

# SCIENTIFIC REPORTS



OPEN

## The mediating effect of immune markers on the association between ambient air pollution and adult-onset asthma

Nahid Mostafavi<sup>1</sup>, Ayoung Jeong<sup>2,3</sup>, Jelle Vlaanderen<sup>1</sup>, Medea Imboden<sup>2,3</sup>, Paolo Vineis<sup>4,5</sup>, Debbie Jarvis<sup>6</sup>, Manolis Kogevinas<sup>7</sup>, Nicole Probst-Hensch<sup>2,3</sup> & Roel Vermeulen<sup>1,5,8</sup>

We aim to investigate to what extent a set of immune markers mediate the association between air pollution and adult-onset asthma. We considered long-term exposure to multiple air pollution markers and a panel of 13 immune markers in peripheral blood samples collected from 140 adult cases and 199 controls using a nested-case control design. We tested associations between air pollutants and immune markers and adult-onset asthma using mixed-effects (logistic) regression models, adjusted for confounding variables. In order to evaluate a possible mediating effect of the full set of immune markers, we modelled the relationship between asthma and air pollution with a partial least square path model. We observed a strong positive association of IL-1RA [OR 1.37; 95% CI (1.09, 1.73)] with adult-onset asthma. Univariate models did not yield any association between air pollution and immune markers. However, mediation analyses indicated that 15% of the effect of air pollution on risk of adult-onset asthma was mediated through the immune system when considering all immune markers as a latent variable (path coefficient ( $\beta$ ) = 0.09; 95% CI: (−0.02, 0.20)). This effect appeared to be stronger for allergic asthma (22%;  $\beta$  = 0.12; 95% CI: (−0.03, 0.27)) and overweight subjects (27%;  $\beta$  = 0.19; 95% CI: (−0.004, 0.38)). Our results provides supportive evidence for a mediating effect of the immune system in the association between air pollution and adult-onset asthma.

Asthma is a chronic inflammatory disease of the airways which has a significant impact on quality of life<sup>1</sup>. Current research suggests a complex etiology for asthma<sup>2</sup>, and there is emerging evidence that air pollution is one of the important environmental factors that may be involved in disease pathogenesis<sup>3</sup>. However, previous studies are inconsistent with respect to the role of air pollution in the development of asthma in adults<sup>4,5</sup>.

The mechanism by which air pollution may lead to asthma is an active research area<sup>6</sup>. Exposure to air pollutants induces oxidative stress and activates inflammatory pathways<sup>7,8</sup>, which have also been implicated in asthma development<sup>9</sup>. Research has also indicated that interleukins (IL), tumor necrosis factor (TNF), and pro-inflammatory cytokines may have an active role in the recognition of air pollutants and subsequent inflammatory response<sup>10</sup>.

Several studies have reported on associations between long-term exposure to air pollution and chronic changes in immune markers<sup>8,11,12</sup>. For example, in the ESCAPE study, living close to busy traffic was found to be associated with increased C-reactive protein concentrations<sup>13</sup>. In another study long-term exposure to NO<sub>x</sub> was reported to be associated with decreased levels of circulating IL-8, IL-10, IL-2, and TNF- $\alpha$  concentrations<sup>8</sup>. Human airway cells or bronchial epithelial cells have been shown to release anti-inflammatory cytokines and

<sup>1</sup>Division of Environmental Epidemiology, Institute for Risk Assessment Sciences, Utrecht University, 3584 CM, Utrecht, the Netherlands. <sup>2</sup>Swiss Tropical and Public Health (TPH) Institute, Basel, Switzerland. <sup>3</sup>Department of Public Health, University of Basel, Basel, Switzerland. <sup>4</sup>Italian Institute for Genomic Medicine (IIGM), Turin, Italy. <sup>5</sup>Medical Research Council-Public Health England Centre for Environment and Health, Department of Epidemiology and Biostatistics, Imperial College London, London, United Kingdom. <sup>6</sup>Department of Public Health Sciences, King's College, London, UK. <sup>7</sup>ISGlobal, Barcelona, Spain. <sup>8</sup>Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands. Nahid Mostafavi and Ayoung Jeong contributed equally. Nicole Probst-Hensch and Roel Vermeulen jointly supervised this work. Correspondence and requests for materials should be addressed to R.V. (email: [r.c.h.vermeulen@uu.nl](mailto:r.c.h.vermeulen@uu.nl))

Characteristic	Control (N = 199)	Cases (N = 140)	Differences P-value <sup>a</sup>
<b>Sex (N sample (%))</b>			
Male	96 (48)	53 (38)	
Female	103 (52)	87 (62)	0.074
BMI (Kg/m <sup>2</sup> ; median and P <sub>25</sub> -P <sub>75</sub> )	24 (22.4, 27.2)	26 (23.1, 29.6)	0.002
Age (years; median and P <sub>25</sub> -P <sub>75</sub> )	57.1 (48.7, 64.8)	59.5 (48.6, 67.9)	0.26
<b>Education (N sample (%))</b>			
Primary school	2 (1)	4 (3)	
Secondary school, middle school, or apprenticeship	126 (63)	87 (62)	
Technical college or university	71 (36)	49 (35)	0.44
<b>Physical activity<sup>b</sup></b>			
Insufficiently active	40 (20)	38 (27)	
Sufficiently active	156 (80)	101 (72)	0.18
<b>Season (N sample)</b>			
1: spring (21/3-20/6)	52 (26)	33 (24)	
2: summer (21/6-20/9)	58 (29)	39 (28)	
3: autumn (21/9-20/12)	48 (24)	30 (22)	
4: winter (21/12-20/3)	41 (21)	37 (27)	0.63
<b>Area (N sample)</b>			
Basel	26 (13)	19 (14)	
Wald	56 (28)	19 (14)	
Davos	21 (11)	13 (9)	
Lugano	21 (11)	27 (19)	
Montana	21 (11)	19 (14)	
Payerne	12 (6)	5 (4)	
Aarau	27 (14)	20 (14)	
Geneva	15 (8)	18 (13)	0.02

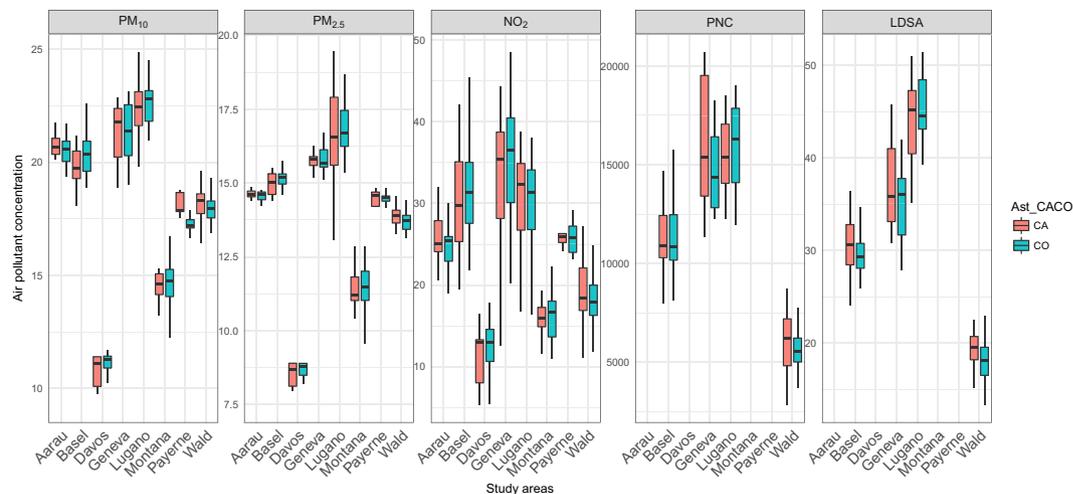
**Table 1.** Characteristics of study participants. <sup>a</sup>P-value for difference was calculated using the chi-squared test for categorical baseline variables and the student's t-test for continuous variables. <sup>b</sup>Sufficiently active: either moderate physical activity  $\geq 150$  min/week, vigorous physical activity  $\geq 60$  min/week, or combined duration (duration of moderate physical activity +  $2 \times$  duration of vigorous physical activity)  $\geq 150$  min/week; Insufficiently active: otherwise.

