



# The association between air pollution and the incidence of idiopathic pulmonary fibrosis in Northern Italy

Sara Conti <sup>1</sup>, Sergio Harari<sup>2</sup>, Antonella Caminati<sup>2</sup>, Antonella Zanobetti<sup>3</sup>, Joel D. Schwartz<sup>3</sup>, Pietro A. Bertazzi<sup>4</sup>, Giancarlo Cesana<sup>1</sup> and Fabiana Madotto<sup>1</sup>

**Affiliations:** <sup>1</sup>Research Centre on Public Health, Dept of Medicine and Surgery, University of Milano-Bicocca, Milan, Italy. <sup>2</sup>Unità Operativa di Pneumologia e Terapia Semi-Intensiva Respiratoria, Servizio di Fisiopatologia Respiratoria ed Emodinamica Polmonare, Ospedale San Giuseppe-MultiMedica IRCCS, Milan, Italy. <sup>3</sup>Dept of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, MA, USA. <sup>4</sup>Dept of Preventive Medicine, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy.

**Correspondence:** Sergio Harari, UO di Pneumologia e Terapia Semi-Intensiva Respiratoria-Servizio di Fisiopatologia Respiratoria ed Emodinamica Polmonare, Ospedale San Giuseppe-Multimedica IRCCS, Via San Vittore, 12, I-20123 Milan, Italy. E-mail: sharari@hotmail.it



@ERSpublications

**The inhalation of traffic-related pollutants, namely NO<sub>2</sub>, might be involved in the development of IPF**  
<http://ow.ly/qfLE30ge36C>

**Cite this article as:** Conti S, Harari S, Caminati A, *et al.* The association between air pollution and the incidence of idiopathic pulmonary fibrosis in Northern Italy. *Eur Respir J* 2018; 51: 1700397 [<https://doi.org/10.1183/13993003.00397-2017>].

**ABSTRACT** Acute exacerbations and worsening of idiopathic pulmonary fibrosis (IPF) have been associated with exposure to ozone (O<sub>3</sub>), nitrogen dioxide (NO<sub>2</sub>) and particulate matter, but chronic exposure to air pollution might also affect the incidence of IPF. We investigated the association between chronic exposure to NO<sub>2</sub>, O<sub>3</sub> and particulate matter with an aerodynamic diameter <10 µm (PM<sub>10</sub>) and IPF incidence in Northern Italy between 2005 and 2010.

Daily predictions of PM<sub>10</sub> concentrations were obtained from spatiotemporal models, and NO<sub>2</sub> and O<sub>3</sub> hourly concentrations from fixed monitoring stations. We identified areas with homogenous exposure to each pollutant. We built negative binomial models to assess the association between area-specific IPF incidence rate, estimated through administrative databases, and average overall and seasonal PM<sub>10</sub>, NO<sub>2</sub>, and 8-hour maximum O<sub>3</sub> concentrations.

Using unadjusted models, an increment of 10 µg·m<sup>-3</sup> in NO<sub>2</sub> concentration was associated with an increase between 7.93% (95% CI 0.36–16.08%) and 8.41% (95% CI –0.23–17.80%) in IPF incidence rate, depending on the season. After adjustment for potential confounders, estimated effects were similar in magnitude, but with larger confidence intervals.

Although confirmatory studies are needed, our results trace a potential association between exposure to traffic pollution and the development of IPF.

---

This article has supplementary material available from [erj.ersjournals.com](http://erj.ersjournals.com)

Received: Feb 24 2017 | Accepted after revision: Oct 23 2017

**Support Statement:** This study was supported by an unrestricted grant by Hoffmann–La Roche Std. The sponsor had no role in the study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the report for publication. Funding information for this article has been deposited with the Crossref Funder Registry.

**Conflict of interest:** Disclosures can be found alongside this article at [erj.ersjournals.com](http://erj.ersjournals.com)

Copyright ©ERS 2018

## Introduction

Idiopathic pulmonary fibrosis (IPF) is a chronic and progressive fibrotic lung disease with a severe prognosis and a median survival of 3–5 years from diagnosis [1].

Although the mechanisms that cause the development of IPF are still under debate [1, 2], there is evidence that both genetic and environmental factors, such as cigarette smoking [3, 4], occupational exposure to metal and wood dusts [5–7] and exposure to viruses [8], such as Epstein–Barr virus, may play a role [9–12].

JOHANSSON *et al.* [13] observed an association between the risk of acute exacerbation of IPF and ozone (O<sub>3</sub>) and nitrogen dioxide (NO<sub>2</sub>), and two recent studies observed that exposure to particulate matter is associated with increased mortality and decreased forced vital capacity in IPF patients [14, 15], suggesting that air pollution may contribute to worsening disease.

These studies and the large body of evidence supporting the short- and long-term effects of air pollution on respiratory diseases, including chronic obstructive pulmonary disease, asthma and lung cancer [16–21], point to the need for further investigation into the relationship between air pollution and IPF [22]. Such studies are specifically important given the recent translational data that identify air pollution as a trigger for telomere shortening and accelerated senescence, surfactant abnormalities, stress of the endoplasmic reticulum, oxidant–antioxidant dysregulation and chronic inflammation. These mechanisms all potentially cause premature exhaustion and abnormal alveolar re-epithelialisation, which are mechanisms responsible for the development of IPF [23–27].

The aim of this study was to investigate the long-term relationship between exposure to three criteria pollutants, namely particulate matter with an aerodynamic diameter <10 µm (PM<sub>10</sub>), NO<sub>2</sub> and O<sub>3</sub>, and the incidence of IPF in Lombardy, Northern Italy, from 2005 to 2010.

Lombardy is characterised by high air pollution levels, heterogeneous among subregional areas. The stagnation of air pollutants over its lowland areas [28, 29] is facilitated by weather, climate and orographic characteristics (figure 1) which cause frequent thermal inversion. In a prior study [30], among the ~10 million inhabitants of the region, we identified >2000 IPF incident cases from 2005 to 2010. The number of IPF cases, together with the variability of environmental exposures makes Lombardy an interesting context for this analysis.

