Air pollution, residential greenspace, and the risk of incident immune thrombocytopenic purpura: a prospective cohort study of 356,482 participants

Peiyang Luo,[1*] Feifan Wang,[1*] Jiacheng Ying,[1] Ke Liu,[2] Baojie Hua,[2] Shuhui Chen,[2] Jiayu Li,[2] Xiaohui Sun,[2] Ding Ye,[2] Baodong Ye,[3] Jinyi Tong,[1#] Keding Shao[3#] and Yingying Mao[2#]

[1]The Fourth School of Clinical Medicine, Zhejiang Chinese Medical University, Hangzhou First People's Hospital; [2]School of Public Health, Zhejiang Chinese Medical University and [3]The First Affiliated Hospital of Zhejiang Chinese Medical University (Zhejiang Provincial Hospital of Chinese Medicine), Hangzhou, China

*PL and FW contributed equally as first authors.
#JT, KS and YM contributed equally as senior authors.

Supplementary Appendix for

Air pollution, residential greenspace and the risk of incident immune thrombocytopenic purpura: a prospective cohort study of 356,482 participants

Peiyang Luo[1#], Feifan Wang[1#], Jiacheng Ying[1], Ke Liu[2], Baojie Hua[2], Shuhui Chen[2], Jiayu Li[2], Xiaohui Sun[2], Ding Ye[2], Baodong Ye[3], Jinyi Tong[1*], Keding Shao[3*], Yingying Mao[2*]

Contents

Supplemental methods

The land use regression (LUR) models

The detail information of residential greenspace assessment

The detail information of additive interaction

Modeling Mediation effects of residential greenspace and air pollutants on ITP risk

Reference

Supplemental Table 1. Details information for covariates.

Supplementary Table 2. The associations between residential greenspace buffer at 1000m and the risk of incident ITP.

Supplementary Table 3. The associations between five air pollutants and residential greenspace buffer at 300m with per 10 unit increase and the risk of ITP.

Supplementary Table 4. Sensitivity analysis after excluding incident ITP within the first year of entering the cohort.

Supplementary Table 5. Sensitivity analysis restricting participants with over 10 years in the same residence.

Supplementary Table 6. Sensitivity analysis after excluding individuals diagnosed with other types of purpura and hemorrhagic conditions before baseline.

Supplementary Table 7. The associations of air pollution, residential greenspace buffer at 300m with the risk of ITP according to quartiles of pollutants and residential greenspace buffer at 300m.

Supplementary Figure 1. The flow chart of the study design.

Supplementary Figure 2. Schoenfeld residual plots for air pollution and residential greenspace.

Supplementary Figure 3. Stratified analyses of the associations of air pollution, residential greenspace buffer at 300 m with the risk of incident ITP.

Supplemental methods

1.The land use regression (LUR) models

The LUR models, developed as part of the European Study of Cohorts for Air Pollution Effects (ESCAPE), utilized the geocoded residential addresses of participants at baseline, incorporating a series of predictor variables (such as traffic intensity, population density, land use and topography) provided by the geographic information system (GIS) to estimate the spatial variation of air pollutant concentrations (http://www.escapeproject.eu/) [1, 2]. Additionally, leave-one-out cross-validation showed that the models exhibit good performance for $PM_{2.5}$, PM_{10}, NO_2 and NO_x with cross-validation R^2 values of 77%, 88%, 87%, and 88%, respectively, and moderate performance for PM_{coarse} with a cross-validation R^2 of 57%.

2. The detail information of residential greenspace assessment

The Generalized Land Use Database (GLUD), issued by the UK Government's Communities and Local Government Department (https://www.gov.uk/government/statistics), provided land use distribution information at the 2001 census output area (COA) level and has been previously used in studies [3, 4]. Each polygon representing a home location was assigned an area-weighted mean of the land use percentage coverage intersecting the home location buffer. The percentage of residential greenspace (categorized as "Greenspace" within the residential location buffer) was then calculated as a proportion of all land use types within 300m and 1000m buffers. These buffer distances were chosen to represent nearby and wide-area residential greenspace relative to participants' household positions. Detailed information about the measurement is available at

https://biobank.ctsu.ox.ac.uk/showcase/ukb/docs/App15374Docs.pdf.