chemokines such as IL-10 or IL-8 upon incubation of ambient air particles<sup>14,15</sup>. Altogether, both observational and experimental studies have provided supportive evidence for the interaction between air pollutants and the immune-system in particular inflammatory pathways.

As air pollution has been linked to inflammatory responses and as these responses have been linked to adult asthma it would follow that the observed effect of air pollution on asthma is potentially mediated by the immune system. There is however a dearth of studies that has investigated this. As immune markers resembling inflammatory responses are pleiotropic it could be argued that such analyses should not be done on single immune markers but on a combined measure of the immune response. We therefore used partial least square path modeling (PLS-PM) to identify the relationship between adult-onset asthma and air pollutants and the possible mediating effect of immune markers simultaneously. PLS-PM enables us to combine information from multiple air pollution and immune markers by using a latent structure pattern for both the air pollutants (i.e. *air pollution*) and immune-markers (i.e. *immune-modulation*). We considered long-term exposure to multiple air pollution markers (particulate matter (PM) smaller than  $2.5 \mu\text{m}$  (PM<sub>2.5</sub>), and smaller than  $10 \mu\text{m}$  (PM<sub>10</sub>), nitrogen dioxide (NO<sub>2</sub>), ultra-fine particulates (UFP) [based on two metrics: particle number concentration (PNC) and lung deposited surface area (LDSA)]) and a panel of 13 immune markers in peripheral blood samples collected from 140 adult cases (asthma diagnosis after the age of 16) and 199 controls using a nested-case control design. Previous analysis in this cohort has shown an association between air pollutants and adult-onset asthma, in particular UFP<sup>16</sup>.

## Results

Our study population consists of 339 non-smoking adults (median age 58 years) from 8 areas in Switzerland of whom 140 participants developed asthma after the age of 16 years old (Table 1). There were no differences between cases and controls for any of the baseline variables investigated except BMI and study area. Median BMI for cases (26 Kg/m<sup>2</sup>; [P<sub>25</sub>-P<sub>75</sub>]: [23.1, 29.6]) was higher than controls (24 Kg/m<sup>2</sup>; [P<sub>25</sub>-P<sub>75</sub>]: [22.4, 27.2]). In Fig. 1, we show the distribution characteristics of different air pollutants by study area. For all air pollutants, substantial variation was observed between (particularly PM<sub>10</sub>, PM<sub>2.5</sub>) as well as within (particularly NO<sub>2</sub>, PNC, and LDSA) areas. UFP exposures (both PNC and LDSA) were only measured and estimated in 4 of the 8 areas (Basel, Geneva,



**Figure 1.** Box plots of distribution of air pollution concentrations by case (CA; red) and control (CO; green) per study area. Each panel shows one air pollution marker;  $PM_{10}$ ,  $PM_{2.5}$  ( $PM_{2.5}$  and  $PM_{10}$  both, estimated from the PolluMap dispersion models;  $\mu\text{g}/\text{m}^3$ ),  $\text{NO}_2$  (estimated from LUR model;  $\mu\text{g}/\text{m}^3$ ), PNC (particle number concentration; particles/ $\text{cm}^3$ ), and LDSA (lung deposited surface area;  $\mu\text{m}^2/\text{cm}^3$ ). Horizontal lines correspond to medians, and boxes to the 25th–75th percentiles; whiskers extend to data within the interquartile range times 1.5.

Air pollution metric <sup>a</sup>	Adjusted for BMI		Not adjusted for BMI	
	P-value	OR <sup>b</sup> [95% CI]	P-value	OR <sup>b</sup> [95% CI]
$PM_{10}$	0.83	1.17 [0.29, 4.68]	0.36	1.62 (0.58, 4.71)
$PM_{2.5}$	0.88	1.13 [0.24, 5.33]	0.41	1.63 (0.51, 5.35)
$\text{NO}_2$	0.37	1.37 [0.69, 2.76]	0.09	1.65 (0.94, 2.94)
PNC	0.001	3.23 [1.62, 6.43]	0.001	3.18 (1.65, 6.40)
LDSA	0.0002	5.22 [2.16, 12.6]	0.0001	5.26 (2.28, 12.74)

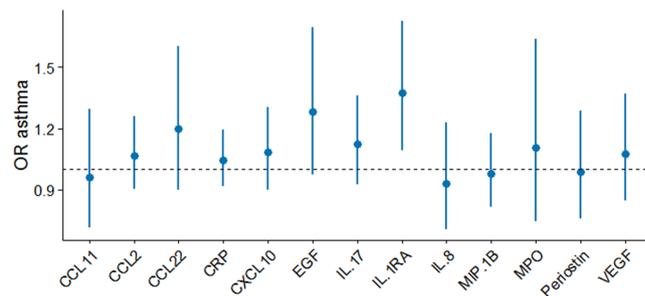
**Table 2.** Association between air pollution and adult-onset of asthma using random-effect logistic regression analysis. Note: models adjusted for age, sex, education level, and study area as random effect. <sup>a</sup>All air pollution metrics have been natural log-transformed (N = 338 for  $PM_{10}$ , N = 339 for  $\text{NO}_2$  and  $PM_{2.5}$ ; N = 189 for PNC and LDSA). <sup>b</sup>Odds ratio (OR) for adult-onset asthma per one unite increase in the natural-logarithm of each air pollutants.

Lugano, and Wald) and were available for a subset of samples (76 cases, 113 controls). Median concentrations of air pollution were generally higher in the cities of Geneva and Lugano and were lower in the more rural and smaller communities Davos (for  $PM_{10}$ ,  $PM_{2.5}$  and  $\text{NO}_2$  concentrations) and Wald (for PNC and LDSA). We observed high correlations between air pollutants within the study population (Fig. S1) especially between PNC and LDSA ( $r = 0.95$ ) and between  $PM_{10}$  and  $PM_{2.5}$  ( $r = 0.95$ ).

