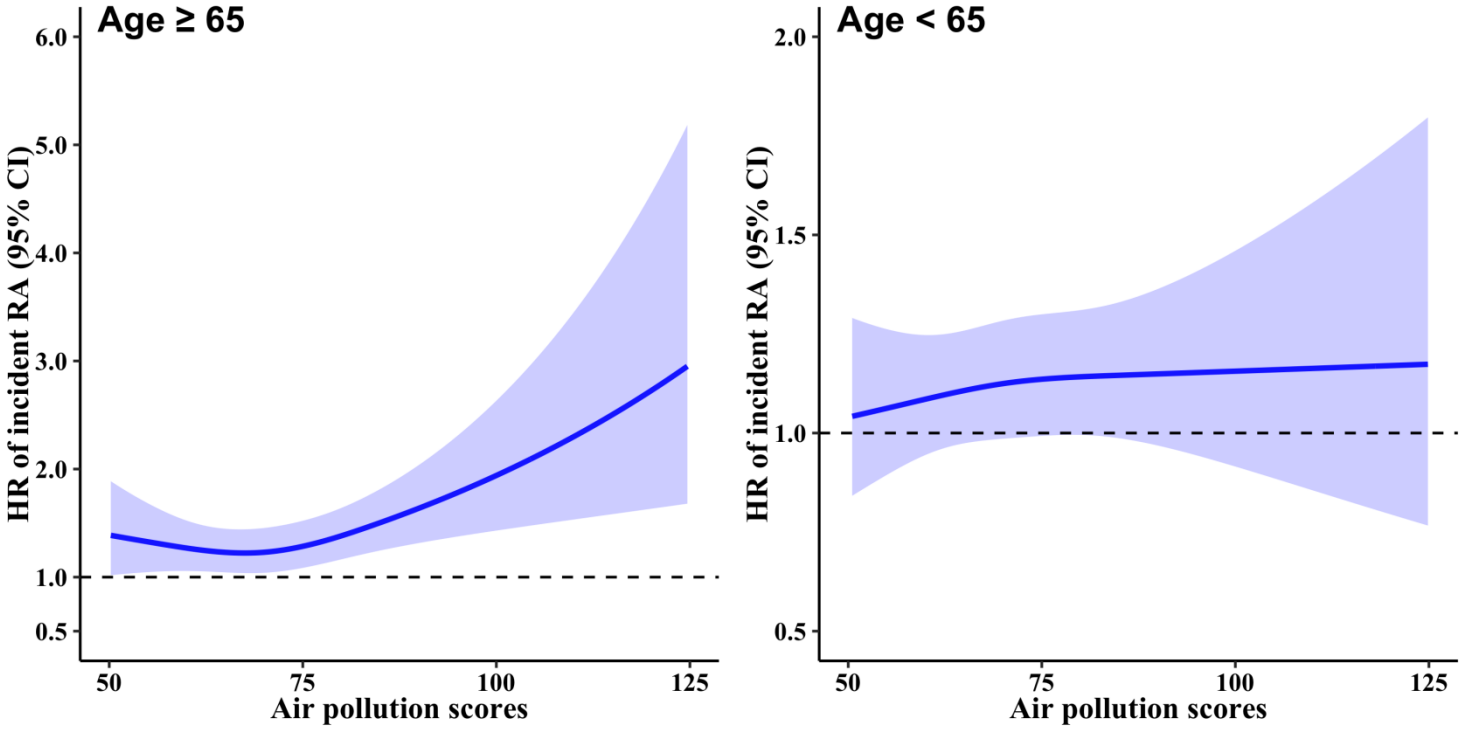


Supplemental Figure S3 Joint distribution of air pollution scores and polygenic risk scores between RA and non-RA among UK Biobank 342,973 participants (2,034 incident RA events).

The marginal density plot represents the distribution of air pollution score and polygenic risk score according to incident RA status. The ridge density plot represents the 95% range for the bivariate distribution. Incident RA (red points) are also shown. The source data used to generate the figure was in Excel Table S3



Supplemental Figure S4 Validation of association between air pollution score and incident rheumatoid arthritis by 10-fold cross-validation.

