

<https://doi.org/10.1038/s41531-025-00935-y>

Prominent role of PM10 in the link between air pollution and incident Parkinson's Disease



Alessandro Gialluisi^{1,2}✉, Simona Costanzo^{1,3}, Maria Ilenia De Bartolo^{4,5}, Giovanni Veronesi³, Matteo Renzi⁶, Assuntina Cembalo⁷, Alfonsina Tirozzi¹, Stefania Falciglia⁸, Moreno Ricci⁸, Americo Bonanni¹, Francesco Martone⁷, Gaetano Zazzaro⁷, Antonietta Pepe¹, Daniele Belvisi^{4,5}, Marco Mario Ferrario³, Francesco Gianfagna³, Chiara Cerletti¹, Maria Benedetta Donati¹, Stefania Massari⁹, Alfredo Berardelli^{4,5}, Giovanni de Gaetano¹ & Licia Iacoviello^{1,2}

Air pollution has been associated with Parkinson's Disease (PD) risk, although this relationship remains unclear. We estimated yearly levels of exposure to ten air pollutants (period 2006–2018) in an Italian population cohort, the Moli-sani study ($N = 24,325$; ≥ 35 years; 51.9% women), and derived three principal components, testing their associations with incident PD risk over 23,841 participants (213 cases, median(IQR) follow-up 11.2(2.0) years). This revealed a statistically significant association of PC1 (explaining 38.2% of common variance, tagging PM10 levels), independent on sociodemographic, professional and lifestyles covariates (Hazard Ratio [95%CI] = 1.04[1.02–1.07]). The association was confirmed testing average PM10 levels during follow-up (18[13–24]% increase of PD risk per 1 $\mu\text{g}/\text{m}^3$ increase of PM10). Among different circulating markers, lipoprotein a explained a significant proportion of this association (2.8[0.9; 8.4]%). These findings suggest PM10 as a target to lower PD risk at the population level and a potential implication of lipoprotein a in PD etiology.

In the last decades, pollution – prominently air pollution – has represented a hotspot of investigation and received increasing interest from policy makers, in view of the notable burden that it implies for public health and welfare systems¹. Air pollution is defined as a combination of various compounds originating from different anthropogenic and biogenic emission sources, ranging from particulate matter (PM, with different components based on their aerodynamic diameter) to gases (carbon monoxide, nitrogen dioxide, etc.), and organic chemical compounds (e.g., hydrocarbons)². Not only chronic exposure to air pollution is associated with increased cardiovascular and respiratory diseases risk³, but recent findings have supported a link between air pollution and neurological disorders, including dementia of different origins and Parkinson's disease (hereafter called PD)^{4,5}. PD represents the second most common neurodegenerative disease, characterized by alpha-synuclein aggregates accumulation in dopaminergic neurons of the substantia nigra, which leads to progressive, irreversible loss of these neurons⁶.

While recent reviews and meta-analyses have reported growing evidence suggesting that chronic exposure to air pollution – especially to particulate matter with aerodynamic diameter $< 2.5 \mu\text{m}$ (PM2.5) and nitrogen dioxide (NO2) – is associated with an increased risk of both incident and prevalent neurodegenerative disorders^{7–10}, these associations are often not replicated, especially for PD, and only partly concordant across the different studies, for several reasons. Among them, the heterogeneity of study settings probably represents one of the main hindrances to results concordance, with very scarce longitudinal studies carried out in large population cohorts (e.g., refs. 11–13). Indeed, studies reported so far were mostly carried out within administrative cohorts (e.g., ref. 14), in case-control (e.g., refs. 15–17) and geostatistical studies (e.g., ref. 18). Most of these analyses tested a handful of pollutants simultaneously, but only few analyzed them together with other potential factors influencing PD risk. Therefore, most of these studies often suffer from residual confounding bias due to scarce adjustment for other potential confounders of the association

¹Research Unit of Epidemiology and Prevention, IRCCS NEUROMED, Pozzilli, Italy. ²Department of Medicine and Surgery, LUM University, Casamassima, Italy.

³Department of Medicine and Surgery, EPIMED Research Center, University of Insubria, Varese, Italy. ⁴IRCCS NEUROMED, Pozzilli, Italy. ⁵Department of Human Neurosciences, Sapienza University of Rome, Rome, Italy. ⁶Department of Epidemiology, Lazio Region Health Service/ASL Roma 1, Rome, Italy.

⁷CIRA-Italian Aerospace Research Centre, Capua, Italy. ⁸UOC Governance del Farmaco, Azienda Sanitaria Regionale del Molise –ASREM,

Campobasso, Italy. ⁹Department of Occupational and Environmental Medicine, Epidemiology and Hygiene, Italian Workers' Compensation Authority (INAIL), Rome, Italy. ✉e-mail: alessandro.gialluisi@gmail.com



between air pollution and PD risk, like lifestyles and professional covariates. Other studies are mostly based on a case-control approach – due to the relatively low prevalence of PD (about 1% in the general population over 60 years of age)¹⁹ – and hence may suffer from reverse causality bias. Overall, studies on the environmental epidemiology of PD are quite scarce^{7–10} and

only few of them are based on deeply characterized longitudinal cohorts (e.g., refs. 11–13). These may allow to disentangle the cluster of risk/protective factors for neurodegenerative disorders, including socio-demographic, lifestyles and professional factors.

The main aim of this study was to test the influence of air pollution on the incident risk of PD independent on other potential risk/protective factors. We did this in a general Italian population cohort, with a deep phenotypic assessment, by simultaneously analyzing several air pollutants, along with sociodemographic, lifestyle, and professional exposures, over a 12 years follow-up. We also tested potential mediation of pathways including circulating inflammation, cardiovascular, kidney function, lipid- and platelet-related biomarkers in the link identified, to provide hints on the pathophysiological mechanisms underlying PD.

Table 1 | Baseline characteristics of the population under study

Variable	Population under study (N = 23,841)	Removed participants (N = 484)	P for difference (analyzed vs removed)
Sex (men, %)	48.17	44.83	0.15
Age (y, mean (SD), range)	55.64 (11.91), 34.6–98.7	62.93 (12.30), 35.8–87.5	<0.0001
Education (%)			
Primary or less	25.46	41.11	<0.0001
Lower secondary	27.87	20.04	
Upper secondary	34.22	20.87	
Post-secondary	12.45	9.92	
Missing	-	8.06	
Health conditions (%)			
CVD	5.88	13.74	<0.0001
Cancer	3.46	4.51	<0.0001
Diabetes	4.89	10.12	<0.0001
Hypertension	28.45	41.74	<0.0001
Hyperlipidaemia	7.80	14.38	<0.0001
Lifestyle factors			
Smokers/previous smokers (%)	23.06/27.43	18.64/27.12	0.05
MeDi Score (0–9)	4.35 (1.64)	4.21 (1.61)	0.07
Physical activity (MET-h/day)	3.49 (4.03)	2.99 (3.62)	0.0001
BMI (kg/m ²)	28.05 (4.78)	28.66 (4.69)	0.005
Circulating markers			
PDW (fL)	15.64 (1.78)	15.35 (2.00)	0.01
Cys-C (mg/L)	0.98 (0.25)	1.07 (0.28)	<0.0001
Lp(a) (mg/dL)	18.05 (18.82)	19.21 (19.37)	0.10

Here we report frequency (%) for categorical variables, or, alternatively, mean values and standard deviations (SD) for continuous variables. The population under study results from the exclusion of participants residing outside the Molise region, with lack of follow-up information, or sociodemographic covariates required for basic Cox PH models (education), or with a prevalent PD condition. P-values resulting from statistical comparisons of the analyzed vs non-analyzed participants are reported, rounded to the second decimal place, unless they were significant. Chi-squared test was applied to education levels and smoking; Fisher Exact Test to sex, CVD, cancer, diabetes, hypertension and hyperlipidemia; unpaired t-test to age, Cys-C, MeDi score and BMI; and Wilcoxon signed rank test to PDW, Lp(a) and physical activity levels. All the variables are defined in Supplementary Information. Abbreviations: BMI body mass index, CVD cardiovascular disease, MeDi adherence score to Mediterranean Diet⁴⁰, PDW platelet distribution width, Cys-C cystatin C, Lp(a) lipoprotein a.