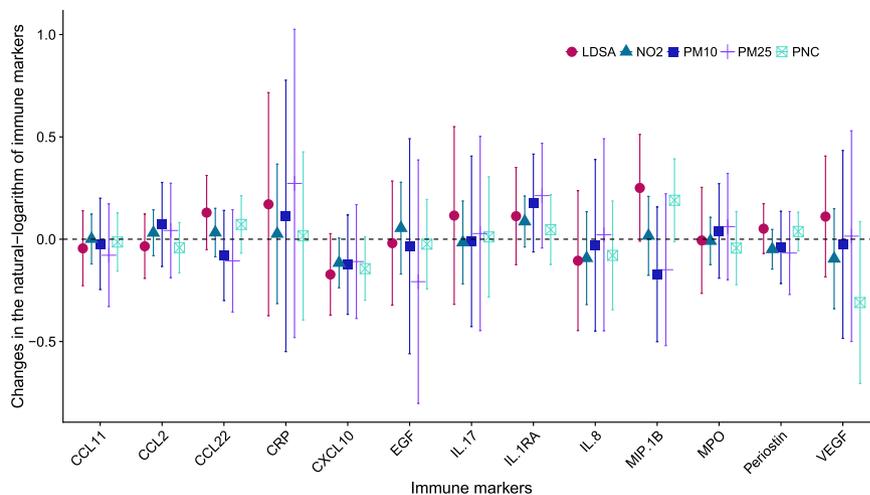
Exposure to UFP was associated with an increased risk of adult-onset asthma as previously reported (Jeong *et al.*, 2018) (Table 2). The odds ratio (OR) was 3.23 [95% CI (1.62, 6.43)] for one unit increase in natural log PNC, and 5.22 [95% CI (2.16, 12.6)] for one  $\mu\text{m}^2/\text{cm}^3$  increase in natural-logarithm of LDSA (Table 2). These associations remained similar when not correcting for the effect of BMI in the model (Table 2). The odds ratio (OR) was 3.18 [95% CI (1.65, 6.40)] for one particles/ $\text{cm}^3$  increase in natural-logarithm of PNC, and 5.26 [95% CI (2.28, 12.74)] for one  $\mu\text{m}^2/\text{cm}^3$  increase in natural-logarithm of LDSA. No associations were identified for any of the other air pollution metrics.

Higher concentrations of IL-1RA, EGF, CCL22, CCL2, CRP, CXCL10, IL-17, and MPO were associated with an increased risk of adult-onset asthma (Fig. 2, Table S1). The strongest association was seen for IL-1RA and adult-onset asthma with a corresponding OR of 1.37 [95% CI (1.09, 1.73); FDR = 0.08] for IQR (224.84 pg/mL) increase in natural-logarithm of IL-1RA. Noteworthy were the results for EGF OR = 1.28 [95% CI (0.97, 1.96)] and CCL22 OR = 1.20 [95% CI (0.9, 1.60)] which did not reach the FDR threshold (<0.2). Interestingly, CRP OR = 1.05 [95% CI (0.92, 1.19)] showed no clear association with adult-onset asthma. (Fig. 2, Table S1).

Multiple air pollutants were associated with perturbations in immune markers although none of these associations reached the threshold (FDR < 0.2) (Fig. 3). Higher concentration of IL-1RA were associated with higher exposure to all air pollutants. MIP1- $\beta$  showed positive association with PNC and LDSA, and negative association with  $PM_{2.5}$  and  $PM_{10}$ .



**Figure 2.** Odds ratios (OR) and 95%-confidence intervals for adult-onset asthma per IQR increase in the natural-log of each immune markers. IL-1RA was associated (p-value = 0.01, FDR = 0.08) with risk of adult-onset asthma.

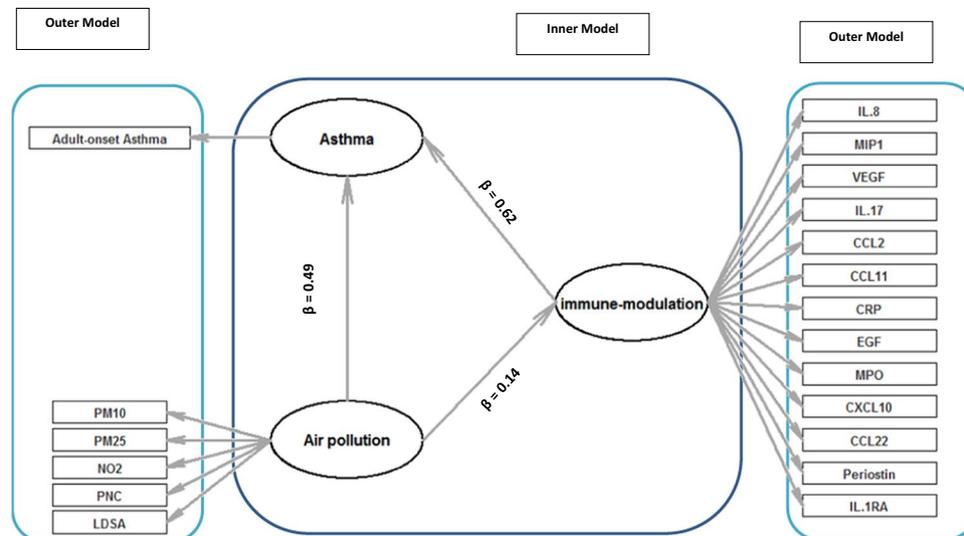


**Figure 3.** Effect estimates and 95%-confidence intervals for changes in the natural-logarithm of immune markers in per unit increase in the natural-logarithm of air pollution markers (N = 338 for PM10, N = 339 for NO<sub>2</sub> and PM<sub>2.5</sub>; N = 189 for PNC and LDSA).

**Partial least square path modeling (PLS-PM).** Figure 4, shows PLS path diagrams of the latent variables *immune-modulation* and *air pollution* for adult-onset asthma. We conducted PLS-PM on the subset of data where UFP estimates were available (76 cases, 113 controls). In the measurement (outer) model, the loading and the paths revealed a quantitative relationship between the latent and observed variables. All air pollution markers had positive effects on the latent variable *air pollution* and each explained around 86–96% variance of the corresponding latent variable (Fig. 4, Table S2). In the block of *immune-modulation*, IL-1RA, CRP, VEGF, MIP1- $\beta$ , and CCL22 explained the highest variability with a positive effect on the corresponding latent variable (each more than 40%; Fig. 4, Table S2). Latent variables *air pollution* and *immune-modulation* were constructed as linear combinations of corresponding observed variables using loadings in Table S2 as weights. In the structural (inner) model, the relationships between the latent variables were quantified with the standardized path coefficients ( $\beta$ ) (Fig. 4). *Air pollution* had a positive effect on *asthma* ( $\beta = 0.49$ ; 95% CI (0.18, 0.82) OR = 1.6; for one-unit increase in the *air pollution latent normalized structure*); Table 3) and on *immune-modulation* ( $\beta = 0.14$ ; 95% CI (-0.27, 0.32); Table 3). The association between *Immune-modulation* and *asthma* was positive ( $\beta = 0.62$ ; 95% CI (0.29, 0.98); OR = 1.9; for every one-unit increase in the *immune-modulation*; Table 3). The overall positive association of *air pollution* with *asthma* can be partially explained by *immune-modulation* (path coefficient = 0.58). This overall effect can be calculated by summing up the indirect (path coefficients =  $0.14 \times 0.62$ ) and direct (path coefficient = 0.49) effect of *air pollution* on *asthma*.