3.The detail information of additive interaction

Additive interaction was evaluated with relative excess risk due to interaction (RERI) and the attributable proportion due to interaction (AP), and 95% confidence intervals (CIs) included 0 indicated no additive interactions. RERI was calculated via the formula $RERI = RR_{11} - RR_{10} - RR_{01} + 1$, whereas RR_{11} denoted the relative risk of an individual exposed to both factors (air pollution and residential greenspace buffer at 300m), and RR_{10} and RR_{01} represented the relative risks for individuals exposed to either air pollution or residential greenspace, respectively. The estimation of 95% CIs was performed using the delta method outlined by Hosmer and Lemeshow [5].

4.Modeling Mediation effects of residential greenspace and air pollutants on ITP risk

We initially employed a linear model, with residential greenspace as the independent variable and air pollutants as the dependent variable (mediation model). Then we fitted a Cox proportional hazard model, treating residential greenspace, air pollutants and the interaction term as independent variables, with incident ITP as the dependent variable (outcome model) [6, 7]. All models were adjusted for age, sex, alcohol drinking status, BMI, household income, physical activity, smoking status, education level, ethnicity.

Reference:

1. Beelen, R., et al., Development of NO2 and NOx land use regression models for estimating air pollution exposure in 36 study areas in Europe – The ESCAPE project. Atmospheric Environment, 2013. 72: p. 10-23.

2. Eeftens, M., et al., Development of Land Use Regression Models for PM2.5, PM2.5 Absorbance, PM10 and PMcoarse in 20 European Study Areas; Results of the ESCAPE Project. Environmental Science & Technology, 2012. 46(20): p. 11195-11205.

3. Alcock, I., et al., Longitudinal effects on mental health of moving to greener and less green urban areas. Environ Sci Technol, 2014. 48(2): p. 1247-55.

4. White, M.P., et al., Would You Be Happier Living in a Greener Urban Area? A Fixed-Effects Analysis of Panel Data. Psychological Science, 2013. 24(6): p. 920-928.

5. Hosmer, D.W. and S. Lemeshow, Confidence interval estimation of interaction. Epidemiology, 1992. 3(5): p. 452-6.

6. Huang, Y.T. and H.I. Yang, Causal Mediation Analysis of Survival Outcome with Multiple Mediators. Epidemiology, 2017. 28(3): p. 370-378.

7. Valeri, L. and T.J. VanderWeele, SAS macro for causal mediation analysis with survival data. Epidemiology, 2015. 26(2): p. e23-4.

Supplemental Table 1. Details information for covariates.

Covariates		Field ID	Definition
Sociodemographic factors	Age	21022	Continuous variable. Data from the UK biobank baseline touchscreen questionnaire.
	Sex	31	Data from the UK biobank baseline touchscreen questionnaire. Participants were categorized as male or female.
	Ethnicity	21000	Data from the UK biobank baseline touchscreen questionnaire. Participants were categorized into two groups: White and Others.
	Education level	6138	Participants from the UK Biobank baseline touchscreen questionnaire were classified into seven groups based on their reported educational qualifications: (1) College or University degree; (2) A levels/AS levels or equivalent; (3) O levels/GCSEs or equivalent; (4) CSEs or equivalent; (5) NVQ, HND, HNC or equivalent; (6) Other professional qualifications (e.g., nursing, teaching); and (7) None of the above. For analysis, groups 1, 5, and 6 were classified as 'higher vocational qualifications or more', while the remaining groups were categorized as 'lower vocational qualifications or less'.
Lifestyle factors	Smoking status	20116	Data from the UK biobank baseline touchscreen questionnaire. Participants were categorized into two groups based on their responses of questionnaires: Never smoking and Previous/current smoking.
	Alcohol drinking status	20117	Data from the UK biobank baseline touchscreen questionnaire. Participants were categorized into two groups based on their responses of questionnaires: Never drinking and Previous/current drinking.
	Physical activity	884, 894, 904, 914	Data from the UK biobank baseline touchscreen questionnaire. Regular physical activity was defined as meeting one of the following criteria: (i) Frequency: vigorous activity once or moderate physical activity at least 5 days per week; (ii) Time: vigorous activity for at least 75 minutes or moderate activity for 150 minutes per week.
	Body mass index	21001	BMI, measured in kg/m^2, was calculated based on height and weight measured at the initial assessment center visit. Continuous variable.
Household characteristics	Average total before-tax household income	738	Data from the UK biobank baseline touchscreen questionnaire. We classified participants into two groups: < £31000 and >= £31000.