Results

Associations between air pollution and incident PD

Table 1 reports the main sociodemographic and epidemiological characteristics of the analyzed sample and the comparison to participants excluded from the analysis. The analyzed subjects were younger (55.64 (11.91) vs 62.93 (12.30) years) and showed a generally higher education level and a lower prevalence of chronic health conditions ($p < 0.0001$).

After 11.2 (2.0) years (median (IQR)) of follow-up, 213 incident cases of PD were identified, with a standardized incidence rate 0.61/1000 person/year in females (95% Confidence Interval = [0.49–0.75]; 1.04 in males ([0.88–1.24]) and 2.4 in participants ≥ 65 years ([2.00–2.80]).

A PCA of air pollutants levels available within the Molise region revealed three main principal components scores, explaining >5% of common variance across all environmental variables tested: PC1 (38.2%), PC2 (19.5%) and PC3 (8.5%; Fig. S1). While PC1 was quite clearly tagging PM10 levels, PC2 and PC3 showed moderate to high loadings of ozone (O₃), carbon monoxide (CO), sulfur dioxide (SO₂), and of nitric oxides (NO_x) and of BTX hydrocarbons (benzene, toluene and xylene), respectively, although less clearly (Fig. S2a–c). This pattern was consistent with Spearman's correlations across all the pollutants tested together, averaged across the years (Fig. S3).

We observed a significant association of PC1 with an increased PD risk (HR [95% CI] = 1.04 [1.02–1.06] per unitary increase of PC1 score), which was stable across models adjusted for professional factors (1.05 [1.02–1.07]) smoking (1.05 [1.03–1.07]) and living in urban settings (1.04 [1.02–1.07]). No other PC score showed significant associations surviving Bonferroni correction for multiple testing of three PC scores ($\alpha = 0.017$; Table 2, Table S1). Since PC1 showed high loadings of PM10 and was therefore clearly tagging the levels of this pollutant, we tested directly average PM10 levels for association with incident PD risk, which revealed relative risks consistent with those observed for PC1. Indeed, this approach was preferred over testing yearly levels of PM10 as time-varying exposure, due to the high correlations across the years of follow-up (Fig. S4). Each unitary ($\mu\text{g}/\text{m}^3$) increase of PM10 was associated with a 18% (95% CI = (13–24)%) increase of PD risk in the most conservative model. The association was stable in a model further enriched for lifestyles (or their proxies) other than smoking which may influence the outcome, including drinking status, adherence to Mediterranean Diet, physical activity and BMI (19 (13–24)%) increase of PD risk per $\mu\text{g}/\text{m}^3$ increase of PM10).

Table 2 | Results of the main analysis of incident PD risk vs principal components of air pollutants exposure

HR [95% CI] Model 1	HR [95% CI] Model 2	HR [95% CI] Model 3	HR [95% CI] Model 4	Exposure	Tagged pollutants
1.04 [1.02–1.06]	1.05 [1.02–1.07]	1.05 [1.03–1.07]	1.04 [1.02–1.07]	PC1	PM10
1.02 [0.99–1.04]	1.01 [0.98–1.04]	1.01 [0.98–1.04]	1.00 [0.97–1.04]	PC2	O ₃ /CO/SO ₂
0.99 [0.95–1.03]	0.99 [0.95–1.03]	0.99 [0.95–1.03]	0.99 [0.95–1.03]	PC3	NO _x /BTX hydrocarbons

Hazard Ratios and 95% confidence intervals HR [95% CI] are reported for each model and each PC score tested, while association p-value is reported only for the most adjusted model (see below). Significant associations surviving Bonferroni correction for multiple testing of three PC scores ($\alpha = 0.017$) are highlighted in bold. Model 1: age and sex; Model 2: Model 1 + education level completed + professional exposure to toxic compounds and working class; Model 3: Model 2 + smoking status; Model 4: Model 3 + urban living. Legend: PM10 = particulate matter with aerodynamic diameter $< 10 \mu\text{m}$; NO_x nitrogen oxides, SO₂ sulfur dioxide, CO Carbon monoxide, O₃ ozone, BTX hydrocarbons benzene, toluene and xylene.

Sensitivity analyses

Sensitivity analyses removing participants with PD onset in the first 2, 5, and 10 years after the start of follow-up (to remove incident cases possibly due to other chronic exposures not directly linked to air pollution) showed consistent associations, as did the analysis after removal of early onset PD cases (Table S2). Since PM10 showed a bimodal distribution (Fig. S5a–j), to ensure against potential biases resulting from departures from normality, we compared participants above and below the median level of PM10 in the analyzed population (11.65 µg/m³; Table S3), which revealed that participants exposed to average PM10 concentrations >11.65 µg/m³ showed a ~14 × (9–24) increase of PD risk, compared to subjects exposed to PM10 concentrations ≤11.65 µg/m³ (Fig. 1).

Mediation analysis

When we tested potential mediations of the association between PM10 levels and incident PD, only three circulating markers showed at least suggestive associations (*p* < 0.1) with both PM10 levels and incident PD risk, namely platelet distribution width (PDW), cystatin C (cys-C) and lipoprotein a (Lp(a)). Only the latter showed explained a significant proportion of the association between PM10 levels and incident PD (2.8 (0.9; 8.4)%, *p* = 0.004), being negatively associated with both incident PD risk and PM10 levels (Table S4). No other putative mediators tested explained a significant proportion of the association between PM10 and PD risk, although plateletcrit approached significance (Table 3).

Discussion

In the present manuscript, we analyzed the relationship between air pollution and incident PD risk in an Italian population, identifying a notable influence of PM10 levels on an increased risk of PD. This influence was independent on diverse other factors representing known risk/protective factors for PD,

including lifestyles like smoking²⁰, and on professional factors like exposure to toxic compounds. While a link between PM10 and neurodegenerative risk has been long hypothesized and an involvement of PM10 in PD mechanisms has been reported in animal models²¹, only few studies supported a clear association with PD, while recent meta-analyses revealed contrasting results and a yet unclear relationship^{7,10,22}, reflecting inconsistent findings. A geospatial analysis of PD cases and controls reported a significant difference in the mean annual NO2 and PM10 levels between areas where PD cases were concentrated (hotspots) and those where they were not (coldspots)¹⁸. These findings are in line with a large nested case-control study from a Chinese health insurance cohort, where PM10 was the only pollutant reported to influence an increase in PD risk, among many others¹⁵. However, other cohort studies found no evidence of association with PM10^{11,13,23,24} as well as case-control studies^{16,17,25,26}, and even meta-analyses^{27,28}. Other works identified associations between PM2.5, incident PD risk (although limited to women living in a specific US region)^{12,13} and clinical events related to PD, like an increased mortality in PD patients¹⁰. Conversely, the Rome Longitudinal Study showed no association of PM2.5 and PM10 levels with Parkinson’s disease¹⁴, in spite of the evidence reported for vascular dementia²⁹. Other studies found associations between exposure to air pollutants – particularly PM2.5, PM10, and NO2 – and neuropathological hallmarks of PD, but also with neuroimaging endophenotypes of neurodegeneration like a reduced prefrontal cortex, hippocampal volume, and cortical thickness⁷. Interestingly, a recent umbrella review and meta-analysis revealed a significant association of PM10 levels with vascular dementia but not with Alzheimer’s Disease risk³⁰. This, along with the inconsistent findings from epidemiological studies on PD and with the scarcity of studies on animal PD models, suggests caution in the interpretation of our findings.

Of note, one of the pollutants most reported to have an influence on PD risk by previous studies, NO2^{7,12}, did not reveal any significant association in our population. This may depend on the different approach used here to investigate air pollution exposure while controlling for collinearity bias – namely through PC scores rather than at single pollutant levels. Still, our finding is not that surprising in view of the inconsistent results published so far for this and other compounds, which may be due to different factors, such as the duration of the exposure, the age of participants, occupational backgrounds and other professional exposures to pollutants and toxic compounds²¹.