The results of the bootstrap sampling indicated that the original coefficient path value and the one obtained from the bootstrap are close (Table 3). The average coefficient of determination  $R^2$  value for two latent variables in our analysis, *immune-modulation*, and *asthma* (pseudo-Nagelkerke  $R^2$ ), was found to be 0.02 and 0.11 (Table 3) respectively. This indicated that *immune-modulation* and *air pollution* together explained 11% of variation in *asthma*.

Our results indicated that of the total effect of *air pollution* on *asthma* 15% (ratio of an indirect effect on total effect) is explained by the indirect effect of *immune-modulation* although this effect did not reach the threshold (P-value < 0.05) (path coefficient = 0.09, Z-statistics = 1.64, 95% CI (-0.02, 0.20).



**Figure 4.** Path diagram indicating the conceptual model behind the relations among latent variables and their manifest variables. Rectangles refer to the manifest variables (outer model) and the ellipses refer to the latent variables (Inner model). Arrows show assumed causations among the variables (either latent or manifest), and the direction of the arrow defines the assumed direction of the relation. Path coefficients ( $\beta$ 's) indicate the quantification of the relationship between latent variables. Corresponding ORs for the one unit increase in immune-modulation and air pollution on asthma are 1.9 and 1.6, respectively.

	Path Coefficients (Using data)	Estimated p-value	R <sup>2</sup>	Path Coefficients (Using 200 data set in Bootstrap)	SE	95 LCI	95 UCI	OR <sup>b</sup>
Direct Effect								
$\beta_{\text{Air pollution} \rightarrow \text{immune-modulation}}$	0.14	0.057	0.02	0.08	0.20	-0.27	0.32	—
$\beta_{\text{Air pollution} \rightarrow \text{Asthma}}$	0.49	0.002	0.11 <sup>a</sup>	0.49	0.16	0.18	0.82	1.6
$\beta_{\text{immune-modulation} \rightarrow \text{Asthma}}$	0.62	0.0004		0.62	0.18	0.29	0.98	1.9

**Table 3.** Partial least square path modeling analysis for the relationships between latent variables. <sup>a</sup>For logistic regression we calculated pseudo-Nagelkerke R<sup>2</sup>. <sup>b</sup>Odds ratios (OR) for adult-onset asthma per one unit increase in the corresponding latent variable.

**Sensitivity analysis.** Sensitivity analysis indicated that the associations of LDSA and PNC with adult-onset of asthma were generally robust to the set of confounders considered. The association of IL-1RA with adult-onset of asthma that we observed in the main analysis was recovered in all of the sensitivity analysis. The effect was weaker when we included BMI in the model as a confounder or in the analysis stratified by BMI status (Fig. S2B). The association between air pollution and immune markers did not change when we stratified analysis by BMI or included BMI as a confounder in the model (Fig. S3). Results of PLS-PM among the overweight subgroup indicated a stronger association between *Immune-modulation* and *asthma* ( $\beta = 0.66$ ; 95% CI (0.18, 1.18); OR = 1.93; for one-unit increase in the *immune-modulation*; Table S5) as well as between *air pollution* and *Immune-modulation*,  $\beta = 0.29$ ; 95% CI (-0.39, 0.48); Table S5), compared to the main analysis. This analysis also showed that a higher proportion (27%) of the total effect of *air pollution* on *asthma* is explained by the indirect effect of *immune-modulation* (path coefficient = 0.19, Z-statistics = 1.92, 95% CI (-0.004, 0.38)) in comparison with the main analysis (15%; 95% CI (-0.02, 0.20)).

The magnitude of the risks (ORs) of immune markers associated with adult-onset asthma was often stronger for allergic asthma than non-allergic asthma (Fig. S2C). In the allergic group, we observed a positive association of CCL22 and IL-1RA with adult-onset asthma. The odds ratios were 1.55 [95% CI 1.07, 2.23] for IQR (245.6 pg/mL) increase in natural-logarithm of CCL22, and 1.38 [95% CI 1.06, 1.80] for IQR (220.4 pg/mL) increase in natural-logarithm of IL-1RA. The associations of EGF and Periostin with adult-onset asthma were positive but weak. The odds ratios were 1.36 [95% CI 0.97, 1.90] for IQR (56.8 pg/mL) increase in natural-logarithm of EGF, and 1.34 [95% CI 0.96, 1.88] for IQR (33862.4 pg/mL) increase in natural-logarithm of Periostin. For allergic asthma, PLS-PM also indicated a stronger association between *Immune-modulation* and *asthma* (path coefficient = 0.82 (OR = 2.27; for one-unit increase in the *immune-modulation*), P-value = 0.0003; Table S4). This analysis also showed that a higher proportion (22%) of the total effect of *air pollution* on *asthma* is explained by the indirect effect of *immune-modulation* (path coefficient = 0.12, Z-statistics = 1.6, P-value = 0.1) in comparison with the main analysis (15%).

## Discussion

We investigated the relationship between exposure to air pollution, immune markers, and adult-onset asthma using data from the Swiss Cohort Study on Air Pollution and Lung and Heart Diseases in Adults (SAPALDIA). Previous studies have investigated the associations of immune markers with either air pollution or adult-onset asthma separately, which limits causal interpretation. The main aim of the current study was to better understand the biological mechanism linking air pollution exposure to adult-onset asthma and to study if immune-markers mediate this association.

We previously published on the positive association between biannual mean exposure to UFP (both PNC and LDSA) with adult-onset asthma<sup>16</sup>. In these extended analyses, we observed non-significant positive associations between PM<sub>10</sub>, PM<sub>2.5</sub> and NO<sub>2</sub> and adult-onset asthma. Additionally, we observed an increased risk of asthma with elevated concentrations of IL-1RA, and to some extent with EGF, and CCL22. Of these IL-1RA achieved Benjamini-Hochberg (BH-FDR) < 0.2 in main analysis as well as in sensitivity analyses. We observed limited evidence for a univariate association between air pollution and immune markers. IL-1RA was positively associated with all air pollutants and higher concentration of MIP1-β and CRP and lower concentration of IL-8 was observed with UFP exposures. Our study further suggested that 15% of the total effect of air pollution on adult-onset asthma is explained by the indirect effect of immune markers. These associations were robust after inclusion of different confounders.

A previous review of epidemiological studies reported that the evidence to support a causal relationship between air pollution and adult-onset asthma was inconsistent and identified that there was a need for large-scale cohort studies and the inclusion of local-scale traffic pollutants like UFP<sup>17</sup>. Although our study population is still modest in size (N = 339, Cases = 140), we were able to study its association with two metrics of UFP (LDSA and PNC). The results indicated that UFP contributes to increases in adult-onset asthma independently of other potential risk factors. The association, which we reported previously, with UFP in our study was in contrast with a recent study that assessed long-term exposure to UFP in Toronto, where a city-specific LUR model was developed and linked to health registry data of 1.1 million adult city residents. Within this study, association between UFP exposure and respiratory disease incidence, including adult-onset asthma was not reported<sup>18</sup>. A study of young adult Italians (aged 20–44 years) reported a positive (albeit weak) relationship between NO<sub>2</sub> exposure and asthma prevalence (OR 1.13, 95% CI 0.98–1.32 per 18.3 μg/m<sup>3</sup>)<sup>19</sup>. Further, a study of 12 177 Australian females reported no association between the 3-year mean annual NO<sub>2</sub> and asthma prevalence (RR: 0.97, 95% CI: 0.86–1.10)<sup>20</sup>.