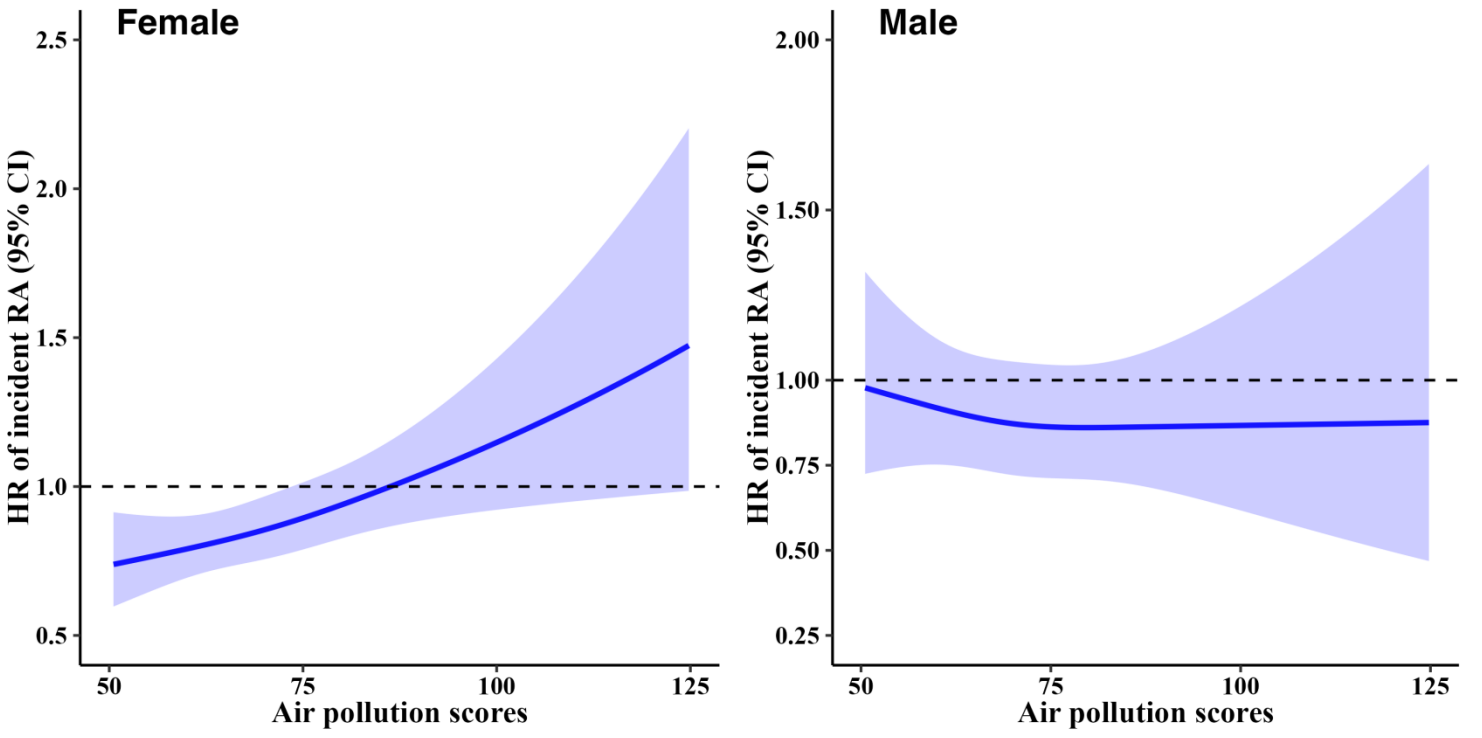


Supplemental Figure S5 Exposure-response relationship of air pollution score with incident rheumatoid arthritis according to age.

HR, hazard ratio; CI, confidence interval; RA, rheumatoid arthritis.

The exposure-response relationship of air pollution scores with hazard ratios for incident RA was assessed using restricted cubic spline analysis with 3 knots. The source data used to generate the figure was in Excel Table S4.

Both were adjusted for age, sex, UK Biobank assessment center, household income education, smoking status, body mass index, alcohol consumption, sedentary time, physical activity duration, healthy diet score, polygenic risk score, first ten genetic principal components and genotyping batch.