Despite the promising findings on PM10, the molecular mechanism of action of this pollutant on the central nervous system remains largely unclear. Among the potential pathways tested, lipoprotein a was the only putative mediator explaining a significant proportion of their link. Although this proportion is relatively small (2.8%), this still provides interesting insights into the pathophysiology of PD and its relationship with air pollution. Lp(a) is a cardiovascular risk marker with a fundamental role in binding other lipoproteins to transport phospholipids in the blood and may have dual influence on inflammation³¹. Interestingly, Lp(a) was found to play a fundamental role in the binding of α-synuclein and high density lipoproteins (HDL) in the human plasma, a potential mechanism through which α-synuclein may be transported through the blood brain barrier³²,

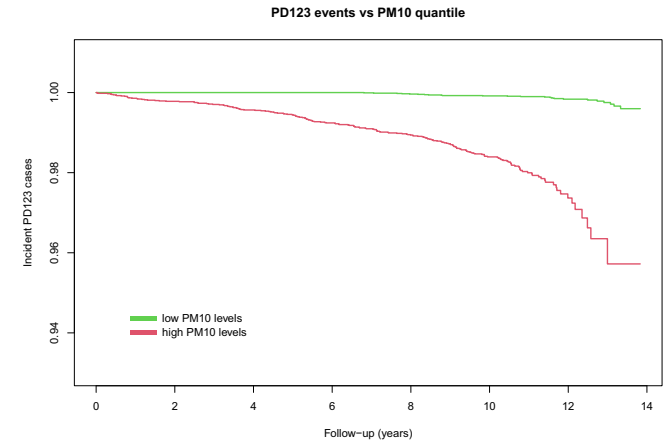


Fig. 1 | Kaplan Meier curves of incident PD events vs PM10 quantiles. Incident Parkinson’s disease events are compared between the two quantiles of exposure to PM10: below the median (11.65 µg/m³; green) and above the median (red).

Table 3 | Results of the mediation analysis of circulating blood markers in the association between PM10 and incident Parkinson’s Disease

Putative Mediator	Total Effect HR [95% CI]	Pure Natural Direct Effect HR [95% CI]	Pure Natural Indirect Effect HR [95% CI]	% mediation [95% CI] (p-value)
PDW	1.095 [1.045; 1.144]	1.103 [1.026; 1.172]	0.993 [0.951; 1.046]	−7.9 [−73.7; 59.2]% <i>P</i> = 0.75
Cys-C ^a	1.088 [1.019; 1.141]	1.083 [1.011 1.136]	1.005 [1.000; 1.010]	5.5 [−0.2; 31.9]% <i>P</i> = 0.06
Lp(a) ^a	1.089 [1.032; 1.138]	1.087 [1.030; 1.136]	1.002 [1.001; 1.005]	2.8 [0.9; 8.4]% <i>P</i> = 0.004

Hazard Ratios and proportion of association mediated are reported for each marker tested, along with 95% Confidence Intervals, as computed through CMAverse package⁷⁷. HRs represent the counterfactual effects conditional on covariates specified in the fully adjusted Cox PH model (Model 4). % mediation was computed as *Direct Effect* × (*Indirect Effect* − 1)/(*Total Effect* − 1). Significant mediations (*p* < 0.05) are highlighted in bold.

PDW platelet distribution width, Cys-C cystatin C, Lp(a) lipoprotein a.

^aFor Cys-C, estimates refer to the counterfactual model accounting also for exposure-mediator interaction, since significant evidence of such interaction was detected (*p* < 0.05).

which may explain the inverse association between PD and Lp(a). Alternatively, this relationship may be explained by the protective influence of Lp(a) against diabetes^{31,33}, which is increasingly reported as a PD risk factor³⁴. A handful of studies specifically investigating the relationship between Lp(a), PD and its endophenotypes have reported contrasting results, like weak positive associations with PD symptoms severity, no association with disease staging³⁵ and a negative correlation with depressive symptoms³⁶ and cognitive performance^{37,38}.

To our knowledge, no studies investigated the relationship between exposure to PM10 and Lp(a), which warrants further investigations. Similarly, in light of the large genetic influence on circulating Lp(a) levels³³, gene-by-environment interactions should also be investigated, to further clarify how pollution and other exposome layers may influence incident PD risk.

This study presents strengths like the comprehensive approach to multiple air pollutants from diverse sources and the simultaneous assessment of many sociodemographic, lifestyle, and professional factors, which allowed us to identify a clear influence of PM10 on incident PD risk. Moreover, we are not aware of any study testing potential explanations of the link between pollution and any neurodegenerative disorder through a vast range of circulating biomarkers as presented here, tagging different pathways. However, our work also has some limitations. First, the use of land measurements to interpolate exposure maps may not be as precise as maps integrating land and satellite data, as well as the fact that data for the first year of recruitment (2005) were not available in the ARPA Molise database. Still, Molise is a rather rural and upland region with does never experience sudden massive changes in the degree of anthropic activities, hence the levels of air pollutants tend to remain stable across the years. On the one hand this poses questions on the generalizability of our findings but on the other hand this may explain the large effect sizes observed. Some residual unmeasured confounding may still explain part of this effect size, but this possibility is quite remote in light of the E-values computed (29.05, see URLs)³⁹. Second, the lack of PM2.5 data – which was in fact not available for the period of interest – hindered a deeper analysis of the influence of different components of particulate matters on PD risk. We plan to further investigate these associations using environmental data from other independent European sources, which will be made available in the future. Third, the algorithm used for defining incident PD cases may have led to classify as cases some subjects affected by related disorders (e.g., parkinsonisms) or with a yet unclear diagnosis, a problem often affecting neurological clinical practice. However, our PD incidence rate was in line with the data reported in other Italian regions^{40,41} and the diagnosis was clinically confirmed in a subgroup of patients with further clinical data available from our center. Fourth, the simultaneous assessment of environmental exposures and circulating markers does not allow us to unambiguously infer direction of causality between them. Last, other environmental factors like exposure to pesticides (not currently available in our cohort) may help untangling the investigated relationship, although proxies like agricultural activity were not previously associated with PD in our cohort³⁴. We are now working on improving many of these aspects, so to run more powerful analyses with more precise exposure estimates and longer and validated neurological follow-up in the near future.

To conclude, we identified a significant direct influence of PM10 levels on incident PD risk in an Italian population cohort. This association was partly explained by lipoprotein a, suggesting that this may play a fundamental role in the link between air pollution and neurodegenerative risk. These findings point towards PM10 as a potential target to lower neurodegenerative risk in the Italian population and warrant further investigations on lipids-related pathways to clarify potential mechanisms underlying PD etiology.

Methods

Population of study

The study population consisted of participants to the Moli-sani project (N = 24,325; 51.9% women, 35–99 years), a prospective cohort of Italian residents recruited from the general population of the Molise Region (Central-Southern Italy), between March 2005 and April 2010. Exclusion

criteria were pregnancy at the time of recruitment, inability to understand terms of participation, current poly-traumas (i.e., simultaneous injury to several organs or body systems), coma, or refusal to sign the informed consent⁴². The Moli-sani Study was approved by the ethical committee of the Catholic University of Rome (approval nr: P99, A-931/03-138-04/CE/2004, 11 February 2004; ClinicalTrials.gov ID: NCT03119142) and all the participants provided written informed consent.

Geolocalization

For this study, participants were geolocalized based on their residence address – available for all the participants – with the use of complementary dedicated software APIs, like Geokettle, QGIS and Here (see URLs). The data were cleaned up manually correcting the errors detected, such as incorrect postal codes. To facilitate the geocoding process, the addresses were transformed by replacing the abbreviations. Using an automated procedure based on the Here API, it was possible to assign latitude and longitude coordinates for all the subjects in the database. A handful of participants with residence outside the Molise region ($n = 17$) were geolocated accordingly and hence removed from the analysis. The results of the geocoding were verified using the shapefiles of the municipalities' boundaries of the region. Addresses placed outside the limits of the city of residence were manually geolocated using Google Maps. With this procedure, it was possible to assign latitude and longitude coordinates for 24,303 subjects in the database, with a high level of confidence. The subjects were then linked to air pollution maps (built as described below), which allowed us to estimate the amount of exposure to each pollutant at the physical coordinates where subjects reported their residence.