Supplementary Table 2. The associations between residential greenspace buffer at 1000m and the risk of incident ITP.

Exposure	Total	Events/Person years	Model a		Model b	
			HR (95% CI)	*P* value	HR (95% CI)	*P* value
Residential greenspace buffer at 1000m						
Per IQR (32.83%) increment	356482	500/4711602	0.74 (0.65,0.86)	3.70×10^{-5}	0.76 (0.66,0.88)	2.54×10^{-4}
Per 10% increment	356482	500/4711602	0.91 (0.88,0.95)	3.70×10^{-5}	0.92 (0.88,0.96)	2.54×10^{-4}
Q1 [4.49%, 27.60%]	89121	146/1162739	Ref		Ref	
Q2 (27.60%, 42.00%]	89128	120/1176779	0.78 (0.61,0.99)	0.039	0.76 (0.60,0.97)	0.030
Q3 (42.00%, 60.40%]	89118	137/1184563	0.85 (0.68,1.08)	0.182	0.85 (0.67,1.08)	0.182
Q4 (60.40%, 99.19%]	89115	97/1187520	0.59 (0.46,0.77)	7.47×10^{-5}	0.62 (0.48,0.81)	3.84×10^{-4}

Model a: age and sex.

Model b: age, sex, ethnicity, BMI, education level, household income, smoking status, alcohol drinking status, and physical activity.

Abbreviations: BMI: body mass index; CI: confidence intervals; HR: hazard ratio; IQR: interquartile range; ITP: Immune thrombocytopenic purpura.

Supplementary Table 3. The associations between five air pollutants and residential greenspace buffer at 300m with per 10 unit increase and the risk of ITP.

Exposure	Total	Events/Person years	Model a		Model b	
			HR (95%CI)	*P* value	HR (95%CI)	*P* value
$PM_{2.5}$	356482	500/4711602	4.29 (1.96,9.40)	2.78×10^{-4}	3.03 (1.35,6.80)	0.007
PM_{coarse}	356482	500/4711602	0.90 (0.33,2.44)	0.832	0.81 (0.30,2.21)	0.679
PM_{10}	356482	500/4711602	1.43 (0.91,2.26)	0.124	1.28 (0.80,2.03)	0.300
NO_2	356482	500/4711602	1.27 (1.15,1.42)	5.37×10^{-6}	1.23 (1.10,1.37)	1.83×10^{-4}
NO_X	356482	500/4711602	1.09 (1.04,1.14)	5.42×10^{-4}	1.07 (1.02,1.12)	0.011
Residential greenspace buffer at 300m	356482	500/4711602	0.91 (0.88,0.95)	1.26×10^{-5}	0.92 (0.88,0.96)	7.96×10^{-5}

Model a: age and sex.

Model b: age, sex, ethnicity, BMI, education level, household income, smoking status, alcohol drinking status, and physical activity.

Abbreviations: BMI: body mass index; CI: confidence intervals; HR: hazard ratio; IQR: interquartile range; ITP: Immune thrombocytopenic purpura; NO_2: nitrogen dioxide; NO_x: nitrogen oxides; $PM_{2.5}$: fine particulate matter; PM_{coarse}: coarse particulate matter with aerodynamic diameters ranging between 2.5 and 10 μm; PM_{10}: particulate matter with diameters of less than 10-micrometer.

Supplementary Table 4. Sensitivity analysis after excluding incident ITP within the first year of entering the cohort.