## Material and methods

### Outcome assessment

To identify IPF incident cases, we used the broad case definition (BCD), as described in our previous study [30]: we relied on regional healthcare administrative databases in order to identify all patients who had at least one hospitalisation or outpatient visit with a diagnosis of IPF (International Classification of Diseases, ninth revision, clinical modification (ICD-9-CM) code 516.3 [31]) between 2005 and 2010, and we used the date of the first traceable diagnosis as a proxy for the disease onset. We excluded all patients who had at least one visit with a diagnosis of another interstitial lung disease (ICD-9-CM codes 135, 237.7, 272.7, 277.3, 277.8, 446.21, 446.4, 495, 500–505, 506.4, 508.1, 508.8, 515, 516.0–516.2, 516.8–516.9, 517.2, 517.8, 518.3, 555, 710.0–710.4, 714.81, 720.0 and 759.5) after onset of IPF, and in order to distinguish incident from prevalent cases, we required a 5-year washout period. For each case we had information on the municipality of residence.

### Exposure assessment

We retrieved daily predictions of PM<sub>10</sub> concentrations in Lombardy on a 10×10 km grid, for the years 2000–2009, from a model based on satellite aerosol optical depth measurements and land-use regression variables [28].

For each municipality, we estimated the yearly average PM<sub>10</sub> concentration (online supplementary material).

We obtained hourly NO<sub>2</sub> and O<sub>3</sub> concentrations measured from 2000 to 2010 at background and traffic monitoring stations [32] that fulfilled predefined completeness criteria (online supplementary material) from ARPA (Regional Environmental Protection Agency) of Lombardy (<http://ita.arpalombardia.it/ita/index.asp>).

Population average exposures are usually assessed based on background stations [33]; however, it is well known that NO<sub>2</sub> is a proxy for traffic-related exposure [18], therefore we decided to also rely on traffic monitors. For each municipality, we estimated the long-term exposure to NO<sub>2</sub> by averaging measures from all background monitors located within 10 km of the municipality limits and all traffic monitors located within 5 km.

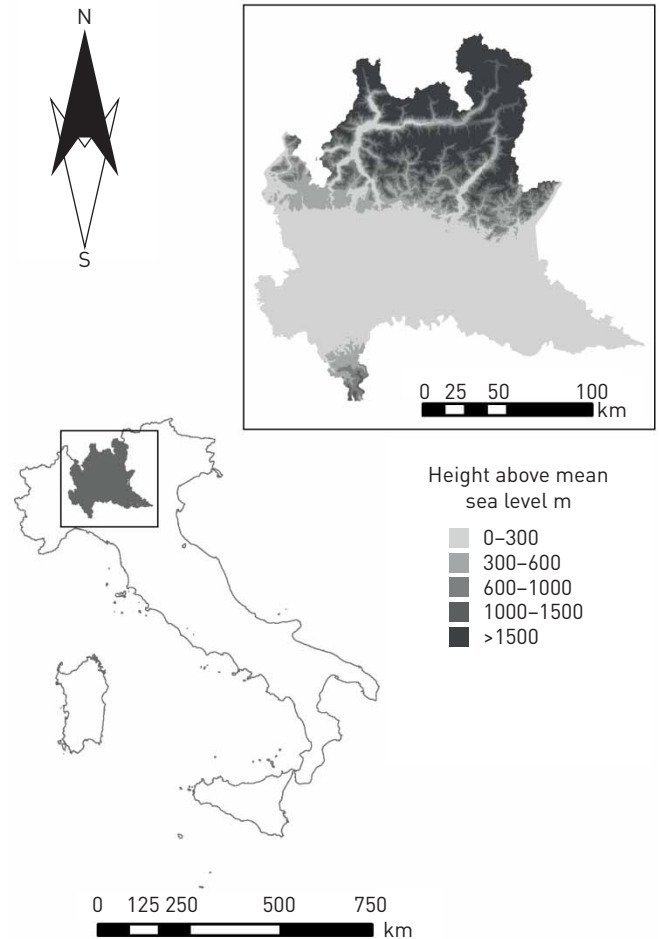


FIGURE 1 Location of the study area. Elevation raster data of Lombardy downloaded from the website of the network of the Italian national environmental information system ([www.sinanet.isprambiente.it/it/sia-ispra/download-mais/dem75/view](http://www.sinanet.isprambiente.it/it/sia-ispra/download-mais/dem75/view)). Shapefile of the Italian administrative regions downloaded from the website of the Italian National Institute of Statistics ([www.istat.it/it/archivio/24613](http://www.istat.it/it/archivio/24613)).

For  $O_3$ , we considered only background monitors, and given that it is a seasonal pollutant with very low concentrations during the cold season [34], we estimated the municipality-specific yearly average 8-hour maximum concentration of  $O_3$  over the warm season, based on measures from monitors located within 10 km of the municipality limits.

After evaluating regional coverage and average concentrations over different time windows, we decided to use the average concentrations over 2005–2010 for  $NO_2$  and  $O_3$ , and over 2005–2009 for  $PM_{10}$  as proxies for chronic exposures (online supplementary material).

#### ***Assessment of confounders***

Age is a well-known risk factor for IPF and some studies highlighted that sex might also play a role [30, 35]. In order to account for potential confounding due to these two factors, we retrieved the age and sex distribution of the population on January 1, 2005 and 2011 for each municipality from the Italian National Institute of Statistics (ISTAT; [www.istat.it/it/](http://www.istat.it/it/)).