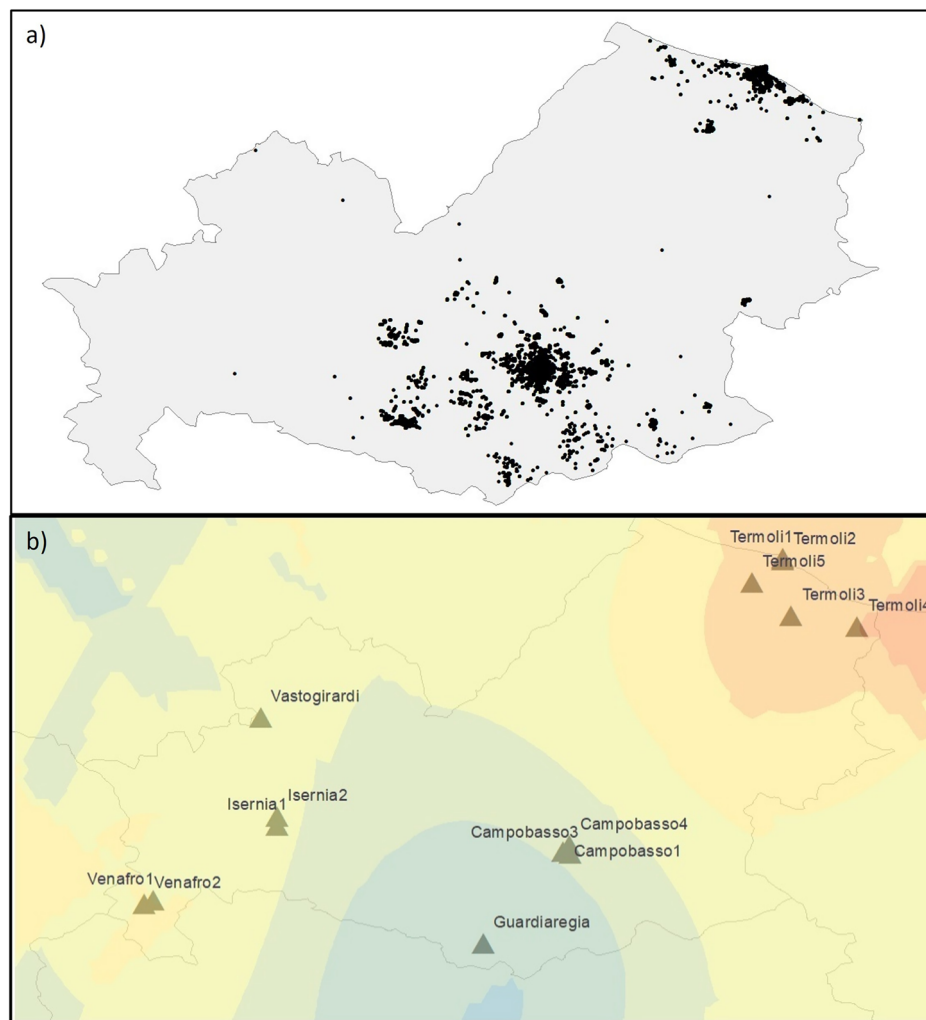
Exposure: air pollution

We estimated yearly levels of exposure to ten different pollutants, including Nitrogen oxides (NOX, NO, NO₂), Ozone (O₃), particulate matter with aerodynamic diameter <10 μm (PM10), Sulfur dioxide (SO₂), Carbon monoxide (CO) and BTX hydrocarbons (benzene, toluene and xylene) in 2006–2018, applying an ordinary Kriging interpolation algorithm – a geostatistical data interpolation method used to infer unsampled points across the spatial field⁴³ (see URLs) – to fourteen land measurements of air pollutants available from the regional environmental authority (ARPA Molise, see URLs). To apply geostatistical algorithms and display analysis results on maps, we used ESRI ArcGIS tool, a proprietary geographic information system (GIS; see URLs) which allowed us to display air monitoring stations on the regional Molise map and finally build maps of exposure to different air pollutants (see Fig. 2a, b). PM2.5 was available for a small period of the follow-up range (nine months in 2018), in only four land measurement centers of the network, and was therefore excluded from the analysis. After defining the variogram to evaluate the spatial autocorrelation of observed data, Kriging algorithm was applied to each pollutant for each year of the follow-up time considered, over participants with physical coordinates available. Algorithm parameters (Semivariogram model, Lag Size, and Nugget) were fine-tuned so to minimize the Root Mean Squared Standardized Error, and the results were validated through cross-validation. Overall, 130 different exposures (10 pollutants × 13 years) were obtained for 24,303 participants. Toluene levels in 2017 presented a single level of exposure and were not varying across subjects, therefore this variable was removed. Since Kriging returned exposure intervals which were largely overlapping but with variable limits across the years, to reduce collinearity among variables, we decided to apply a principal component analysis (PCA) to derive latent variables which could capture most of their shared variance. To this end, we computed the mean of the two limits for each level of exposure (or rank) and each pollutant, so to have point estimates of exposure. The levels of the analyzed pollutants (Fig. S5a–j, Table S3) were comparable to the Italian estimates as reported by the Italy Air Quality Index (AQI) database and often lower than thresholds recommended by the World Health Organization (WHO) air quality guidelines (see URLs).

We carried out a preliminary Kaiser-Meyer-Olkin test of sampling adequacy, which revealed a very good factorability (MSA = 0.97). A

Fig. 2 | Geolocalization of participants and air pollutants maps in the Moli-sani cohort.

a Geographical location (based on residence address) of each participant to the Moli-sani cohort. **b** Example of air pollutant exposure map built in the study, based on Kriging algorithm (PM10, year 2007). Triangles represent land boxes used to build the map, whose data are publicly available from the regional environmental authority (Agenzia Regionale per la Protezione Ambientale del Molise - ARPA Molise, see URLs). Gradient colors represent increasing PM10 levels ($\mu\text{g}/\text{m}^3$), from low (blue shades) to high (red shades).



Bartlett's Test of Sphericity suggested a significant discrepancy with the identity matrix, hence notable reciprocal correlations across the variables, which further supported our choice. PCA was performed through Singular Value Decomposition with orthogonal (varimax) rotation (built-in *prcomp()* function in R⁴⁴). The resulting PC scores were then tested as main exposures in survival analyses, as detailed below.

Outcome: incident PD cases

Ascertainment of the Parkinson's disease status was carried out through linkage with Electronic Health Records (EHRs) databases using fiscal code of each participant as unique identifier. The Molise regional drug prescription registry was first used to identify patients starting PD therapy during the follow-up. According to the national health service, anyone prescribed a drug is entered into a specific registry maintained by the local health authority. Only patients with chronic use of medications with Anatomical Therapeutic Chemical (ATC) classification N04XX (anti-Parkinson drugs) were considered. A similar method for the identification of PD cases has been used in other studies (e.g., ref. 45).

Then, drug registry identification of cases was integrated through the regional registry of hospital discharge records (HDRs). Hospitalization was defined as any stay lasting ≥ 24 h in any hospital, clinics, emergency room, or similar, be it private or public. Hospitalizations showing an International Classification of Diseases, 9th revision (ICD-9) code for PD (332.0) in primary and secondary diagnoses were classified as incident cases. The earliest date indicating a PD status in the above mentioned registries was used to establish the onset of PD. Three independent neurologists with experience in movement disorders further validated PD diagnoses by

combining information on anti-parkinsonian drugs in use (levodopa, dopamine agonists, MAO-B inhibitors) and hospital discharge records. Movement disorder specialists evaluated prescriptions to ensure that only those with dosages indicative of a progressive neurodegenerative disorder were considered. This included evidence of gradual dose escalation or the addition of other anti-parkinsonian medications over time, reflecting disease progression. Consequently, individuals with chronic low-dose dopamine agonist prescriptions, without signs of dose adjustment or combination therapy, were excluded, thereby minimizing the risk of misclassifying cases of Restless Legs Syndrome. Follow-up of each cohort member continued until the date of the diagnosis of Parkinson's disease, death, or 31 December 2018.