Exposure	Total	Events/Person years	Model a		Model b	
			HR (95% CI)	*P* value	HR (95% CI)	*P* value
$PM_{2.5}$	355828	478/4711222	1.20 (1.09,1.33)	4.05×10^{-4}	1.15 (1.04,1.28)	0.007
PM_{coarse}	355828	478/4711222	0.99 (0.91,1.07)	0.731	0.98 (0.90,1.06)	0.604
PM_{10}	355828	478/4711222	1.06 (0.98,1.15)	0.145	1.04 (0.96,1.13)	0.311
NO_2	355828	478/4711222	1.26 (1.14,1.41)	1.57×10^{-5}	1.23 (1.10,1.37)	3.03×10^{-4}
NO_x	355828	478/4711222	1.14 (1.05,1.24)	0.001	1.11 (1.02,1.21)	0.016
Residential greenspace buffer at 300m	355828	478/4711222	0.74 (0.65,0.85)	1.22×10^{-5}	0.76 (0.66,0.87)	5.99×10^{-5}

Model a: age and sex.

Model b: age, sex, ethnicity, BMI, education level, household income, smoking status, alcohol drinking status, and physical activity.

Abbreviations: BMI: body mass index; CI: confidence intervals; HR: hazard ratio; IQR: interquartile range; ITP: Immune thrombocytopenic purpura; NO_2: nitrogen dioxide; NO_x: nitrogen oxides; $PM_{2.5}$: fine particulate matter; PM_{coarse}: coarse particulate matter with aerodynamic diameters ranging between 2.5 and 10 μm; PM_{10}: particulate matter with diameters of less than 10-micrometer.

Supplementary Table 5. Sensitivity analysis restricting participants with over 10 years in the same residence.

Exposure	Total	Events/Person years	Model a		Model b	
			HR (95% CI)	*P* value	HR (95% CI)	*P* value
$PM_{2.5}$	241534	367/3183216	1.26 (1.12,1.41)	1.40×10^{-4}	1.19 (1.06,1.34)	0.005
PM_{coarse}	241534	367/3183216	1.01 (0.92,1.10)	0.058	1.00 (0.91,1.09)	0.174
PM_{10}	241534	367/3183216	1.09 (1.00,1.20)	0.831	1.07 (0.97,1.17)	0.978
NO_2	241534	367/3183216	1.35 (1.20,1.52)	1.18×10^{-6}	1.29 (1.13,1.46)	1.06×10^{-4}
NO_x	241534	367/3183216	1.20 (1.09,1.31)	9.05×10^{-5}	1.15 (1.05,1.27)	0.004
Residential greenspace buffer at 300m	241534	367/3183216	0.72 (0.62,0.84)	3.95×10^{-5}	0.75 (0.64,0.87)	2.95×10^{-4}

Model a: age and sex.

Model b: age, sex, ethnicity, BMI, education level, household income, smoking status, alcohol drinking status, physical activity.

Abbreviations: BMI: body mass index; CI: confidence intervals; HR: hazard ratio; IQR: interquartile range; ITP: Immune thrombocytopenic purpura; NO_2: nitrogen dioxide; NO_x: nitrogen oxides; $PM_{2.5}$: fine particulate matter; PM_{coarse}: coarse particulate matter with aerodynamic diameters ranging between 2.5 and 10 μm; PM_{10}: particulate matter with diameters of less than 10-micrometer.

Supplementary Table 6. Sensitivity analysis after excluding individuals diagnosed with other types of purpura and hemorrhagic conditions before baseline.

Exposure	Total	Events/Person years	Model a		Model b	
			HR (95% CI)	*P* value	HR (95% CI)	*P* value
PM$_{2.5}$	356185	472/4708290	1.19 (1.08,1.32)	8.49×10^{-4}	1.14 (1.03,1.27)	0.013
PM$_{coarse}$	356185	472/4708290	1.00 (0.92,1.08)	0.948	0.99 (0.91,1.07)	0.798
PM$_{10}$	356185	472/4708290	1.07 (0.98,1.16)	0.112	1.05 (0.97,1.14)	0.258
NO$_2$	356185	472/4708290	1.26 (1.13,1.40)	2.19×10^{-5}	1.22 (1.09,1.37)	4.42×10^{-4}
NO$_x$	356185	472/4708290	1.14 (1.05,1.24)	0.002	1.11 (1.02,1.21)	0.019
Residential greenspace buffer at 300m	356185	472/4708290	0.76 (0.66,0.87)	4.25×10^{-5}	0.77 (0.67,0.89)	2.05×10^{-4}

Model a: age and sex.