Since temperature is another known confounder of the short-term relationship between respiratory morbidity and air pollution [29], we obtained the monthly average temperatures predicted at a  $1 \times 1$  km grid by the SCIA project ([www.scia.isprambiente.it/home\\_new.asp](http://www.scia.isprambiente.it/home_new.asp)) developed by the Italian Institute for Environmental Protection and Research (ISPRA). The 2005–2010 municipality-specific average temperatures were assessed using the same methodology used for  $PM_{10}$  (online supplementary material).

In order to account for the potential confounding due to socioeconomic status, we retrieved data on the municipality-specific average per-capita income from 2005 to 2010 from the website of the Italian Ministry of Economy and Finance ([www1.finanze.gov.it/finanze2/pagina\\_dichiarazioni/dichiarazioni.php](http://www1.finanze.gov.it/finanze2/pagina_dichiarazioni/dichiarazioni.php)).

Finally, since there is evidence of a relationship between IPF incidence and exposure to farming or livestock [7], for each municipality we computed the area covered by natural, agricultural and urban/industrial lands, based on data estimated for 2006 through the CORINE land cover model [36] available on the ISPRA website ([www.sinanet.isprambiente.it/it/sia-ispra/download-mais/corine-land-cover/corine-land-cover-2006/view](http://www.sinanet.isprambiente.it/it/sia-ispra/download-mais/corine-land-cover/corine-land-cover-2006/view)).

### *Statistical analysis*

We observed that the distribution of the municipality-specific 2005–2010 cumulative count of IPF incident cases followed a negative binomial distribution, with a high proportion of zeroes, due to the presence of many municipalities with a small population. In order to reduce this proportion, we stratified the distribution of each pollutant in classes of  $2 \mu\text{g}\cdot\text{m}^{-3}$  for  $\text{PM}_{10}$  and  $\text{NO}_2$ , and 2 ppb for  $\text{O}_3$ .

For each pollutant, we first aggregated adjacent municipalities belonging to the same class, creating new areas with bigger populations. Then, for each aggregated area, we used the municipality-specific data to compute cumulative number of IPF incident cases, mean age, percentage of males, average per-capita income, main use of the land (natural, agricultural or urban/industrial), average temperature and average exposure to the pollutant of interest over the study period. Finally, we applied negative binomial regression models and investigated the relationship between the logarithm of the outcome of interest and predictors [37].

For each pollutant, we built an unadjusted model accounting for no confounders, a fully adjusted model, accounting for all available confounders, and the best model, chosen based on the Akaike information criterion, including only confounders that improved the model fitting (online supplementary material).

Results are reported as percentage change in the incidence rate, with associated 95% confidence intervals for a 10-unit change in pollutant concentrations.

There is evidence of seasonal differences in the short-term effects of  $\text{PM}_{10}$  and  $\text{NO}_2$ . This could be due to differential sources in the case of  $\text{PM}_{10}$ , the presence of unaccounted effect modifiers or differences in the accuracy of measured concentrations [38]. In order to verify if such differences persist when addressing long-term effects, analyses were repeated considering the average warm (April–September) and cold (October–March) season concentrations of these two pollutants.

All statistical analyses were performed using SAS version 9.4 (SAS Institute, Cary, NC, USA) and R version 3.1.3 (R Project for Statistical Computing, [www.R-project.org](http://www.R-project.org)); maps were created using ESRI ArcMAP 10.5 (ESRI, Redlands, CA, USA).

### *Sensitivity analysis*

In order to test how the inclusion of  $\text{NO}_2$  traffic monitors affected the estimated municipality-specific exposures, we compared these measures with those assessed based only on background monitors located within 10 km of the city limits.

Furthermore, we tested whether our results were robust to case definition by repeating the analysis using two alternative definitions: one less stringent (generic case definition), and one more stringent (narrow case definition) (online supplementary material).

Given the results observed for  $\text{NO}_2$ , as a descriptive analysis we defined two macro-areas based on the average 2005–2010  $\text{NO}_2$  concentration using a cut-off of  $40 \mu\text{g}\cdot\text{m}^{-3}$ , which corresponds to the European Union (EU) yearly average standard (<http://ec.europa.eu/environment/air/quality/standards.htm>). We then compared age- and sex-adjusted IPF incidence rates between the two areas. The same analysis was repeated using a cut-off of  $20 \mu\text{g}\cdot\text{m}^{-3}$ , the limit recommended by the World Health Organization [39].

## **Results**

In our previous study [30], we identified 2093 incident cases satisfying the BCD, distributed among 1545 municipalities (online supplementary table S1). We identified a mean $\pm$ SD  $1.4\pm 11.3$  cases per municipality, but intermunicipality variability was very high, as shown by the coefficient of variation, which was 8.4. This was due to the variability of the municipality-specific population:  $6249\pm 34855$  inhabitants with a coefficient of variation of 5.6.

Average exposures over the time windows are reported in online supplementary tables S2–S4. Using the window 2005–2010, we could assign  $\text{PM}_{10}$  exposure to 99.1% of the municipalities,  $\text{NO}_2$  exposure to 68.7% and  $\text{O}_3$  exposure to 57.7%. The location of monitoring stations selected for the analysis and the average municipality-specific exposure are reported in online supplementary figures S1–S5.

The aggregation procedure significantly lowered the number of areas with no IPF cases: the percentage of zeroes fell from 56.6% to 36.2% for PM<sub>10</sub>, from 51.1% to 21.5% for NO<sub>2</sub> and from 49.8% to 12.5% for O<sub>3</sub> (online supplementary figures S6–S8).

In table 1, we report the description of outcomes and confounders after the aggregation process for each pollutant: we identified 398 areas based on PM<sub>10</sub> concentration, 303 based on NO<sub>2</sub> and 112 based on O<sub>3</sub>. The average number of IPF cases per area was five for the aggregation based on PM<sub>10</sub>, six for NO<sub>2</sub> and 15 for O<sub>3</sub>.