In a subsample ($n = 99$) of the 213 incident PD cases identified using these criteria, it was possible to further validate PD diagnosis through medical records of the IRCCS Neuromed outpatients Parkinson Center (records of periodic outpatient visits, clinical diaries, discharge letters, DATSCAN examinations, drug prescriptions), which confirmed a PD clinical diagnosis and hence correct classification in 100% of these cases.

Statistical analysis

All statistical analyses of the present manuscript were carried out in R (see URLs)⁴⁴. After removing prevalent PD ($n = 52$) cases, as well as those participants with missing information on incident events (372), socio-demographic covariates (38) and environmental exposure (22, see above), we finally tested 23,841 subjects for incident PD risk vs air pollution exposure through multivariable Cox Proportional Hazards (PH) regression models⁴⁶. To test potential influences of air pollution independent on other

clusters of risk/protective factors, multivariable models were built, incrementally adjusted for (i) age and sex (Model 1); (ii) education level completed and professional factors like working class and exposure to toxic compounds (Model 2); (iii) lifestyles like smoking (Model 3) and (iv) living in urban or rural settings, since this may affect both the levels of exposures to air pollutants and the access to movement disorders specialists (Model 4). Except for age and sex, these covariates were identified as confounders based on Directed Acyclic Graph (DAG) analysis, through DAGitty v3.1 (see URLs and Fig. S6). All the covariates are described in detail below. The main pollution exposure was initially tested as PC scores, as computed through the PCA over all yearly pollutant levels, which were tested all together in multivariable models. If any PC score showed a significant association, levels of the single pollutant/s tagged by that PC were tested, averaged over the actual years of follow-up for each participant. All the models showed Variance Inflation Factors (VIF) < 2, suggesting negligible collinearity bias.

A sensitivity model with additional lifestyles other than smoking (namely drinking status, adherence to Mediterranean Diet, physical activity in leisure time) or their proxies (e.g., body mass index, BMI) was performed to analyze the stability of the association in enriched models with other exposures potentially influencing the outcome. Similarly, additional sensitivity analyses were carried out, removing participants with PD onset in the first 2, 5 and 10 years after the start of follow-up – to rule out the influence of other chronic exposures not directly linked to air pollution – and removing cases with early onset (below 50, 55 and 60 years of age), which were likely under strong genetic influence.

Definition of covariates

Smoking status was defined as a three-class variable based on cigarette smoking habits of participants: *smokers*, *ex-smokers* (i.e., subjects who quit at least one year before the interview) and *never-smokers* (reference class).

Leisure-time *physical activity* was assessed through a structured questionnaire and expressed as daily energy expenditure in metabolic equivalent task-hours (MET-h/day)⁴⁷.

Food intake was assessed through the validated Italian EPIC food frequency questionnaire⁴⁸. Adherence to *Mediterranean Diet* was defined according to the Mediterranean Diet Score, ranging from 0 (low adherence) to 9 (high adherence)⁴⁹. Then we defined three adherence classes: low (Trichopoulos score 0–3, used as reference), moderate (4–6) and high (7–9). The EPIC questionnaire also allowed to compute *alcohol consumption* habits along with additional questions, as described by ref. 50. Participants were classified based on the estimated daily alcohol intake in the year before enrollment: life-time abstainers, former drinkers, occasional drinkers and current drinkers who drank 1–12 (reference class), 12.1–24, 24.1–48 and >48 g/day.

Height and weight were measured for each participant, as well as waist circumference (cm), which was measured in the middle between the 12th rib and the iliac crest, while hip circumference (cm) was measured around the buttocks. Then *Body mass index (BMI)* was calculated, and participants were grouped into three categories: <25, 25–29.9, and ≥30 kg/m².

Socioeconomic status (SES) information was self-reported and/or collected by a structured questionnaire administered by trained personnel. *Educational attainment* was defined as the education level completed and subjects were divided into four classes: None/Primary (reference), Lower secondary, Upper secondary and Post-secondary. Current *occupational social class* was classified based on the Registrar General's occupation classification scheme, and ranked as previously described by ref. 51. Based on this rank, five occupational classes were defined: professional/managerial, skilled non-manual, skilled manual, partly skilled/unskilled and unemployed/unclassified (reference).

Professional exposure to potentially toxic compounds, like concrete, eternit, aluminum, starchy and paper-derived compounds was also self-reported by each participant (yes/no).

Prevalent *diabetes* and *hyperlipidemia* were defined as dichotomous variables (Yes/No), based on the reported and verified use of specific drugs

for the treatment of these disorders. Prevalent *cardiovascular disease (CVD)* and *cancer* were defined as binary variables, classified into subjects reporting medical history of the disease (possibly supported by medical documentation or by the use of specific drugs), and those with no medical history of the disease.

Urban or rural setting was defined based on the urbanization level of the place of residence, as described by the European Institute of Statistics (EUROSTAT definition) and as defined in the Atlante Statistico dei Comuni, 2014 Edition (<https://www.istat.it/notizia/atlane-statistico-dei-comuni/>), by the Italian National Institute of Statistics (ISTAT).

Mediation analysis

We tested the role of different biomarkers available in the Moli-sani study as potential mediation pathways in the link between air pollution and PD risk. Circulating inflammation has been hypothesized to play a role in the above mentioned link⁷, and was therefore tested through a composite blood-based inflammation index, called INFLA-score, capturing both serum and cellular-circulating inflammation⁵². Additional non-composite blood markers were also tested, which were previously reported to influence PD risk. These included (i) lipid biomarkers, like total cholesterol, triglycerides, high (HDL) and low density lipoprotein (LDL)^{53–55}, lipoprotein a (Lp(a)), apolipoprotein A1 (Apo-A1) and B (Apo-B)^{32,35,56–58}; (ii) markers of glucose metabolism^{59,60}: glucose, C-peptide, insulin; (iii) cardiovascular risk markers: N-Terminal Pro-B-Type Natriuretic Peptide (NT-proBNP) and high-sensitivity cardiac troponin I⁶¹; (iv) other hormones: testosterone⁶² and vitamin D⁶³; (v) renal markers: uric acid, albumin, creatinine, c statin C (Cys-C)^{64–66}; (vi) other inflammation markers like high sensitivity C-reactive protein (CRP) and immunity-related markers like total white blood cells (WBC), lymphocytes (LY), neutrophils (NE) and neutrophil-to-lymphocyte ratio (NLR)^{57,67–71}; (vii) red blood cell traits like red cell distribution width (RDW)⁷², and hemoglobin levels (Hgb)⁷³; (viii) platelet parameters like mean platelet volume (MPV), plateletcrit (PCT), platelet count (Plt) and distribution width (PDW)^{61,74–76}.

All these biomarkers were preliminarily screened as putative mediators by testing associations with both the exposure (PM10 levels) and the outcome (incident PD risk). Only three markers showed at least suggestive associations with both the exposure and the outcome ($p < 0.1$; see Results section) – namely PDW, Lp(a) and Cys-C – and hence underwent a formal mediation analysis. This was carried out through the CMAverse package⁷⁷, using a counterfactual approach. This framework is more robust than other traditional approaches since it protects against mediator-outcome confounding affected by the exposure, exposure-mediator interaction, and exposure-outcome confounding⁷⁸. The counterfactual approach allows simultaneous estimation of the direct (the PM10-incident PD association not due to the mediator), indirect (association due to the mediator), and total effects (the sum of the direct and indirect effects). Using the *cmest()* function (regression-based method), we estimated these effects in fully adjusted Cox PH regressions (Model 4), for each putative mediator separately. Confidence intervals for the indirect effects were estimated through a bootstrap-based approach, over 1000 bootstrap samples with replacement. Where significant exposure-mediator interactions were detected ($p < 0.05$, for Cys-C), the direct, indirect and total effects were estimated under this assumption (option *EMint = T*). The proportion of association explained by each putative mediator was calculated as $\text{Direct Effect} \times (\text{Indirect Effect} - 1) / (\text{Total Effect} - 1)$, and proportions with $p < 0.05$ were considered significant.