Model b: age, sex, ethnicity, BMI, education level, household income, smoking status, alcohol drinking status, and physical activity.

Abbreviations: BMI: body mass index; CI: confidence intervals; HR: hazard ratio; IQR: interquartile range; ITP: Immune thrombocytopenic purpura; NO$_2$: nitrogen dioxide; NO$_x$: nitrogen oxides; PM$_{2.5}$: fine particulate matter; PM$_{coarse}$: coarse particulate matter with aerodynamic diameters ranging between 2.5 and 10 µm; PM$_{10}$: particulate matter with diameters of less than 10-micrometer.

Supplemental Table 7. The associations of air pollution, residential greenspace buffer at 300m with the risk of ITP according to quartiles of pollutants and residential greenspace buffer at 300m.

Exposure	Total	Events/Person years	Model a HR (95%CI)	Model a P value	Model b HR (95%CI)	Model b P value
$PM_{2.5}$ concentration						
Q1 [8,17μg/m³, 9.28μg/m³]	89845	98/1192383	Ref		Ref	
Q2 (9.28μg/m³, 9.92μg/m³]	89026	128/1172721	1.35 (1.04,1.76)	0.024	1.29 (0.99,1.68)	0.057
Q3 (9.92μg/m³, 10.55μg/m³]	89001	132/1172384	1.43 (1.10,1.86)	0.007	1.33 (1.02,1.73)	0.034
Q4 (10.55μg/m³, 21.31μg/m³]	88610	142/1174114	1.59 (1.23,2.06)	$3.86×10^{-4}$	1.44 (1.11,1.87)	0.006
P for trend				$4.77×10^{-4}$		0.009
PM_{coarse} concentration						
Q1 [5.57μg/m³, 5.84μg/m³]	90907	118/1211340	Ref		Ref	
Q2 (5.84μg/m³, 6.10μg/m³]	87662	139/1161587	1.24 (0.97,1.59)	0.082	1.21 (0.95,1.55)	0.126
Q3 (6.10μg/m³, 6.62μg/m³]	89426	122/1172139	1.10 (0.85,1.42)	0.467	1.06 (0.82,1.37)	0.638
Q4 (6.62μg/m³, 12.82μg/m³]	88487	121/1166535	1.10 (0.85,1.42)	0.466	1.06 (0.82,1.37)	0.655
P for trend				0.704		0.930
PM_{10} concentration						
Q1 [11.78μg/m³, 15.20μg/m³]	89660	97/1192342	Ref		Ref	
Q2 (15.20μg/m³, 16.00μg/m³]	89576	145/1189950	1.52 (1.17,1.96)	0.001	1.46 (1.13,1.89)	0.004
Q3 (16.00μg/m³, 17.00μg/m³]	88471	132/1161497	1.45 (1.12,1.89)	0.005	1.38 (1.06,1.79)	0.017
Q4 (17.00μg/m³, 31.39μg/m³]	88775	126/1167813	1.39 (1.06,1.81)	0.016	1.32 (1.01,1.72)	0.043
P for trend				0.037		0.098
NO_2 concentration						
Q1 [12.93μg/m³, 21.30μg/m³]	89283	99/1190163	Ref		Ref	
Q2 (21.30μg/m³, 26.00μg/m³]	88983	111/1178463	1.15 (0.88,1.51)	0.313	1.09 (0.83,1.43)	0.534