The 2005–2009 average PM<sub>10</sub> concentration at an aggregated area level (table 2) was 39 µg·m<sup>-3</sup>, the 2005–2010 average NO<sub>2</sub> concentration was 41 µg·m<sup>-3</sup> and the 2005–2010 average 8-hour maximum O<sub>3</sub> concentration during the warm season was 107 ppm. Estimated PM<sub>10</sub> and NO<sub>2</sub> concentrations were significantly higher during the cold season. Figures 2–4 report the estimated pollutant concentrations in the aggregated areas covered by the exposure assessment, clearly showing the heterogeneity of the levels registered over the study area.

Table 3 presents the results of the association between incidence of IPF and chronic exposure to each pollutant. The incidence of IPF was not associated with the concentration of PM<sub>10</sub> and O<sub>3</sub>.

As for the effect of NO<sub>2</sub>, in unadjusted models we detected a significant association between the incidence of IPF and cold season NO<sub>2</sub> concentration: an increment of 10 µg·m<sup>-3</sup> in NO<sub>2</sub> was associated with a

TABLE 1 Distribution of the outcome and of the potential confounders in the aggregated areas

	Aggregation based on PM <sub>10</sub>	Aggregation based on NO <sub>2</sub>	Aggregation based on O <sub>3</sub>
<b>Areas</b>	398	303	112
<b>IPF cases with assessed exposure</b>	2090	1842	1702
<b>IPF cases per area</b>			
Mean±SD	5±24.0	6±26.0	15±42.8
Median (IQR)	1 [0–3]	2 [1–5]	4.5 [1–15]
Range	0–451	0–438	0–433
Coefficient of variation	4.66	4.27	2.82
<b>Average population per area 1000s</b>			
Mean±SD	24.0±82.2	28.1±80.6	70.0±136.8
Median (IQR)	5.9 [2.2–16.2]	14.7 [6.1–29.3]	30.3 [12.8–86.8]
Range	0.04–1353.85	0.13–1322.34	1.68–1311.78
Coefficient of variation	3.40	2.87	1.96
<b>Males per area %</b>			
Mean±SD	49.5±1.2	49.3±0.9	49.2±0.6
Median (IQR)	49.4 [48.8–50.1]	49.3 [48.8–49.8]	49.2 [48.9–49.7]
Range	45.3–53.7	46.4–54.1	47.3–51.0
Coefficient of variation	0.02	0.02	0.01
<b>Mean age per area years</b>			
Mean±SD	42.8±2.8	41.8±1.9	41.9±1.4
Median (IQR)	42.4 [40.9–44.1]	41.6 [40.6–42.8]	41.9 [41.1–42.7]
Range	37.2–58.2	35.9–50.7	38.3–46.0
Coefficient of variation	0.07	0.05	0.03
<b>Average income per area €1000</b>			
Mean±SD	18.9±3.1	20.6±2.7	20.9±2.4
Median (IQR)	18.9 [16.9–20.9]	20.5 [19.3–21.7]	21.1 [19.6–22.3]
Range	10.0–33.6	11.6–37.3	14.6–29.6
Coefficient of variation	0.17	0.13	0.11
<b>Main land use within the area n (%)</b>			
Natural	187 [46.98]	52 [17.16]	18 [16.07]
Agricultural	171 [42.96]	196 [64.69]	68 [60.71]
Urban and industrial	40 [10.05]	55 [18.15]	26 [23.21]
<b>Average temperature °C</b>			
Mean±SD	11±2.8	13±1.4	12±1.9
Median (IQR)	13 [9.9–13.1]	13 [12.4–13.2]	13 [12.6–13.2]
Range	0–14	2–14	2–14
Coefficient of variation	0.26	0.11	0.15

Data are presented as n, unless otherwise stated. PM<sub>10</sub>: particulate matter with an aerodynamic diameter <10 µm; NO<sub>2</sub>: nitrogen dioxide; O<sub>3</sub>: ozone; IPF: idiopathic pulmonary fibrosis; IQR: interquartile range.

TABLE 2 Distributions of daily concentrations of air pollutants at an aggregated area level over the study period

	Overall	Warm season	Cold season
<b>Average PM<sub>10</sub> concentration<sup>#</sup> <math>\mu\text{g}\cdot\text{m}^{-3}</math></b>			
Mean $\pm$ SD	39 $\pm$ 4.3	27 $\pm$ 3.1	51 $\pm$ 5.8
Median (IQR)	40 (34.9–42.1)	27 (24.0–29.4)	52 (46.1–55.3)
Range	30–50	20–34	40–65
<b>Average NO<sub>2</sub> concentration<sup>¶</sup> <math>\mu\text{g}\cdot\text{m}^{-3}</math></b>			
Mean $\pm$ SD	41 $\pm$ 9.0	30 $\pm$ 8.1	53 $\pm$ 10.2
Median (IQR)	41 (35.2–47.7)	29 (24.3–34.6)	52 (45.5–60.5)
Range	16–65	9–55	24–74
<b>Average O<sub>3</sub> concentration<sup>¶</sup> ppm</b>			
Mean $\pm$ SD		107 $\pm$ 6.8	
Median (IQR)		107 (101.2–112.8)	
Range		91–120	

PM<sub>10</sub>: particulate matter with an aerodynamic diameter <10  $\mu\text{m}$ ; NO<sub>2</sub>: nitrogen dioxide; O<sub>3</sub>: ozone; IQR: interquartile range. <sup>#</sup>: 2005–2009; <sup>¶</sup>: 2005–2010.

7.93% (95% CI 0.36–16.08%) increase in the incidence rate of IPF; in addition, we observed a marginally significant ( $p < 0.10$ ) association with incidence rate (95% CI) for the overall (8.41%,  $-0.23$ – $17.80\%$ ) and the warm season (8.22%,  $-1.44$ – $18.82\%$ ) concentrations. A marginally significant association with the cold season concentration persisted when using the model with the best fit (7.55%,  $-0.76$ – $16.56\%$ ). Even though the estimated effect did not reach statistical significance, point estimates were quite robust to adjustment, and only the width of the confidence interval varied (figure 5). This was confirmed by the sensitivity analyses involving cases identified using the generic and narrow case definitions (online supplementary figures S9 and S10).