There is growing evidence linking systemic inflammation in the pathogenesis of asthma<sup>9</sup>. Additionally, the higher susceptibility of individuals with asthma to air pollution may be attributable to a combination of local inflammation, and oxidative stress<sup>21</sup>. The current study investigated a relatively wide panel of inflammatory biomarkers to evaluate the effects of long-term exposure on systemic inflammatory status in adults. Our results are consistent, in terms of direction, with other studies on immune markers measured in adults showing elevated levels of CRP, and MIP1-β<sup>12,13,22</sup> and decreased levels of IL-8 and EGF<sup>8,22</sup> in relation to long-term exposure to air pollution. These markers were not noteworthy in our study. Our study also provides evidence for perturbation of plasma concentration of IL-1RA in association with exposure to air pollution in particular UFP. IL-1RA is a pleiotropic cytokine which has an anti-inflammatory effect. By suppressing the production of multiple pro-inflammatory cytokines (including TNF-α, IL-1β, IL-6, and IL-8) it regulates the progression of the immune response<sup>23,24</sup>. To date, this immune marker has not been reported in previous studies on immune markers and air pollution.

Our study shows that plasma concentrations of IL-1RA, EGF, and CCL22 were elevated in asthma patients compared to the control group. This is in line with previously reported increases in expression of IL-1RA in asthmatic airway epithelium<sup>25</sup>. Moreover, IL-1, of which IL-1RA is the receptor antagonist, is a critical mediator of the inflammatory process of multiple diseases, including asthma<sup>26</sup>. Therefore, it is reasonable to postulate that some of these markers may be responsible, at least in part, for a dysregulated inflammatory response in asthma<sup>26</sup>.

In addition to univariate regression analysis, we applied PLS path modeling that is designed to test for associations in a network of causal relationships. This approach overcomes statistical limitations by combining information from multiple variables rather than assessing them one by one. We focused on PLS path modeling in our analysis because we were interested in identifying to what extent the association between air pollution and asthma can be explained by air pollution induced immune modulation. Moreover, PLS path modeling enabled a joint assessment of the full set of immune biomarkers and air pollution exposures and does not suffer from instability in estimates and failure to converge due to multicollinearity. Our study also demonstrated that PLS path modeling is effective in assessing the total effects and relative importance of various pathways, as reflected in the latent variables, on adult-onset asthma. As univariate analysis indicated a stronger association of UFP with asthma, we also tested the mediating effect of immune markers when only using PNC and LDSA in the outer model of PLS-PM (Table S3). Interestingly, this model indicated a similar mediating effect of UFP (as measured by PNC and LDSA) on adult-onset asthma (15%). Sensitivity analyses using a two-stage regression approach to correct first for confounding factors and then apply a PLS-PM on the residuals lead to essential similar results (not shown).

Epidemiological studies have shown that obesity represents a pro-inflammatory state<sup>27</sup> which itself may increase susceptibility to air pollution by increasing the inflammatory response. Being overweight has been found to increase susceptibility to the respiratory effects of air pollution<sup>28–30</sup>. In our analysis, the effect of immune markers on asthma weakened when we corrected for the effect of BMI. As BMI may be on the causal path, inclusion of BMI in the model will in that scenario lead to an underestimation of the effect of air pollution on immune markers. We therefore put more emphasis on the non BMI corrected analyses. Interestingly, PLS mediation analysis among the overweight sub group showed a stronger mediation effect of the immune system on the relation

between air pollution and asthma (27% versus 15% in the overall analyses). This result may provide additional support to the hypothesis that obesity may indeed increase susceptibility to air pollution.

The strengths of the current study include a relatively large and well-characterized study population, individual assessment of exposure to long-term air pollution based on locally optimized models, precise measurements of plasma immune markers concentrations, and evaluation of combined mediating effect of pleiotropic immune biomarkers using PLS-PM.

Our study was also subject to some (potential) limitations. Prerequisites for mediation analysis are an association between exposure (e.g. air pollution) and outcome (e.g. asthma), exposure and mediator (e.g. an immune marker), and an association between mediator and asthma. In our analyses on an individual marker level these prerequisites were not met: there was a mis-alignment in the immune markers that were associated to air pollution and those that were associated to asthma. However, when we view the total panel of immune markers that was included in our study as a proxy for the immune system as a whole, prerequisites for mediation analysis on group level were met. Furthermore, we assumed that the weak associations that we observed between air pollution, asthma and individual immune markers was at least partly explained by measurement error of these markers. By focusing on the underlying latent structures of the immune markers, the impact of this measurement error is reduced, which likely contributed to our ability to observe an association between air pollution and asthma and the latent structure of the immune markers.

Even though we investigated a relatively large set of immune markers in relation to air pollution exposure and adult-onset asthma, these 13 immune markers may not fully represent all aspects of the immune system. As such the effect of the immune system on adult-onset asthma may be underestimated as well as its mediating effect of air pollution on asthma. Future studies should attempt to explore a broader range of inflammatory cytokines in order to expand our understanding of underlying mechanisms and the interactions between cytokines/chemokines in the air pollution/asthma inflammatory response.

This study is also limited by examining not incident but prevalent asthma cases. Asthma is a complex chronic disease that can be unnoticed for long time, grow out, and resurface, complicating accurate definition of the disease onset. Therefore this study is cross-sectional in nature, making it difficult to draw causal inference, although we restricted to the adult-onset asthma that is less susceptible to reverse causation. Additionally, the case definition of asthma in our study was based on self-report that may be prone to recall bias.

Air pollution estimates used in our analyses reflect exposure to UFP at the time of blood sample collection and therefore do not reflect a precise estimation of participants' exposure before they had developed asthma. However, work in The Netherlands has suggested that current levels of UFP are relevant for past exposures at least in relative sense<sup>31</sup>. Occupational exposure may contribute to asthma development in adults and it is conceivable that occupational asthma may constitute a distinct asthma phenotype in adults<sup>32</sup>. In our study, however, the odds of adult-onset asthma did not differ by occupational exposure to vapors, gas, dust, and fumes (Fisher's test  $p$ -value > 0.99).

## Conclusion

The current study adds to the growing body of evidence demonstrating the inflammatory effects of exposure to long-term air pollution. Furthermore, the current study provides supportive evidence for the mediatory effect of the immune system in the association between air pollutants (in particular UFP) and adult-onset asthma.

## Material and Methods

**Study population.** SAPALDIA is a population-based study that was initiated in 1991 to investigate the effect of long-term exposure to air pollution on the respiratory health of the adult Swiss population (18–60 years,  $N = 9651$ ). As previously described in detail<sup>33,34</sup>, the study comprised of eight distinct rural and urban areas covering the environmental and geographic diversity of Switzerland (Basel, Geneva, Davos, Aarau, Payerne, Montana, Wald, and Lugano). A subset of baseline participants completed two follow-up assessments in 2002–2003 (SAPALDIA2,  $n = 8047$ ) and in 2010–2011 (SAPALDIA3,  $n = 6088$ ). In both follow-up assessments blood was drawn and stored at  $-80\text{ }^{\circ}\text{C}$  in the SAPALDIA biobank. The present study included individuals selected among participants of SAPALDIA3. Information on age, sex and level of education were provided by self-reported questionnaire data collected at the study examinations. Smokers were excluded so that all participants had not smoked for at least one year before blood was drawn at SAPALDIA2.