Measurement of circulating biomarkers

Blood samples were obtained between 7 and 9 AM from participants who had fasted overnight and had refrained from smoking for at least 6 h (h). Biochemical analyses were performed in the centralized Moli-sani laboratory, as described elsewhere⁷⁹. Haemochromocytometric analyses were performed by cell counter (Coulter HMX, Beckman Coulter, Milan, Italy) within 3 h from venepuncture. All haemochromocytometric analyses (white blood cell and platelet counts, granulocyte % and lymphocyte %) were performed by cell counter (Coulter HMX, Beckman Coulter, Milan,

Italy) within 3 h from venipuncture. CRP was measured in fresh serum, by a latex particle-enhanced immunoturbidimetric assay (ILab 350 Instrumentation Laboratory, Milan, Italy).

Analyses of creatinine, apolipoproteins, C-peptide, cys-C, insulin, Lp(a), NT-proBNP, testosterone, troponin I and vitamin D were performed within the BiomarCaRE Consortium and have been described in details elsewhere (<https://cordis.europa.eu/project/id/278913/reporting/it>). Main information and descriptive statistics for all the circulating biomarkers currently available in the Moli-sani study are reported in⁸⁰.

URLs

ARPA Molise: <https://www.arpamolise.it/>, Kriging algorithm: <https://desktop.arcgis.com/en/arcmap/10.3/tools/3d-analyst-toolbox/how-kriging-works.htm>, GeoKettle: <http://www.geokettle.org/>, HERE: <https://www.here.com/>, QGIS: <https://www.qgis.org/en/site/>, Italy Air Quality Index (AQI) database: <https://www.aqi.in/dashboard/italy>, World Health Organization (WHO) air quality guidelines: <https://www.eea.europa.eu/publications/status-of-air-quality-in-Europe-2022/europes-air-quality-status-2022/world-health-organization-who-air>, R: <https://www.r-project.org/>, DAGitty: <https://www.dagitty.net/>, E-value online calculator: <https://www.evalue-calculator.com/>.

Data availability

The data underlying this article will be shared upon reasonable request to the corresponding author. The data are stored in an institutional repository (<https://repository.neuromed.it>) and access is restricted by the ethics approval and the legislation of the European Union.

Code availability

The analysis code underlying the findings reported in this article will be shared upon reasonable request to the corresponding author.

Abbreviations

PD	Parkinson's Disease;
PM	particulate matter
PM2.5/PM10	particulate matter with aerodynamic diameter < 2.5/10 µm
NO2	nitrogen dioxide
NO	nitrogen monoxide
NOx	nitrogen oxides
O3	ozone
CO	carbon monoxide
SO2	sulfur dioxide
BTX hydrocarbons	benzene, toluene and xylene
PCA	principal component analysis
MSA	measure of sampling adequacy
HR	Hazard Ratio
CI	Confidence Interval
IQR	interquartile range
EHRs	Electronic Health Records
ATC	Anatomical Therapeutic Chemical
HDRs	hospital discharge records
ICD-9	International Classification of Diseases, 9th revision
VIF	Variance Inflation Factor
BMI	body mass index
CVD	cardiovascular disease
MeDi	adherence score to Mediterranean Diet
HDL/LDL	high/low density lipoprotein
Lp(a)	lipoprotein a
APO-A1/-B	apolipoprotein A1/B
NT-proBNP	N-Terminal Pro-B-Type Natriuretic Peptide
Cys-C	cystatin C
CRP	high sensitivity C-reactive protein
WBC	total white blood cells (WBC)
LY/NE	lymphocytes/neutrophils proportions over WBC

NLR	neutrophil-to-lymphocyte ratio
RDW	red cell distribution width
Hgb	hemoglobin levels
MPV	mean platelet volume
PCT	plateletcrit
Plt	platelet count
PDW	platelet distribution width.

Received: 1 October 2024; Accepted: 2 April 2025;

Published online: 07 May 2025

References

1. Organización Mundial de la Salud (OMS). WHO global air quality guidelines. *Coast. Estuar. Process.* 1–360 (2021).
2. Akimoto, H. Global air quality and pollution. *Science* **302**, 1716–1719 (2003).
3. Thurston, G. D. et al. A joint ERS/ATS policy statement: what constitutes an adverse health effect of air pollution? An analytical framework. *Eur. Respir. J.* **49**, 1600419 (2017).
4. Xu, X., Ha, S. U. & Basnet, R. A review of epidemiological research on adverse neurological effects of exposure to ambient air pollution. *Front. Public Heal.* **4**, 157 (2016).
5. Block, M. L. & Calderón-Garcidueñas, L. Air pollution: mechanisms of neuroinflammation and CNS disease. *Trends Neurosci.* **32**, 506–516 (2009).
6. Jankovic, J. Parkinson's disease: clinical features and diagnosis. *J. Neurol. Neurosurg. Psychiatry* **79**, 368–376 (2008).
7. Pini, L. et al. The pollutome-connectome axis: a putative mechanism to explain pollution effects on neurodegeneration. *Ageing Res. Rev.* **86**, 101867 (2023).
8. Palacios, N. Air pollution and Parkinson's disease - Evidence and future directions. *Rev. Environ. Health* **32**, 303–313 (2017).
9. Zhao, Y. L. et al. Environmental factors and risks of cognitive impairment and dementia: a systematic review and meta-analysis. *Ageing Res. Rev.* **72**, 101504 (2021).
10. Cole-Hunter, T. et al. Long-term air pollution exposure and Parkinson's disease mortality in a large pooled European cohort: an ELAPSE study. *Environ. Int.* **171**, 107667 (2023).
11. Jo, S. et al. Association of NO2 and other air pollution exposures with the risk of parkinson disease. *JAMA Neurol.* **78**, 800–808 (2021).
12. Cao, Z. et al. Air pollutants and risk of Parkinson's disease among women in the sister study. *Environ. Health Perspect.* **132**, 17001 (2024).
13. Yu, Z. et al. Air pollution, surrounding green, road proximity and Parkinson's disease: a prospective cohort study. *Environ. Res.* **197**, 111170 (2021).
14. Cerza, F. et al. Residential exposure to air pollution and incidence of Parkinson's disease in a large metropolitan cohort. *Environ. Epidemiol.* **2**, e023 (2018).
15. Chen, C. Y. et al. Long-term exposure to air pollution and the incidence of Parkinson's disease: a nested case-control study. *PLoS ONE* **12**, 1–14 (2017).
16. Toro, R. et al. Parkinson's disease and long-term exposure to outdoor air pollution: a matched case-control study in the Netherlands. *Environ. Int.* **129**, 28–34 (2019).
17. Rumrich, I. K. et al. Long-term exposure to low-level particulate air pollution and Parkinson's disease diagnosis - a Finnish register-based study. *Environ. Res.* **229**, 115994 (2023).
18. Fleury, V. et al. Geospatial analysis of individual-based Parkinson's disease data supports a link with air pollution: a case-control study. *Park. Relat. Disord.* **83**, 41–48 (2021).
19. Gialluisi, A. et al. Whole exome sequencing study of Parkinson disease and related endophenotypes in the italian population. *Front. Neurol.* **10**, 1362 (2020).