Finally, descriptive analyses on adjusted rates confirmed that IPF incidence rates are significantly higher where the average concentration of NO<sub>2</sub> is  $>40 \mu\text{g}\cdot\text{m}^{-3}$ . In addition, an increase emerges using a  $20 \mu\text{g}\cdot\text{m}^{-3}$  cut-off, but it is not significant due to the limited size of the population exposed to a concentration below this level (online supplementary figures S11 and S12).

## Discussion

In our study, we investigated the association between the incidence of IPF and chronic exposure to PM<sub>10</sub>, NO<sub>2</sub> and O<sub>3</sub> in Northern Italy. While no association was detected with PM<sub>10</sub> and O<sub>3</sub>, we observed a

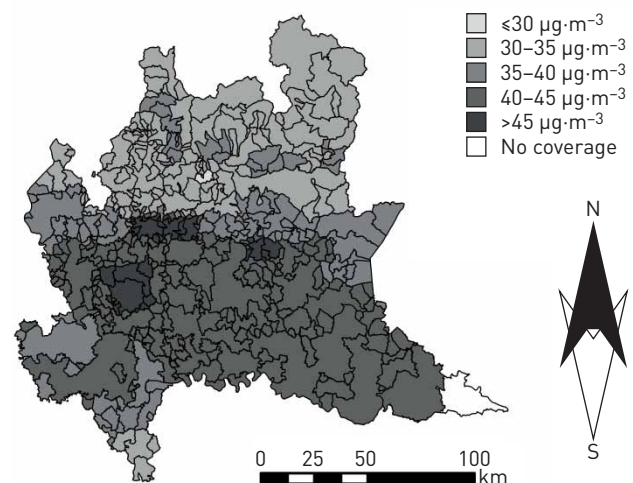


FIGURE 2 Map of the estimated daily average particulate matter with an aerodynamic diameter <10  $\mu\text{m}$  concentration for the period 2005–2009 at an aggregated area level. Shapefile of the Italian administrative regions and municipalities downloaded from the website of the Italian National Institute of Statistics ([www.istat.it/it/archivio/24613](http://www.istat.it/it/archivio/24613)).

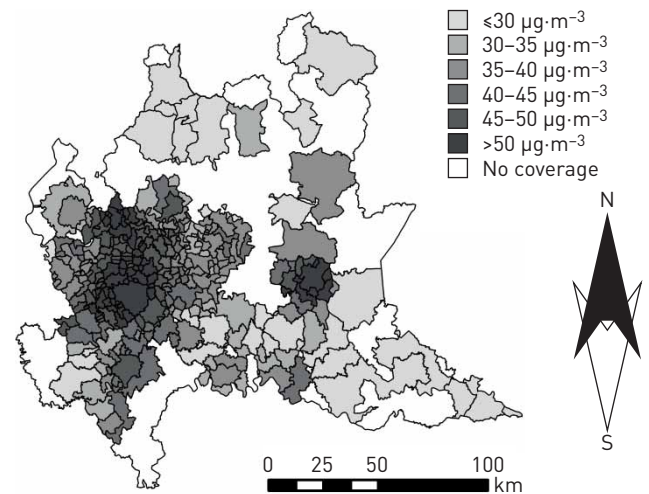


FIGURE 3 Map of the estimated daily average nitrogen dioxide concentration for the period 2005–2010 at an aggregated area level. Shapefile of the Italian administrative regions and municipalities downloaded from the website of the Italian National Institute of Statistics ([www.istat.it/it/archivio/24613](http://www.istat.it/it/archivio/24613)).

positive, although often marginally significant, association with overall and cold season  $\text{NO}_2$  concentration.

Although estimated effects do not always reach statistical significance, they should be interpreted as a strong signal that the association exists: point estimates are quite robust, as they are moderately affected by different model specifications and case definitions; only precision decreases when using adjusted models, as confirmed by wider confidence intervals (figure 5 and online supplementary figures S9 and S10). This probably means that the model power is insufficient to estimate a significant effect of the pollutant while simultaneously accounting for all confounders of interest. If we could assess the exposure for the whole region, instead of covering only 68% of the municipalities, the number of areas included in the models would grow, as would the power of the model. This might also be true for  $\text{O}_3$ . Furthermore, the descriptive analysis of adjusted rates confirmed that IPF incidence is significantly higher in areas where the average  $\text{NO}_2$  concentration is above the EU limit (online supplementary figures S11 and S12).

Finally, effects estimated using the cold season concentrations were similar, but slightly more precise than those of the overall analyses (figure 5 and online supplementary figures S9 and S10). These results suggest that the cold season concentrations better represent the overall real exposure. Indeed, instruments used in ARPA stations to monitor  $\text{NO}_2$  have been shown to overestimate concentrations during the warm season,

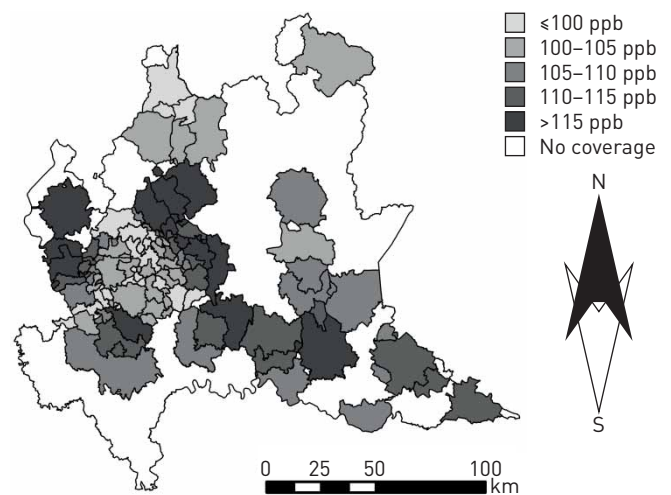


FIGURE 4 Map of the estimated daily average 8-h maximum ozone concentration for the period 2005–2010 at an aggregated area level. Shapefile of the Italian administrative regions and municipalities downloaded from the website of the Italian National Institute of Statistics ([www.istat.it/it/archivio/24613](http://www.istat.it/it/archivio/24613)).