This study complies with the Declaration of Helsinki principles and all participants provided informed consent. The overall SAPALDIA study protocol was approved by the Swiss Academy of Medical Sciences and the regional committees for each study center.

**Asthma definition.** Asthma cases ( $N = 203$ ) were defined as participants who reported asthma in SAPALDIA3. In this study we limited our analyses to adult-onset asthma, which refers to participants who reported the onset of asthma at age 16 years or older (140 out of 203 cases). Controls ( $N = 199$ ) were randomly sampled among the participants who never reported asthma or asthma related symptoms since the first follow up in SAPALDIA1. Specifically, controls were selected among those who never reported any of the following: self-reported asthma; physician-diagnosed asthma; asthma attack in the last 12 months; current asthma medication; wheezing without cold; or three or more asthma-related symptoms in the last 12 months (breathless while wheezing; attack of shortness of breath at rest or after exercise; or woken up with a feeling of chest tightness or shortness of breath).

Participants were tested for atopy at baseline via skin-prick test. Atopy was defined as developing a wheal of diameter 3 mm or larger to one or more of eight common inhalant allergens (cat, timothy grass, parietaria, birch, house-dust mite, *Alternaria tenuis*, *Cladosporium herbarum*, and dog).

**Exposure assessment.** Mean biennial exposure to UFP was estimated using multi-area land use regression (LUR) models which were derived from SAPALDIA-specific measurement campaigns conducted in 2011–2012<sup>35</sup>. Particle number concentration (PNC) and lung deposited surface area (LDSA) were used as metrics of UFP exposure and covered 4 out of 8 SAPALDIA study areas (covering 76 cases and 113 controls)<sup>35</sup>. LDSA is defined as the particle surface area concentration per unit volume of air, weighted by the deposition probability in the lung and was estimated according to the International Commission on Radiological Protection (ICRP)<sup>36</sup>. The surface area of particles is hypothesized to be a more biologically relevant metric than particle mass or number. NO<sub>2</sub> exposure was estimated by applying a recently developed European LUR model (spatial resolution 100 \* 100 m) which had been derived from measurements collected in 2010<sup>37</sup>. We also obtained annual mean estimates of PM<sub>2.5</sub> and PM<sub>10</sub> in 2010 from the PolluMap dispersion models developed by the Swiss Federal Office for the Environment (spatial resolution 200 \* 200 m)<sup>38</sup>.

**Assessment of immune markers.** Serum levels of a panel of 23 immune markers were measured in blood samples collected from all subjects<sup>39</sup>. A Luminex screening assay was performed following the manufacturer-defined protocol. The panel included interleukin (IL) 1 $\beta$ , IL-4, IL-5, IL-6, IL-8, IL-10, IL-13, IL-17, IL-25, tumor necrosis factor alpha (TNF- $\alpha$ ), C-C motif chemokine 11 (CCL11), IL1 receptor antagonist (IL1ra), CXC chemokine ligand 10 (CXCL10), epidermal growth factor (EGF), fibroblast growth factor beta (FGF- $\beta$ ), granulocyte colony-stimulating factor (G-CSF), melanoma growth stimulatory activity/growth-related oncogene (GRO), chemokine (C-C motif) ligand 2 (CCL2), C-C motif chemokine 22 (CCL22), macrophage inflammatory protein-1 beta (MIP-1 $\beta$ ), vascular endothelial growth factor (VEGF), Myeloperoxidase (MPO), and periostin. Additionally, C-reactive protein (CRP) was assessed using R&D System Solid Phase Swedish ELISA. Within each analytical batch, quality control samples were run in duplicate alongside the study samples.

For 10 immune markers (interleukin IL-1 $\beta$ , IL-4, IL-5, IL-6, IL-10, IL-13, IL-25, TNF- $\alpha$ , GRO, and FGF- $\beta$ ) the serum concentrations were below the limit of quantification (LOQ) in more than 70% of the samples. Therefore, they were excluded from analyses. Values below the LOQ for the remaining markers were imputed using a maximum likelihood estimation procedure<sup>40</sup>. To allow for variations in LOQ between plates, we imputed based on plate-specific LOQ and included the plate as a predictor variable in the imputation model. The maximum percentage of imputed samples was 35% (G-CSF) and 27% (eotaxin), while other markers had less than 4% imputed values. Immune markers were natural log-transformed to approximate a normal distribution and stabilize variance.

**Statistical analysis.** *Univariate analysis.* The heterogeneity of demographic and lifestyle variables between asthma cases and controls were compared using chi-square test. We performed random-effects logistic regressions for all air pollution and immune markers to assess the association with adult-onset asthma. Linear mixed-effects regression analysis was carried out to evaluate the relationship between immune markers and long-term exposure to air pollutants.

All analyses were adjusted for sex, age, body mass index (BMI) (kg/m<sup>2</sup>) [only for air pollution and asthma analysis], education (primary, high school, university) as fixed-effects to control for potential confounding effects. To capture heterogeneity between the study areas, city (Basel, Wald, Davos, Lugano, Montana, Payerne, Aarau, and Geneva) was included as a random-intercept in all analyses.

To capture the nuisance variation generated in the immune marker assessment, we included a random intercept for microtiter plate. Additionally, we adjusted for bench time, fasting time, Fourier-transformed venipuncture time point assuming one or two periods per day, and their multiplicative interaction terms with fasting time as potential confounders in immune marker analysis. As BMI could be on the causal path from air pollution to immune-modulation and asthma we did not adjust for BMI in the main analyses of immune markers. The set of confounders that were considered for each analysis in the current study is consistent with our previous publication<sup>16</sup> using the same case-control dataset.

Statistical analyses were conducted using R version 3.4.0 (packages: *glmer*<sup>41</sup> and *lme4*<sup>41</sup>). The resulting p-values underwent multiple testing adjustments using the false discovery rate method of Benjamini-Hochberg (BH-FDR)<sup>42</sup> at the level of 0.2. A Directed acyclic graph (DAG) visualizing the causal assumptions made in our analytical plan is provided in the Supplementary Information (Fig. S4).

*Sensitivity analysis.* To assess how sensitive our findings were to variations in the confounder model, we conducted a set of additional analysis. We ran a minimally adjusted model (only age and sex included as covariates), a model in which study area was excluded, and a model in which we added season, and physical activity (insufficiently active and active) as covariates. For immune marker analysis, we ran a model in which BMI was included as a covariate and a model in which we excluded Fourier-transformed venipuncture time point from the list of covariates. In analysis of air pollution and asthma we ran a model in which BMI was not adjusted.

Additionally, we examined potential effect modification of the associations by BMI (dichotomized into Overweight [BMI  $\geq$  25 kg/m<sup>2</sup>; N = 164], or Normal weight [18  $\leq$  BMI < 25; N = 170]; skin prick test defined as allergic (positive test [78 cases]) or Non-allergic (negative test [62 cases]) in a further sensitivity analysis, as differences in susceptibility to air pollution by BMI and difference in reaction to air pollution by atopic status has been suggested<sup>29,30</sup>. Few individuals (N = 5) in our study were underweight (BMI < 18 kg/m<sup>2</sup>) and were therefore excluded from the stratified analysis by BMI.