20. Belvisi, D. et al. Modifiable risk and protective factors in disease development, progression and clinical subtypes of Parkinson's disease: what do prospective studies suggest?. *Neurobiol. Dis.* **134**, 104671 (2020).
21. Lee, J. et al. Particulate matter exposure and neurodegenerative diseases: a comprehensive update on toxicity and mechanisms. *Ecotoxicol. Environ. Saf.* **266**, 115565 (2023).
22. Dhiman, V., Trushna, T., Raj, D. & Tiwari, R. R. Is ambient air pollution a risk factor for Parkinson's disease? A meta-analysis of epidemiological evidence. *Int. J. Environ. Health Res.* **00**, 1–18 (2022).
23. Palacios, N. et al. Air pollution and risk of parkinson's disease in a large prospective study of men. *Environ. Health Perspect.* **125**, 1–7 (2017).
24. Palacios, N. et al. Particulate matter and risk of Parkinson disease in a large prospective study of women. *Environ. Heal. A Glob. Access Sci. Source* **13**, 1–9 (2014).
25. Liu, R., Young, M. T., Chen, J. C., Kaufman, J. D. & Chen, H. Ambient Air Pollution Exposures and Risk of Parkinson Disease. *Environ. Health Perspect.* **124**, 1759–1765 (2016).
26. Lee, P. C. et al. Traffic-related air pollution increased the risk of Parkinson's disease in Taiwan: a nationwide Study. *Environ. Int.* **96**, 75–81 (2016).
27. Han, C., Lu, Y., Cheng, H., Wang, C. & Chan, P. The impact of long-term exposure to ambient air pollution and second-hand smoke on the onset of Parkinson disease: a review and meta-analysis. *Public Health* **179**, 100–110 (2020).
28. Wang, Y., Liu, Y. & Yan, H. Effect of long-term particulate matter exposure on Parkinson's risk. *Environ. Geochem. Health* **42**, 2265–2275 (2020).
29. Cerza, F. et al. Long-term exposure to air pollution and hospitalization for dementia in the Rome longitudinal study. *Environ. Heal. A Glob. Access Sci. Source* **18**, 1–12 (2019).
30. Jones, A. et al. Environmental risk factors for all-cause dementia, Alzheimer's disease dementia, vascular dementia, and mild cognitive impairment: an umbrella review and meta-analysis. *Environ. Res.* **270**, 121007 (2025).
31. Safiullah, Z. N., Leucker, T., Jones, S. R. & Toth, P. P. Physiological roles and functions of Lipoprotein(a) BT - Lipoprotein(a) (eds. Kostner, K., Kostner, G. M. & Toth, P. P.) 135–158 (Springer International Publishing). https://doi.org/10.1007/978-3-031-24575-6_8 (2023).
32. Emamzadeh, F. N. & Allsop, D. α -Synuclein Interacts with Lipoproteins in Plasma. *J. Mol. Neurosci.* **63**, 165–172 (2017).
33. Kronenberg, F. et al. Lipoprotein(a) in atherosclerotic cardiovascular disease and aortic stenosis: A European Atherosclerosis Society consensus statement. *Eur. Heart J.* **43**, 3925–3946 (2022).
34. Gialluisi, A. et al. Risk and protective factors in Parkinson's disease: a simultaneous and prospective study with classical statistical and novel machine learning models. *J. Neurol.* **270**, 4487–4497 (2023).
35. Choe, C. et al. Association of lipid levels with motor and cognitive function and decline in advanced Parkinson's disease in the Mark-PD study. *Park. Relat. Disord.* **85**, 5–10 (2021).
36. Dong, M. X., Wei, Y. D. & Hu, L. The disturbance of lipid metabolism is correlated with neuropsychiatric symptoms in patients with Parkinson's disease. *Chem. Phys. Lipids* **239**, 105112 (2021).
37. Röhr, F. et al. Relationship between Lipoprotein (a) and cognitive function – Results from the Berlin Aging Study II. *Sci. Rep.* **10**, 1–10 (2020).
38. Solfrizzi, V. et al. Lipoprotein(a), apolipoprotein E genotype, and risk of Alzheimer's disease. *J. Neurol. Neurosurg. Psychiatry.* **72**, 732–736 (2002).
39. Mathur, M. B., Ding, P., Riddell, C. A. & VanderWeele, T. J. Web Site and R Package for Computing E-values. *Epidemiology* **29**, e45–e47 (2018).
40. Valent, F. et al. The epidemiology of Parkinson's disease in the Italian region Friuli Venezia Giulia: a population-based study with administrative data. *Neurol. Sci.* **39**, 699–704 (2018).
41. Eusebi, P. et al. Incidence and prevalence of Parkinson's disease in the Italian region of Umbria: a population-based study using healthcare administrative databases. *Neurol. Sci.* **40**, 1709–1712 (2019).
42. Di Castelnuovo, A. et al. Distribution of short and lifetime risks for cardiovascular disease in Italians. *Eur. J. Prev. Cardiol.* **19**, 723–730 (2012).
43. Singh, P. & Tyagi, A. Applying kriging approach on pollution data using gis software. *Int. J. Environ. Eng. Manag.* **4**, 185–190 (2013).
44. R Core Team. *R: A Language and Environment for Statistical Computing.* (2020).
45. Hu, G., Jousilahti, P., Bidel, S., Antikainen, R. & Tuomilehto, J. Type 2 diabetes and the risk of Parkinson's disease. *Diabetes Care* **30**, 842–847 (2007).
46. Therneau, T. M. & Grambsch, P. M. *Modeling survival data: extending the Cox model.* (Springer, 2000).
47. Ainsworth, B. E. et al. Compendium of physical activities: an update of activity codes and MET intensities. *Med. Sci. Sports Exerc.* **32**, S498–S504 (2000).
48. Pisani, P., et al. Relative validity and reproducibility of a food frequency dietary questionnaire for use in the Italian EPIC centres. *Int. J. Epidemiol.* **26**, S152–S160 (1997).
49. Trichopoulou, A., Costacou, T., Bamia, C. & Trichopoulos, D. Adherence to a mediterranean diet and survival in a greek population. *N. Engl. J. Med.* **348**, 2599–2608 (2003).
50. Costanzo, S. et al. Alcohol consumption and hospitalization burden in an adult Italian population: prospective results from the Moli-sani study. *Addiction* **114**, 636–650 (2019).
51. Bonaccio, M. et al. Relative contribution of health-related behaviours and chronic diseases to the socioeconomic patterning of low-grade inflammation. *Int. J. Public Health* **62**, 551–562 (2017).
52. Pounis, G. et al. Polyphenol intake is associated with low-grade inflammation, using a novel data analysis from the Moli-sani study. *Thromb. Haemost.* **115**, 344–352 (2016).
53. Huang, X. et al. Serum cholesterol and the progression of parkinson's disease: results from DATATOP. *PLoS ONE* **6**, e22854 (2011).
54. Rozani, V. et al. Higher serum cholesterol and decreased Parkinson's disease risk: a statin-free cohort study. *Mov. Disord.* **33**, 1298–1305 (2018).
55. Park, J. H. et al. Association of high-density lipoprotein cholesterol variability and the risk of developing Parkinson disease. *Neurology* **96**, E1391–E1401 (2021).
56. Chelliah, S. S., Bhuvanendran, S., Magalingam, K. B., Kamarudin, M. N. A. & Radhakrishnan, A. K. Identification of blood-based biomarkers for diagnosis and prognosis of Parkinson's disease: a systematic review of proteomics studies. *Ageing Res. Rev.* **73**, 101514 (2022).
57. Lawton, M. et al. Blood biomarkers with Parkinson's disease clusters and prognosis: the oxford discovery cohort. *Mov. Disord.* **35**, 279–287 (2020).
58. Fang, F. et al. Lipids, Apolipoproteins, and the risk of Parkinson disease: a prospective cohort study and a mendelian randomization analysis. *Circ. Res.* 643–652. <https://doi.org/10.1161/CIRCRESAHA.119.314929> (2020).
59. Chmiela, T. et al. If Not Insulin Resistance so What? – Comparison of Fasting Glycemia in Idiopathic Parkinson's Disease and Atypical Parkinsonism. *Diabetes Metab. Syndr. Obes.* **15**, 1451–1460 (2022).
60. Sabari, S. S. et al. Type 2 Diabetes (T2DM) and Parkinson's Disease (PD): a mechanistic approach. *Mol. Neurobiol.* **60**, 4547–4573 (2023).
61. Yang, Y. et al. The shared genetic landscape of blood cell traits and risk of neurological and psychiatric disorders. *Cell Genom.* **3**, 100249 (2023).
62. Bourque, M., Soulet, D. & Di Paolo, T. Androgens and Parkinson's disease: a review of human studies and animal models. *Androg. Clin. Res. Ther.* **2**, 294–303 (2021).
63. Fullard, M. E. & Duda, J. E. A review of the relationship between Vitamin D and Parkinson disease symptoms. *Front. Neurol.* **11**, 454 (2020).