TABLE 3 Estimated percentage change (95% CI) in the incidence rate for a 10-unit increase in average concentrations of pollutants

	Average PM <sub>10</sub> concentration (10 µg·m <sup>-3</sup> increase)	Average O <sub>3</sub> concentration (10 ppm increase)	Average NO <sub>2</sub> concentration (10 µg·m <sup>-3</sup> increase)
<b>Overall</b>			
Unadjusted model	-3.58 [-19.16-15.01]		8.41 [-0.23-17.80]*
Fully adjusted model <sup>#</sup>	9.66 [-16.38-43.80]		6.15 [-4.48-17.95]
Best model <sup>¶</sup>	17.05 [-7.26-47.72]		7.42 [-1.86-17.57]
<b>Cold season</b>			
Unadjusted model	-2.33 [-14.14-11.10]		7.93 [0.36-16.08]*
Fully adjusted model <sup>#</sup>	5.50 [-13.81-29.14]		6.80 [-2.58-17.10]
Best model <sup>¶</sup>	11.15 [-6.20-31.72]		7.55 [-0.76-16.56]*
<b>Warm season</b>			
Unadjusted model	-5.88 [-27.11-21.52]	-7.94 [-19.95-5.86]	8.22 [-1.44-18.82]*
Fully adjusted model <sup>#</sup>	16.40 [-18.50-66.27]	-2.69 [-14.81-11.17]	4.25 [-7.19-17.10]
Best model <sup>¶</sup>	25.28 [-9.49-73.42]	-4.39 [-16.24-9.13]	6.39 [-3.62-17.45]

PM<sub>10</sub>: particulate matter with an aerodynamic diameter <10 µm; O<sub>3</sub>: ozone; NO<sub>2</sub>: nitrogen dioxide. <sup>#</sup>: adjusted for area-specific proportion of males, mean age, average income, temperature and main land use; <sup>¶</sup>: adjusted for area-specific mean age and main land use for PM<sub>10</sub> and NO<sub>2</sub> and area-specific proportion of males for O<sub>3</sub>. \*: p<0.1; \*: p<0.05.

because they are sensitive to other products of the oxidation of nitrogen oxides that are more abundant during this season [40].

To our knowledge this is the first study that aimed at investigating the association between exposure to air pollution and IPF incidence. Only a few previous studies have addressed the relationship between criteria pollutants and disease progression in affected patients [13–15], showing associations between exposure to NO<sub>2</sub>, O<sub>3</sub> and particulate matter and the risk of disease worsening or mortality. We observed no significant association between O<sub>3</sub> or PM<sub>10</sub> and the incidence of IPF, suggesting that exposure to these pollutants may trigger adverse events only in susceptible subjects, who already have the disease.

Reports in the literature have presented possible mechanisms by which NO<sub>2</sub> might have a role in the development of IPF, and they include proinflammatory response in bronchial epithelial cells, alteration of the distribution of leukocyte subsets in both blood and bronchoalveolar lavage and telomere shortening [23–27, 41]. Recent findings show that the telomeric DNA regions at the ends of chromosomes are particularly sensitive to the damage caused by reactive oxygen species [42, 43], which are also generated by NO<sub>2</sub>, and telomere shortening can be accelerated by chronic inflammation and increased oxidative stress [44–46], two major mechanisms involved in mediating NO<sub>2</sub> effects on human health [46–50]. Other mechanisms resulting from exposure to air pollution include a dose-dependent upregulation of proteins associated with epithelial–mesenchymal transition [51], and a deregulation of the transforming growth factor-β signalling pathway [52], both proposed mechanisms of IPF.

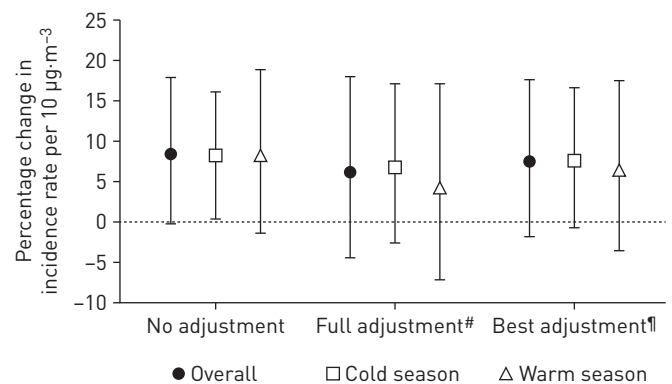


FIGURE 5 Estimated percentage change in the incidence rate (95% CI) of a 10-µg·m<sup>-3</sup> increase in average nitrogen dioxide concentration. <sup>#</sup>: adjusted for area-specific proportion of males, mean age, average income, temperature and main land use; <sup>¶</sup>: adjusted for area-specific mean age and main land use.



Our study had some limitations. Since we relied on administrative databases for the identification of incident cases of IPF, we list the major limitations of this approach.