*Mediating effect of immune markers in the association between air pollution and asthma.* We analyzed the potential mediating effect of immune markers on the association between air pollution and adult-onset asthma by utilizing partial least-squares path modeling (PLS-PM) within the R package *pls*<sup>43</sup>. This technique relies

on taking into account all relationships among unobserved or latent variables (LV), which have each been measured by several observed indicators, generally defined as observed or manifest variables (MV). PLS-PM models are formally defined by two sets of linear equations: structural (Inner) model and measurement (Outer) model. The structural model indicates the relationships among the latent variables “*immune-modulation*”, “*air pollution*”, and “*asthma*”, which are inferred from the observed immune markers, air pollutant markers, and adult-onset asthma in this study, respectively. The measurement (outer) model indicates the relationship between the latent variables and their corresponding manifest variables. In our study, the manifest variables for the latent variable of *air pollution* are based on the six air pollution markers and observed values for the latent variable of *immune-modulation* are based on 13 immune markers. The algorithm includes two steps: first is an iterative estimation of the latent construct scores; second is the final estimation of the PLS-PM coefficients. To estimate the coefficients of the PLS-PM in the final step, we analyzed the association between *asthma* and latent variables of *air pollution* and *immune-modulation* using logistic regression. To evaluate the goodness of fit, we calculated the average coefficient of determination  $R^2$  value for two latent variables in our analysis, *immune-modulation*, and *asthma* (pseudo-Nagelkerke  $R^2$ ). The detailed methodology and algorithm can be found in previous publications<sup>44,45</sup>. In order to determine the confidence intervals for the path coefficient of *air pollution* on *immune-modulation*, a bootstrap analysis was carried out using 200 sample data sets<sup>45</sup>. Additionally, we tested the statistical significance of the indirect effect of air pollution on asthma<sup>46</sup> to establish the mediating effect of *immune-modulation*. PLS-PM does not allow for confounders to be considered in the modeling. To address this issue, we additionally employed a two-stage regression approach. First each exposure block (immune markers and air pollution) were separately regressed on potential confounders, and second, PLS-PM was fit on the obtained residuals.

We also performed sensitivity analyses. We tested the possible mediating effect of immune markers when only using UFP (both PNC and LDSA) to reflect the LV of air pollution. We also tested the mediation effect of immune markers in the subset of allergic cases and overweight individuals.

## Data Availability

Data can be provided via sending an email to the corresponding author.

## References

1. Cai, Y. *et al.* Ambient air pollution, traffic noise and adult asthma prevalence: a BioSHaRE approach. *Ambient air pollution, traffic noise adult asthma Preval. a BioSHaRE approach*, <https://doi.org/10.1183/13993003.02127-2015> (2016).
2. Wu, W., Jin, Y. & Carlsten, C. Inflammatory health effects of indoor and outdoor particulate matter. *J. Allergy Clin. Immunol.* **141**, 833–844 (2018).
3. Gilmour, M. I., Jaakkola, M. S., London, S. J., Nel, A. E. & Rogers, C. A. How exposure to environmental tobacco smoke, outdoor air pollutants, and increased pollen burdens influences the incidence of asthma. *Environmental Health Perspectives*, <https://doi.org/10.1289/ehp.8380> (2006).
4. Guarneri, M. & Balmes, J. R. Outdoor air pollution and asthma. *The Lancet*, [https://doi.org/10.1016/S0140-6736\(14\)60617-6](https://doi.org/10.1016/S0140-6736(14)60617-6) (2014).
5. Jacquemin, B. *et al.* Ambient Air Pollution and Adult Asthma Incidence in Six European Cohorts (ESCAPE). *Environ. Health Perspect.* **123**, 613–21 (2015).
6. Ji, H. *et al.* Air pollution, epigenetics, and asthma. *Allergy, Asthma and Clinical Immunology*, <https://doi.org/10.1186/s13223-016-0159-4> (2016).
7. Kelly, F. J. Oxidative stress: Its role in air pollution and adverse health effects. *Occupational and Environmental Medicine*, <https://doi.org/10.1136/oem.60.8.612> (2003).
8. Mostafavi, N. *et al.* Inflammatory markers in relation to long-term air pollution. *Environ. Int.* **81**, 1–7 (2015).
9. Wouters, E. F. M., Reynaert, N. L., Dentener, M. A. & Vernooy, J. H. J. Systemic and local inflammation in asthma and chronic obstructive pulmonary disease: is there a connection? *Proc. Am. Thorac. Soc.*, <https://doi.org/10.1513/pats.200907-073DP> (2009).
10. Vawda, S. *et al.* Associations between inflammatory and immune response genes and adverse respiratory outcomes following exposure to outdoor air pollution: A huge systematic review. *American Journal of Epidemiology*, <https://doi.org/10.1093/aje/kwt269> (2014).
11. Pope, C. A. *et al.* Exposure to Fine Particulate Air Pollution Is Associated with Endothelial Injury and Systemic Inflammation. *Circ. Res.*, <https://doi.org/10.1161/CIRCRESAHA.116.309279> (2016).
12. Hajat, A. *et al.* Long-term exposure to air pollution and markers of inflammation, coagulation, and endothelial activation a repeat-measures analysis in the multi-ethnic study of atherosclerosis (MESA). *Epidemiology*, <https://doi.org/10.1097/EDE.0000000000000267> (2015).
13. Lanki, T. *et al.* Air pollution from road traffic and systemic inflammation in adults: A cross-sectional analysis in the European escape project. *Environ. Health Perspect.*, <https://doi.org/10.1289/ehp.1408224> (2015).
14. Huang, S. L., Hsu, M. K. & Chan, C. C. Effects of submicrometer particle compositions on cytokine production and lipid peroxidation of human bronchial epithelial cells. *Environ. Health Perspect.*, <https://doi.org/10.1289/ehp.5519> (2003).
15. Alfaro-Moreno, E. *et al.* Induction of IL-6 and inhibition of IL-8 secretion in the human airway cell line Calu-3 by urban particulate matter collected with a modified method of PM sampling. *Environ. Res.*, <https://doi.org/10.1016/j.envres.2009.02.010> (2009).
16. Jeong, A. *et al.* Perturbation of metabolic pathways mediates the association of air pollutants with asthma and cardiovascular diseases. *Prep.* **119**, 334–345 (2018).
17. Jacquemin, B. *et al.* The Role of Air Pollution in Adult-Onset Asthma: A Review of the Current Evidence. *Semin. Respir. Crit. Care Med.* **33**, 606–619 (2012).
18. Weichenthal, S. *et al.* Long-term exposure to ambient ultrafine particles and respiratory disease incidence in Toronto, Canada: A cohort study. *Environ. Heal. A Glob. Access Sci. Source*, <https://doi.org/10.1186/s12940-017-0276-7> (2017).
19. de Marco, R. *et al.* The impact of climate and traffic-related NO<sub>2</sub> on the prevalence of asthma and allergic rhinitis in Italy. *Clin. Exp. Allergy*, <https://doi.org/10.1046/j.1365-2745.2002.01466.x> (2002).
20. Lazarevic, N., Dobson, A. J., Barnett, A. G. & Knibbs, L. D. Long-term ambient air pollution exposure and self-reported morbidity in the Australian Longitudinal Study on Women's Health: A cross-sectional study. *BMJ Open*, <https://doi.org/10.1136/bmjopen-2015-008714> (2015).