64. Hu, W.-D. et al. Elevated Cystatin C levels are associated with cognitive impairment and progression of parkinson disease. *Cognit. Behav. Neurol.* **29**, 144–149 (2016).
 65. Qu, Y. et al. Estimated glomerular filtration rate is a biomarker of cognitive impairment in Parkinson's disease. *Front. Aging Neurosci.* **15**, 1130833 (2023).
 66. Thaler, A. et al. Biochemical markers for severity and risk in GBA and LRRK2 Parkinson's disease. *J. Neurol.* **268**, 1517–1525 (2021).
 67. Jensen, M. P. et al. Lower lymphocyte count is associated with increased risk of Parkinson's Disease. *Ann. Neurol.* **89**, 803–812 (2021).
 68. Muñoz-Delgado, L. et al. Peripheral inflammatory immune response differs among sporadic and familial Parkinson's disease. *npj Park. Dis.* **9**, 12 (2023).
 69. Grillo, P. et al. Neutrophil-to-lymphocyte ratio and lymphocyte count reflect alterations in central neurodegeneration-associated proteins and clinical severity in Parkinson Disease patients. *Parkinsonism Relat. Disord.* **112**, 105480 (2023).
 70. Qiu, X. et al. C-Reactive protein and risk of Parkinson's Disease: a systematic review and meta-analysis. *Front. Neurol.* **10**, 384 (2019).
 71. Muñoz-Delgado, L. et al. Peripheral Immune Profile and Neutrophil-to-Lymphocyte Ratio in Parkinson's Disease. *Mov. Disord.* **36**, 2426–2430 (2021).
 72. Qin, Y. H., Wei, Y. S., Meng, L. Q. & Chen, W. C. The role of red cell distribution width in patients with Parkinson's disease. *Int. J. Clin. Exp. Med.* **9**, 6143–6147 (2016).
 73. Deng, Q. et al. Lower hemoglobin levels in patients with parkinson's disease are associated with disease severity and iron metabolism. *Brain Res.* **1655**, 145–151 (2017).
 74. Izzi, B. et al. Beyond haemostasis and thrombosis: platelets in depression and its co-morbidities. *Int. J. Mol. Sci.* 1–32 <https://doi.org/10.3390/ijms21228817> (2020).
 75. Tirozzi, A. et al. Assessing genetic overlap between platelet parameters and neurodegenerative disorders. *Front. Immunol.* **11**, 2127 (2020).
 76. Tirozzi, A. et al. A multi-trait association analysis of brain disorders and platelet traits identifies novel susceptibility loci for major depression, Alzheimer's and Parkinson's Disease. *Cells* **12**, 245 (2023).
 77. Choirat, C., Coull, B. A., VanderWeele, T. J. & Valeri, L. CMAverse: a suite of functions for reproducible causal mediation analyses. *Epidemiology* **32**, 20–22 (2021).
 78. Richiardi, L., Bellocco, R. & Zugna, D. Mediation analysis in epidemiology: methods, interpretation and bias. *Int. J. Epidemiol.* **42**, 1511–1519 (2013).
 79. Santimone, I. et al. White blood cell count, sex and age are major determinants of heterogeneity of platelet indices in an adult general population: results from the MOLI-SANI project. *Haematologica* **96**, 1180–1188 (2011).
 80. Gialluisi, A. et al. Exploring domains, clinical implications and environmental associations of a deep learning marker of biological ageing. *Eur. J. Epidemiol.* <https://doi.org/10.1007/s10654-021-00797-7> (2021).
- (Brescia, Italy), Livia Rago (Campobasso, Italy). Baseline and Follow-up Data Management: Simona Costanzo (Coordinator), Marco Olivieri (Campobasso, Italy), Teresa Panzera. Data Analysis: Augusto Di Castelnuovo (Coordinator), Marialaura Bonaccio, Simona Costanzo, Simona Esposito, Alessandro Gialluisi, Francesco Gianfagna, Sabatino Orlandi, Emilia Ruggero, Alfonsina Tirozzi. Biobank, Molecular and Genetic Laboratory: Amalia De Curtis (Coordinator), Sara Magnacca, Fabrizia Noro, Alfonsina Tirozzi. Recruitment Staff: Mariarosaria Persichillo (Coordinator), Francesca Braccone, Teresa Panzera. Communication and Press Office: Americo Bonanni. Regional Institutions: Direzione Generale per la Salute - Regione Molise; Azienda Sanitaria Regionale del Molise (ASReM, Italy); Agenzia Regionale per la Protezione Ambientale del Molise (ARPA Molise, Italy); Molise Dati Spa (Campobasso, Italy); Offices of vital statistics of the Molise region. Hospitals: Presidi Ospedalieri ASReM: Ospedale A. Cardarelli - Campobasso, Ospedale F. Veneziale - Isernia, Ospedale San Timoteo - Termoli (CB), Ospedale Ss. Rosario - Venafrò (IS), Ospedale Vietri - Larino (CB), Ospedale San Francesco Caracciolo - Agnone (IS); Casa di Cura Villa Maria - Campobasso; Ospedale Gemelli Molise - Campobasso; IRCCS Neuromed - Pozzilli (IS). Past Moli-sani Study Investigators are available at https://www.moli-sani.org/?page_id=173. We also thank the Associazione Cuore Sano Onlus (Campobasso, Italy) for its cultural support, Marno Srl (Rosignano Marittimo, LI), Innovation Group Srl (Sant'Agapito, IS) for the record linkage with electronic health data, and the IRCCS Neuromed Biobanking Center for the management of samples and for the valuable support and cooperation. The enrollment phase of the Moli-sani Study (2005–2010) was supported by unrestricted research grants from the Pfizer Foundation (Rome, Italy), the Italian Ministry of University and Research (MIUR, Rome, Italy)—Programma Triennale di Ricerca, Decreto no.1588, and Instrumentation Laboratory, Milan, Italy. The follow-up phase of the Moli-sani Study (validation of incident cases) was partially supported the Italian Ministry of Health (PI LI, CoPI MB; grant no. PNRR-MAD-2022-12376167). The present analyses were partially supported by the INAIL-Bric 2019 ID 47 project (Italian National Institute for Insurance against Accidents at Work; call 2019, project CUP code: F24119000630008) and by Next Generation EU, in the context of the National Recovery and Resilience Plan, Investment PE8 – Project Age-It: “Ageing Well in an Ageing Society” [DM 1557 11.10.2022]. No funder had a role in study design, collection, analysis, interpretation of data, writing of the manuscript, and decision to submit this article for publication.

Author contributions

A.G., S.C., L.I., S.M., and F.G. M.B.D., G.d.G., C.C., G.V., M.M.F., and M.R. contributed to the concept and design of the work and/or to the interpretation of data. S.C., T.P., M.R., and S.F. managed data collection, curation and elaboration. A.B., A.C., F.M., and G.Z. carried out elaboration, quality control and curation of environmental data. A.G., A.P., S.C., and A.T. performed data analysis. A.G., A.T., A.C., F.M., and G.Z. wrote a first draft of the manuscript, with critical contributions from all the co-authors. All the co-authors approved the final version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41531-025-00935-y>.

Correspondence and requests for materials should be addressed to Alessandro Gialluisi.

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

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

Acknowledgements

We wish to thank all the Moli-sani Study Investigators. The enrollment phase of the Moli-sani Study was conducted at the Research Laboratories of the Catholic University in Campobasso (Italy), the follow up of the Moli-sani cohort is being conducted at the Department of Epidemiology and Prevention of the IRCCS Neuromed, Pozzilli, Italy. Steering Committee: Licia Iacoviello (Chairperson), Giovanni de Gaetano and Maria Benedetta Donati. Scientific Secretariat: Marialaura Bonaccio, Americo Bonanni, Chiara Cerletti. Simona Costanzo, Amalia De Curtis, Augusto Di Castelnuovo, Alessandro Gialluisi, Francesco Gianfagna, Mariarosaria Persichillo, Teresa Di Prospero (Secretary). Safety and Ethical Committee: Jos Vermeylen (Catholic University, Leuven, Belgium) (Chairperson), Renzo Pegoraro (Pontificia Accademia per la Vita, Rome, Italy), Antonio Spagnolo (Catholic University, Rome, Italy). External Event Adjudicating Committee: Deodato Assanelli

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025