First, IPF patients who never require hospitalisation or whose outpatient visits cannot be traced might have led to an underestimation of incident cases, even if it is unlikely that a subject with such a severe disease never accesses health services. Second, the selection algorithm is based on ICD-9-CM diagnosis codes reported in administrative records, but the accuracy of such codes is unknown and it was impossible to obtain a clinician's validation [30, 53]. However, the BCD algorithm has been used in some of the largest studies on IPF epidemiology [54–56]. Third, we used the date of the incident diagnosis of IPF as a proxy for onset, which might have occurred years earlier; therefore, relevant exposures probably precede 2005. However, considering that in Lombardy the concentrations of the pollutants of interest are declining slowly over time [57], and that we observed only slight variations when computing average exposures over different time-windows (online supplementary tables S2–S4), we are confident that the average 2005–2010 air pollutant concentrations are good indicators of patients' average chronic exposures. Furthermore, we only had data on the municipality of residence during the study period, therefore the estimated exposure could be biased for people who moved before 2005.

Finally, there might be residual confounding due to risk factors for which there were no available municipality-specific data: for example, we had no information on subjects' occupations, on how far they lived from major roads and industries or on the proportion of commuters. Most importantly, we could not assess the proportion of smokers in each municipality and smoking is a major risk factor for the development of IPF [10]. However, we are confident that the observed association is not entirely due to smoking, as this could happen only if the municipalities with higher proportions of smokers are also characterised by higher NO<sub>2</sub> concentrations; there is no evidence of such covariation in the literature.

Despite these limitations, the use of healthcare administrative databases allowed us to study a large and unselected population: Lombardy has almost 10 million inhabitants, and all of them are covered by the Italian healthcare system. In addition to being one of the most polluted areas of Europe, Lombardy is characterised by high variability of exposure, an essential condition for statistical models to succeed in identifying the target association.

Furthermore, even though traffic monitors are usually excluded from the exposure assessment process, their inclusion appears to improve the estimation of NO<sub>2</sub> concentration: our sensitivity analysis shows that concentrations estimated using such monitors are closer to those recently estimated through a model based on emissions within the project called VIAS, funded by the Italian Ministry of Health ([www.vias.it/pagine/concentrazione-degli-inquinanti](http://www.vias.it/pagine/concentrazione-degli-inquinanti)) (online supplementary figures S4 and S13).

In conclusion, we investigated for the first time the relationship between the incidence of IPF and chronic exposure to air pollution, and we found a positive association with NO<sub>2</sub> concentrations.

Longitudinal investigations focusing on subject-specific rather than aggregated data, preferably considering exposures at the subjects' addresses, could substantially improve the reliability of the estimates. Furthermore, study power could be improved by increasing the number of IPF patients involved, which could be accomplished by extending the time window for recruitment or the study population. Our study encourages further investigation of the association between air pollution and IPF, as it provides a strong signal of traffic pollution being involved in the fibrotic process, thereby identifying a potentially removable risk factor that could be controlled for reducing the incidence of the disease.

### Acknowledgements

S. Harari and G. Cesana developed the study concept. S. Conti, A. Zanobetti and J. Schwartz developed the study design. S. Conti and F. Madotto performed the statistical analysis and they interpreted the results from an epidemiological point of view, together with G. Cesana, P. Bertazzi and A. Zanobetti. S. Harari and A. Caminati studied the biological plausibility of the findings and they interpreted the results from a clinical point of view.

All authors contributed to data interpretation and to drafting the paper, revised it critically and approved the final version before the submission. They also checked that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Thomas Piraino (Dept of Respiratory Therapy, St Michael's Hospital, Toronto, ON, Canada and Dept of Anesthesia, Division of Critical Care, McMaster University, Hamilton, ON) revised the language of the final draft.

### References

- 1 Caminati A, Harari S. IPF: new insight in diagnosis and prognosis. *Respir Med* 2010; 104: Suppl. 1, S2–S10.
- 2 Bagnato G, Harari S. Cellular interactions in the pathogenesis of interstitial lung diseases. *Eur Respir Rev* 2015; 24: 102–114.
- 3 Margaritopoulos GA, Harari S, Caminati A, *et al.* Smoking-related idiopathic interstitial pneumonia: a review. *Respirology* 2016; 21: 57–64.
- 4 Raghu G, Collard HR, Egan JJ, *et al.* An official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis: evidence-based guidelines for diagnosis and management. *Am J Respir Crit Care Med* 2011; 183: 788–824.