21. Auerbach, A & Hernandez, M. L. The effect of environmental oxidative stress on airway inflammation. *Curr. Opin. Allergy Clin. Immunol.*, <https://doi.org/10.1097/ACI.0b013e32835113d6>. The (2012).
22. Fiorito, G. *et al.* Oxidative stress and inflammation mediate the effect of air pollution on cardio- and cerebrovascular disease: A prospective study in nonsmokers. *Environ. Mol. Mutagen.*, <https://doi.org/10.1002/em.22153> (2017).
23. Dobрева, Z. G., Miteva, L. D. & Stanilova, S. A. The inhibition of JNK and p38 MAPKs downregulates IL-10 and differentially affects c-jun gene expression in human monocytes. *Immunopharmacol. Immunotoxicol.*, <https://doi.org/10.1080/08923970802626276> (2009).
24. Chung, F. Anti-inflammatory cytokines in asthma and allergy: Interleukin-10, interleukin-12, interferon- $\gamma$ . *Mediators of Inflammation*, <https://doi.org/10.1080/09629350120054518> (2001).
25. Sousa, A. R. *et al.* Effect of inhaled glucocorticoids on IL-1 beta and IL-1 receptor antagonist (IL-1 ra) expression in asthmatic bronchial epithelium. *Thorax* (1997).
26. Whelan, R. *et al.* Role and regulation of interleukin-1 molecules in pro-asthmatic sensitised airway smooth muscle. *Eur. Respir. J.*, <https://doi.org/10.1183/09031936.04.00133803> (2004).
27. Jacobsen, M. J. *et al.* Altered Methylation Profile of Lymphocytes Is Concordant with Perturbation of Lipids Metabolism and Inflammatory Response in Obesity. *J. Diabetes Res.*, <https://doi.org/10.1155/2016/8539057> (2016).
28. Miller, R. L. & Peden, D. B. Environmental effects on immune responses in patients with atopy and asthma. *Journal of Allergy and Clinical Immunology*, <https://doi.org/10.1016/j.jaci.2014.07.064> (2014).
29. Adam, M. *et al.* Adult lung function and long-term air pollution exposure. ESCAPE: a multicentre cohort study and meta-analysis. *Eur Respir J*, <https://doi.org/10.1183/09031936.00130014> (2015).
30. Schikowski, T. *et al.* Improved air quality and attenuated lung function decline: Modification by obesity in the SAPALDIA cohort. *Environ. Health Perspect.*, <https://doi.org/10.1289/ehp.1206145> (2013).
31. Downward, G. S. *et al.* Long-Term Exposure to Ultrafine Particles and Incidence of Cardiovascular and Cerebrovascular Disease in a Prospective Study of a Dutch Cohort. *Environ. Health Perspect.*, <https://doi.org/10.1289/EHP3047> (2018).
32. de Nijs, S. B., Venekamp, L. N. & Bel, E. H. Adult-onset asthma: Is it really different? *European Respiratory Review*, <https://doi.org/10.1183/09059180.00007112> (2013).
33. Martin, B. W. *et al.* SAPALDIA: Methods and participation in the cross-sectional part of the Swiss Study on Air Pollution and Lung Diseases in Adults. *Soz. Praventivmed.* **42**, 67–84 (1997).
34. Ackermann-Lieblich, U. *et al.* Follow-up of the Swiss Cohort Study on Air Pollution and Lung Diseases in Adults (SAPALDIA 2) 1991–2003: methods and characterization of participants. *Soz. Praventivmed.* **50**, 245–63 (2005).
35. Eeftens, M. *et al.* Development of land use regression models for nitrogen dioxide, ultrafine particles, lung deposited surface area, and four other markers of particulate matter pollution in the Swiss SAPALDIA regions. *Environ. Heal.* **15**, 53 (2016).
36. Fierz, M., Houle, C., Steigmeier, P. & Burtscher, H. Design, calibration, and field performance of a miniature diffusion size classifier. *Aerosol Sci. Technol.*, <https://doi.org/10.1080/02786826.2010.516283> (2011).
37. de Hoogh, K. *et al.* Development of West-European PM<sub>2.5</sub> and NO<sub>2</sub> land use regression models incorporating satellite-derived and chemical transport modelling data. *Environ. Res.*, <https://doi.org/10.1016/j.envres.2016.07.005> (2016).
38. Liu, L.-J. S. *et al.* Characterization of source-specific air pollution exposure for a large population-based Swiss cohort (SAPALDIA). *Environ. Health Perspect.* **115**, 1638–45 (2007).
39. Vlaanderen, J. *et al.* Acute changes in serum immune markers due to swimming in a chlorinated pool. *Environ. Int.* **105**, 1–11 (2017).
40. Lubin, J. H. *et al.* Epidemiologic evaluation of measurement data in the presence of detection limits. *Environ. Health Perspect.* **112**, 1691–1696 (2004).
41. Boeck, P. D. *et al.* The Estimation of Item Response Models with the lmer Function from the lme4 Package in R. *J. Stat. Softw.* **39**, 1–28 (2011).
42. Benjamini, Y. & Controlling, Y. H. The False Discovery Rate: a Practical and Powerful Approach to Multiple Testing When researchers tend to select pursuing multiple the (statistically) and support of conclusions. *An unguarded use in a greatly results of single-inference inc.* **57**, 289–300 (1995).
43. Sanchez, G. *PLS Path Modeling with R*. (Berkeley, CA, USA, 2013).
44. Petrarca, F., Russolillo, G. & Trincherà, L. In *Partial Least Squares Path Modeling* pp 259–279, <https://doi.org/10.1007/978-3-319-64069-3> (Springer, Cham, 2017).
45. Tenenhaus, M., Vinzi, V. E., Chatelin, Y.-M. & Lauro, C. PLS path modeling. *Comput. Stat. Data Anal.* **48**, 159–205 (2005).
46. Helm, S., Eggert, A. & Garnefeld, I. In *Handbook of Partial Least Squares*, [https://doi.org/10.1007/978-3-540-32827-8\\_23](https://doi.org/10.1007/978-3-540-32827-8_23) (2010).

## Acknowledgements

This research has received funding from the European Community's Seventh Framework Program (FP7/2007e2011) under Grant Agreement Number: 308610 (Exposomics). The SAPALDIA cohort and biobank is funded by Swiss National Science Foundation Grant Number: 33CS30-148470/1&2 to NPH. The authors thank George Downward for language editing of this manuscript.

## Author Contributions

Nahid Mostafavi contributed to data cleaning, data analysis, interpreting the data and wrote the final manuscript; Roel Vermeulen contributed to data interpretation, editing the manuscript and supervised the study; Nicole Probst-Hensch contributed to the study design, edited the manuscript; Jelle Vlaanderen, contributed to data interpreting and editing the manuscript; Paolo Vineis contributed to editing the manuscript; Ayoung Jeong contributed to part of the data cleaning and data collection; Medea Imboden, Debbie Jarvis, Manolis Kogevinas, contributed to data collection.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-019-45327-4>.

**Competing Interests:** The authors declare no competing interests.

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



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, 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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019