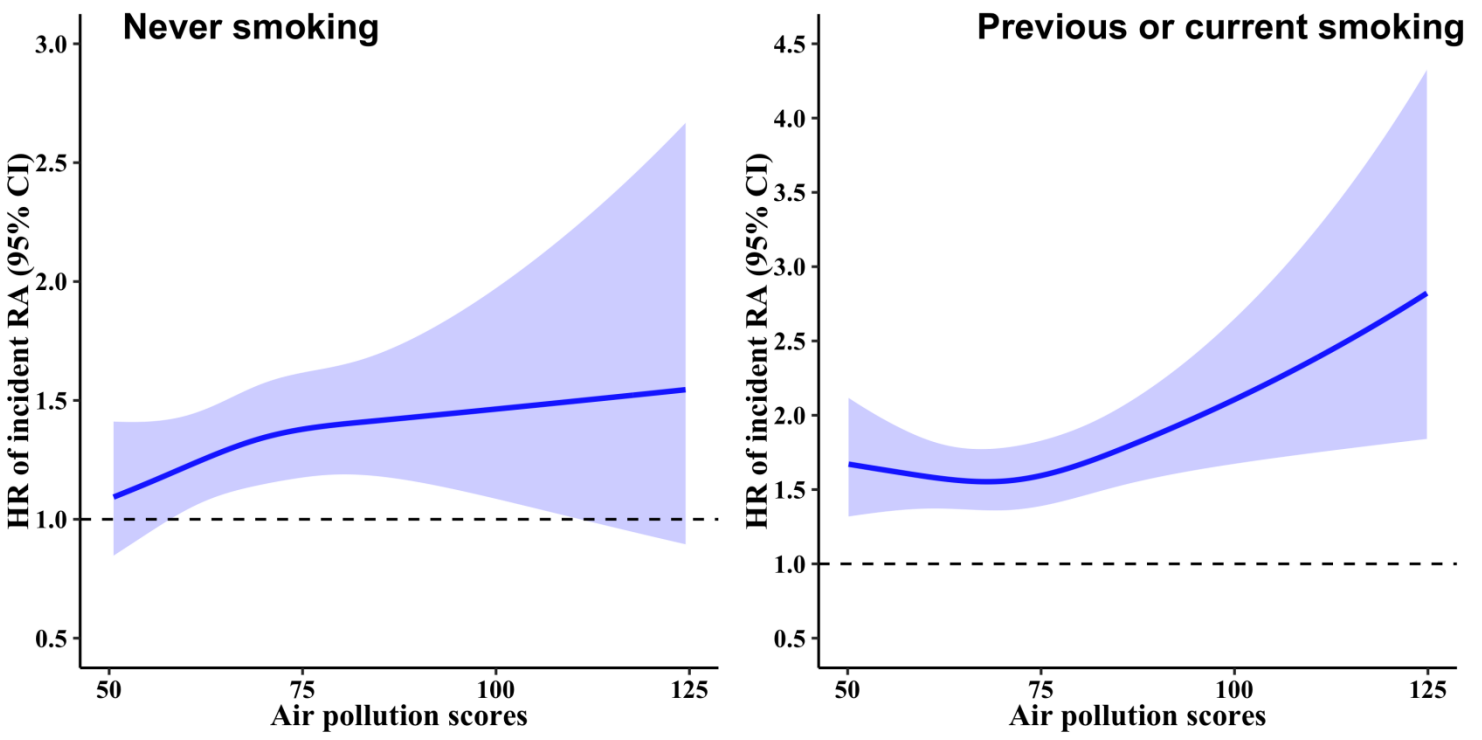
Supplemental Figure S6 Exposure-response relationship of air pollution score with incident rheumatoid arthritis according to sex.

HR, hazard ratio; CI, confidence interval; RA, rheumatoid arthritis.

The exposure-response relationship of air pollution scores with hazard ratios for incident RA was assessed using restricted cubic spline analysis with 3 knots. The source data used to generate the figure was in Excel Table S5.

Both were adjusted for age, UK Biobank assessment center, household income education, smoking status, body mass index, alcohol consumption, sedentary time, physical activity duration, healthy diet score, polygenic risk score, first ten genetic principal components, and genotyping batch.

Additionally adjusted for menopause and hormone replacement therapy in the female subgroup.