|			Model a		Model b	
Q3 (26.00μg/m^3, 31.20μg/m^3]	89173	145/1175729	1.54 (1.20,1.99)	8.56×10^{-4}	1.43 (1.11,1.85)	0.006
Q4 (31.20μg/m^3, 108.49μg/m^3]	89043	145/1167246	1.65 (1.28,2.13)	1.23×10^{-4}	1.51 (1.17,1.96)	0.002
P for trend				1.11×10^{-5}		2.75×10^{-4}
NO$_x$ concentration						
Q1 [19.74μg/m^3, 34.00μg/m^3]	89197	101/1187522	Ref		Ref	
Q2 (34.00μg/m^3, 42.10μg/m^3]	89060	107/1177321	1.09 (0.83,1.43)	0.534	1.04 (0.79,1.36)	0.802
Q3 (42.10μg/m^3, 50.60μg/m^3]	89109	159/1170476	1.68 (1.31,2.15)	5.07×10^{-5}	1.54 (1.20,1.98)	7.29×10^{-4}
Q4 (50.60μg/m^3, 265.94μg/m^3]	89116	133/1176283	1.46 (1.13,1.89)	0.004	1.32 (1.01,1.71)	0.040
P for trend				1.44×10^{-4}		0.003
Residential greenspace buffer at 300m						
Q1 [0.28%, 17.30%]	89126	147/1166815	Ref		Ref	
Q2 (17.30%, 29.80%]	89124	140/1176129	0.93 (0.73,1.17)	0.511	0.89 (0.71,1.13)	0.347
Q3 (29.80%, 48.90%]	89115	124/1182137	0.80 (0.63,1.02)	0.067	0.79 (0.61,0.99)	0.044
Q4 (48.90%, 99.18%]	89117	89/1186521	0.56 (0.43,0.73)	1.47×10^{-5}	0.58 (0.44,0.75)	5.09×10^{-5}
P for trend				8.84×10^{-6}		3.12×10^{-5}

Model a: age and sex.

Model b: age, sex, ethnicity, BMI, education level, household income, smoking status, alcohol drinking status, and physical activity.

Abbreviations: BMI: body mass index; CI: confidence intervals; HR: hazard ratio; IQR: interquartile range; ITP: Immune thrombocytopenic purpura; NO$_2$: nitrogen dioxide; NO$_x$: nitrogen oxides; PM$_{2.5}$: fine particulate matter; PM$_{coarse}$: coarse particulate matter with aerodynamic diameters ranging between 2.5 and 10 μm; PM$_{10}$: particulate matter with diameters of less than 10-micrometer.
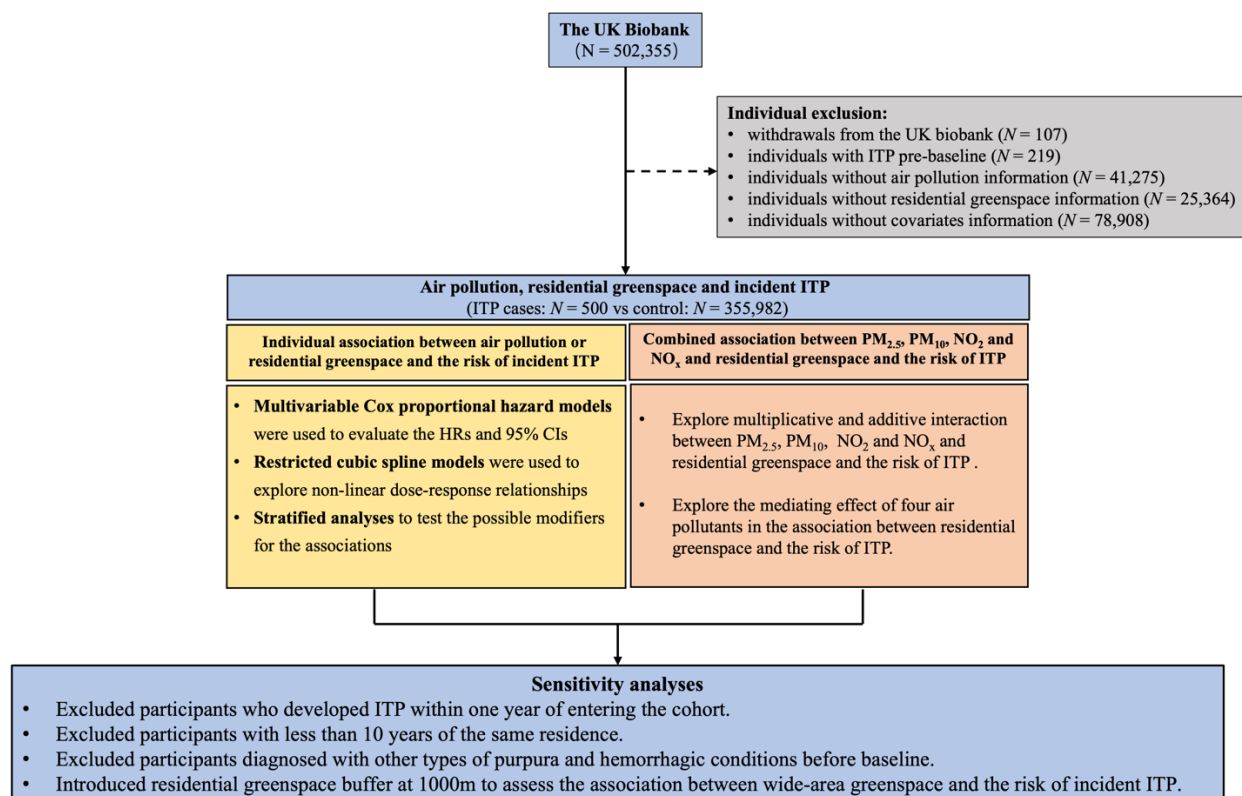
**Supplementary Figure 1.** The flow chart of the study design.

Abbreviations: BMI: body mass index; CIs: confidence intervals; HRs: hazard ratios; ITP: Idiopathic thrombocytopenic purpura; IQR: interquartile range; NO_2: nitrogen dioxide; NO_x: nitrogen oxides; $PM_{2.5}$: fine particulate matter; PM_{10}: particulate matter with diameters of less than 10-micrometer.
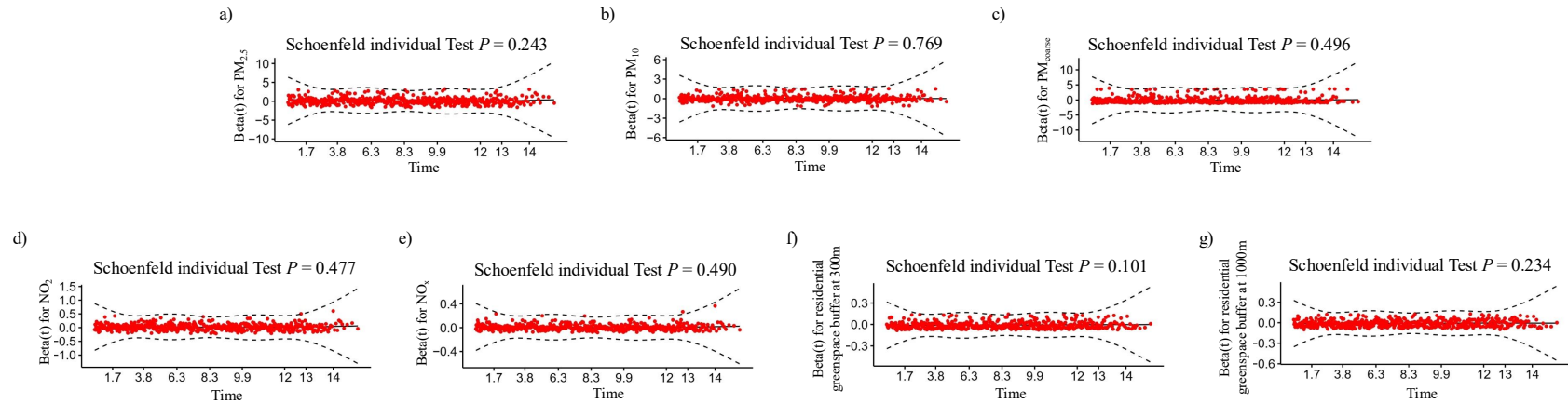
Supplementary Figure 2. Schoenfeld residual plots for air pollution and residential greenspace.

a) $PM_{2.5}$, b) PM_{10}, c) PM_{coarse}, d) NO_2, e) NO_x, f) residential greenspace buffer at 300m, g) residential greenspace buffer at 1000m.