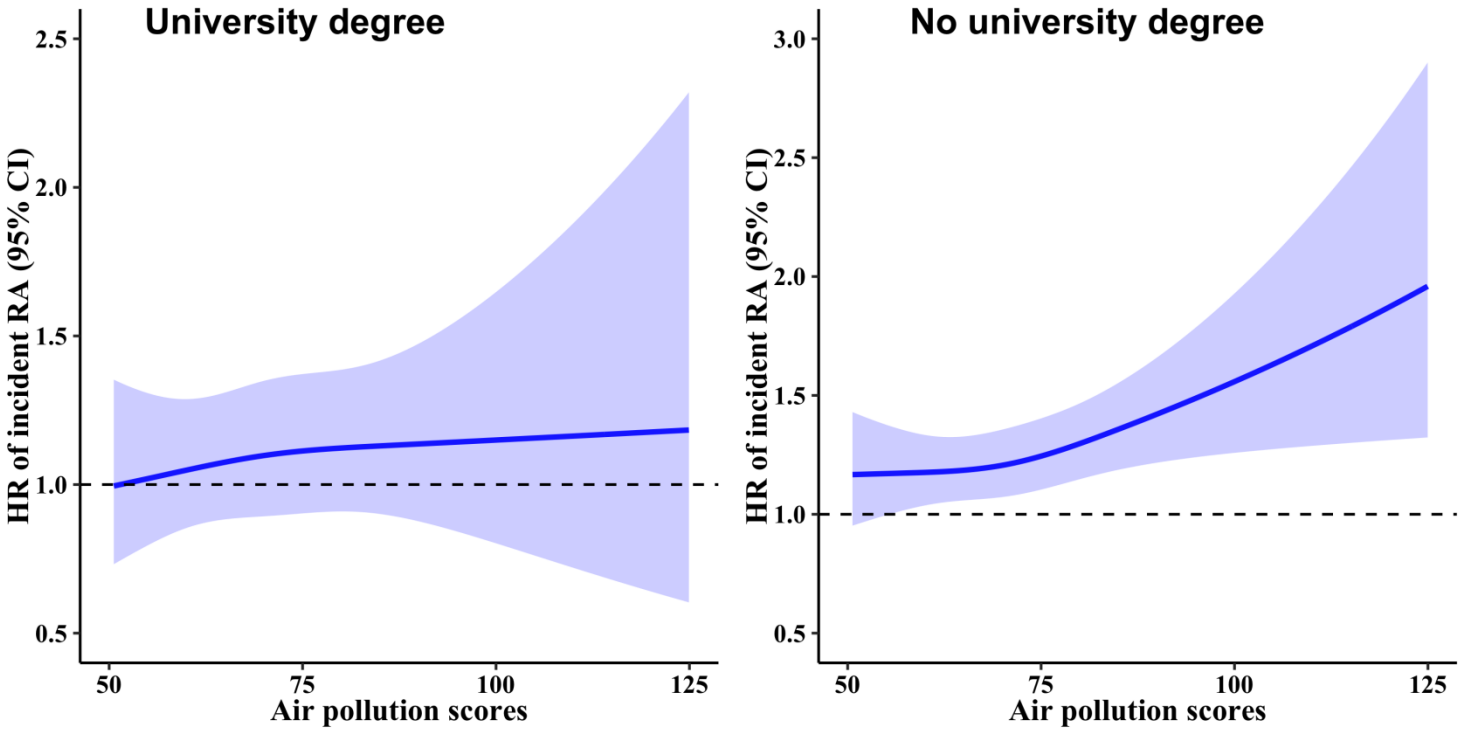


Supplemental Figure S7 Exposure-response relationship of air pollution score with incident rheumatoid arthritis according to smoking status.

HR, hazard ratio; CI, confidence interval; RA, rheumatoid arthritis.

The exposure-response relationship of air pollution scores with hazard ratios for incident RA was assessed using restricted cubic spline analysis with 3 knots. The source data used to generate the figure was in Excel Table S6.

Both were adjusted for age, sex, UK Biobank assessment center, household income education, body mass index, alcohol consumption, sedentary time, physical activity duration, healthy diet score, polygenic risk score, first ten genetic principal components and genotyping batch.