Abbreviations: NO_2: nitrogen dioxide; NO_x: nitrogen oxides; $PM_{2.5}$: fine particulate matter; PM_{coarse}: coarse particulate matter with aerodynamic diameters ranging between 2.5 and 10 μm; PM_{10}: particulate matter with diameters of less than 10-micrometer.
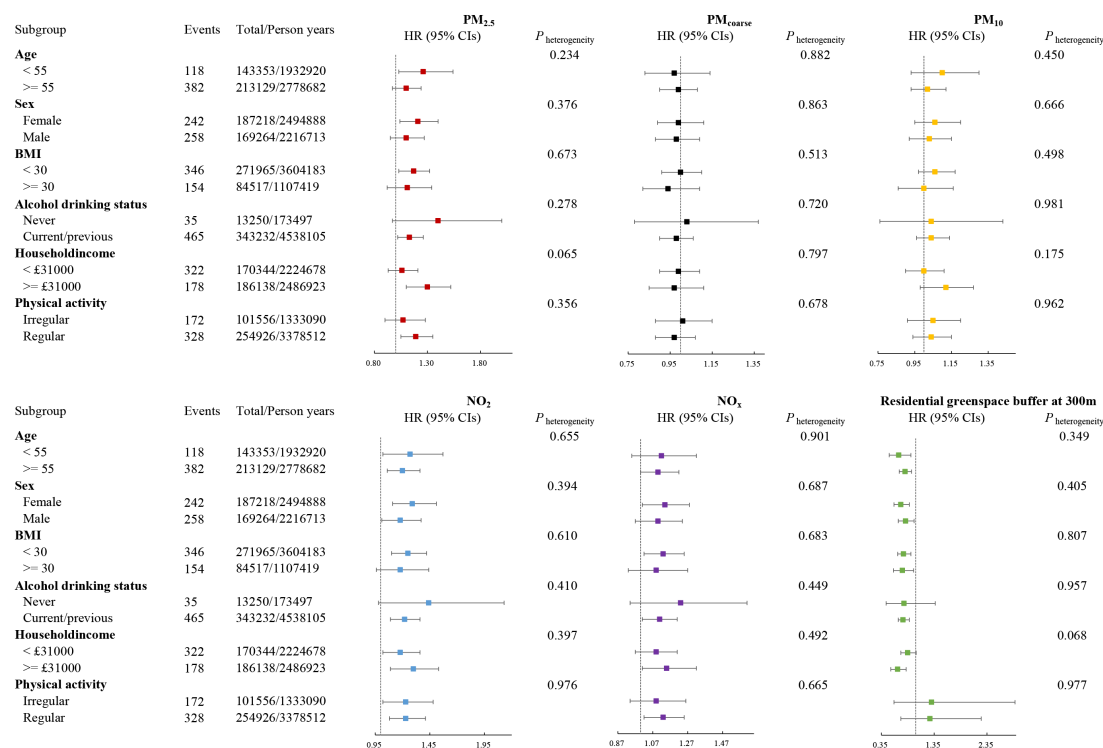
Supplementary Figure 3. Stratified analyses of the associations between air pollution or residential greenspace buffer at 300 m and incident ITP.

[*]Adjusted for age, sex, ethnicity, BMI, education level, household income, smoking status, alcohol drinking status, and physical activity.

Abbreviations: BMI: body mass index; CI: confidence interval; HR: hazard ratio; ITP: Immune thrombocytopenic purpura; NO_2: nitrogen dioxide; NO_x: nitrogen oxides; $PM_{2.5}$: fine particulate matter; PM_{coarse}: coarse particulate matter with aerodynamic diameters ranging between 2.5 and 10 μm; PM_{10}: particulate matter with diameters of less than 10-micrometer.