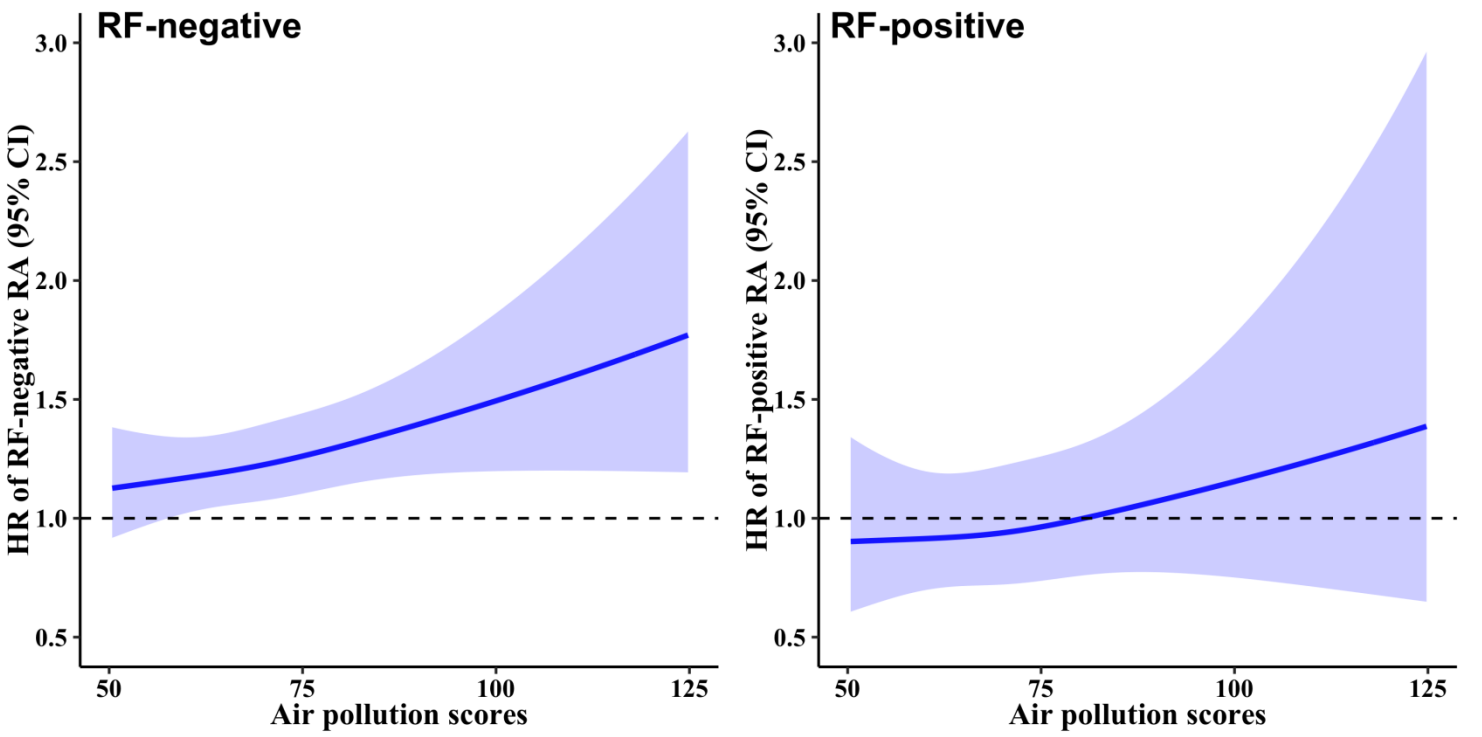


Supplemental Figure S8 Exposure-response relationship of air pollution score with incident rheumatoid arthritis according to education level.

HR, hazard ratio; CI, confidence interval; RA, rheumatoid arthritis.

The exposure-response relationship of air pollution scores with hazard ratios for incident RA was assessed using restricted cubic spline analysis with 3 knots. The source data used to generate the figure was in Excel Table S7.

Both were adjusted for age, sex, UK Biobank assessment center, household income, smoking status, body mass index, alcohol consumption, sedentary time, physical activity duration, healthy diet score, polygenic risk score, first ten genetic principal components, and genotyping batch.



Supplemental Figure S9 Exposure-response relationship of air pollution score with rheumatoid arthritis incidence according to rheumatoid factor (RF) status.

HR, hazard ratio; CI, confidence interval; RA, rheumatoid arthritis.

The exposure-response relationship of air pollution scores with hazard ratios for incident RA was assessed using restricted cubic spline analysis with 3 knots. The source data used to generate the figure was in Excel Table S8.

Both were adjusted for age, sex, UK Biobank assessment center, household income education, smoking status, body mass index, alcohol consumption, sedentary time, physical activity duration, healthy diet score, polygenic risk score, first ten genetic principal components and genotyping batch.