- 5 Pinheiro GA, Antao VC, Wood JM, *et al.* Occupational risks for idiopathic pulmonary fibrosis mortality in the United States. *Int J Occup Environ Health* 2008; 14: 117–123.
- 6 Gulati M, Redlich CA. Asbestosis and environmental causes of usual interstitial pneumonia. *Curr Opin Pulm Med* 2015; 21: 193–200.
- 7 Baumgartner KB, Samet JM, Coultas DB, *et al.* Occupational and environmental risk factors for idiopathic pulmonary fibrosis: a multicenter case-control study. Collaborating Centers. *Am J Epidemiol* 2000; 152: 307–315.
- 8 Molyneaux PL, Maher TM. The role of infection in the pathogenesis of idiopathic pulmonary fibrosis. *Eur Respir Rev* 2013; 22: 376–381.
- 9 Taskar V, Coultas D. Exposures and idiopathic lung disease. *Semin Respir Crit Care Med* 2008; 29: 670–679.
- 10 Garantziotis S, Schwartz DA. Host-environment interactions in pulmonary fibrosis. *Semin Respir Crit Care Med* 2006; 27: 574–580.
- 11 Yang IV. Epigenomics of idiopathic pulmonary fibrosis. *Epigenomics* 2012; 4: 195–203.
- 12 Liu Y, Li H, Xiao T, *et al.* Epigenetics in immune-mediated pulmonary diseases. *Clin Rev Allergy Immunol* 2013; 45: 314–330.
- 13 Johannson KA, Vittinghoff E, Lee K, *et al.* Acute exacerbation of idiopathic pulmonary fibrosis associated with air pollution exposure. *Eur Respir J* 2014; 43: 1124–1131.
- 14 Sesé L, Nunes H, Cottin V, *et al.* Role of atmospheric pollution on the natural history of idiopathic pulmonary fibrosis. *Thorax* 2017; in press [<https://doi.org/10.1136/thoraxjnl-2017-209967>].
- 15 Winterbottom CJ, Shah RJ, Patterson KC, *et al.* Exposure to ambient particulate matter is associated with accelerated functional decline in idiopathic pulmonary fibrosis. *Chest* 2017; in press [<https://doi.org/10.1016/j.chest.2017.07.034>].
- 16 Turner MC, Krewski D, Pope CA 3rd, *et al.* Long-term ambient fine particulate matter air pollution and lung cancer in a large cohort of never-smokers. *Am J Respir Crit Care Med* 2011; 184: 1374–1381.
- 17 Hoek G, Krishnan RM, Beelen R, *et al.* Long-term air pollution exposure and cardio-respiratory mortality: a review. *Environ Health* 2013; 12: 43.
- 18 Faustini A, Rapp R, Forastiere F. Nitrogen dioxide and mortality: review and meta-analysis of long-term studies. *Eur Respir J* 2014; 44: 744–753.
- 19 Guarnieri M, Balmes JR. Outdoor air pollution and asthma. *Lancet* 2014; 383: 1581–1592.
- 20 Schikowski T, Mills IC, Anderson HR, *et al.* Ambient air pollution: a cause of COPD? *Eur Respir J* 2014; 43: 250–263.
- 21 Simoni M, Baldacci S, Maio S, *et al.* Adverse effects of outdoor pollution in the elderly. *J Thorac Dis* 2015; 7: 34–45.
- 22 Jones MG, Richeldi L. Air pollution and acute exacerbations of idiopathic pulmonary fibrosis: back to miasma? *Eur Respir J* 2014; 43: 956–959.
- 23 Diaz de Leon A, Cronkhite JT, Katzenstein AL, *et al.* Telomere lengths, pulmonary fibrosis and telomerase (TERT) mutations. *PLoS One* 2010; 5: e10680.
- 24 Garcia CK. Idiopathic pulmonary fibrosis: update on genetic discoveries. *Proc Am Thorac Soc* 2011; 8: 158–162.
- 25 Ly H. Genetic and environmental factors influencing human diseases with telomere dysfunction. *Int J Clin Exp Med* 2009; 2: 114–130.
- 26 Korfei M, Ruppert C, Mahavadi P, *et al.* Epithelial endoplasmic reticulum stress and apoptosis in sporadic idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2008; 178: 838–846.
- 27 Johannson KA, Balmes JR, Collard HR. Air pollution exposure: a novel environmental risk factor for interstitial lung disease? *Chest* 2015; 147: 1161–1167.
- 28 Nordio F, Kloog I, Coull BA, *et al.* Estimating spatio-temporal resolved PM<sub>10</sub> aerosol mass concentrations using MODIS satellite data and land use regression over Lombardy, Italy. *Atmos Environ* 2013; 74: 227–236.
- 29 Baccini M, Biggeri A, Grillo P, *et al.* Health impact assessment of fine particle pollution at the regional level. *Am J Epidemiol* 2011; 174: 1396–1405.
- 30 Harari S, Madotto F, Caminati A, *et al.* Epidemiology of idiopathic pulmonary fibrosis in Northern Italy. *PLoS One* 2016; 11: e0147072.
- 31 American Medical Association (AMA). International Classification of Diseases, 9th Revision, Clinical Modification: Physician ICD-9-CM, 2002. Chicago, AMA, 2001.
- 32 Snel S. Improvement of Classifications European Monitoring Stations for AirBase. A Quality Control. Bilthoven, European Topic Centre on Air and Climate Change (ETC/ACC); 2004.
- 33 Zauli Sajani S, Hänninen O, Marchesi S, *et al.* Comparison of different exposure settings in a case-crossover study on air pollution and daily mortality: counterintuitive results. *J Expo Sci Environ Epidemiol* 2011; 21: 385–394.
- 34 Masiol M, Agostinelli C, Formenton G, *et al.* Thirteen years of air pollution hourly monitoring in a large city: potential sources, trends, cycles and effects of car-free days. *Sci Total Environ* 2014; 494–495: 84–96.
- 35 Nalysnyk L, Cid-Ruzafa J, Rotella P, *et al.* Incidence and prevalence of idiopathic pulmonary fibrosis: review of the literature. *Eur Respir Rev* 2012; 21: 355–361.
- 36 European Environment Agency. EEA Technical Report No 17/2007, CLC 2006 Technical Guidelines. Copenhagen, European Environment Agency, 2007.
- 37 Allison PD, Waterman RP. Fixed-effects negative binomial regression models. *Sociol Methodol* 2002; 32: 247–265.
- 38 Carugno M, Consonni D, Randi G, *et al.* Air pollution exposure, cause-specific deaths and hospitalizations in a highly polluted Italian region. *Environ Res* 2016; 147: 415–424.
- 39 World Health Organization Regional Office for Europe (WHO). Air Quality Guidelines: Global Update 2005: Particulate Matter, Ozone, Nitrogen Dioxide, and Sulfur Dioxide. Geneva, WHO, 2006.
- 40 Ordóñez C, Richter A, Steinbacher M, *et al.* Comparison of 7 years of satellite-borne and ground-based tropospheric NO<sub>2</sub> measurements around Milan, Italy. *J Geophys Res Atmos* 2006; 111: doi:10.1029/2005JD006305.
- 41 Ayyagari VN, Januszkiewicz A, Nath J. Pro-inflammatory responses of human bronchial epithelial cells to acute nitrogen dioxide exposure. *Toxicology* 2004; 197: 149–164.
- 42 Monaghan P. Telomeres and life histories: the long and the short of it. *Ann NY Acad Sci* 2010; 1206: 130–142.
- 43 von Zglinicki T. Oxidative stress shortens telomeres. *Trends Biochem Sci* 2002; 27: 339–344.
- 44 Hou L, Gawron AJ. Telomere length and cancer risk: cause or consequence, long or short? *Am Assoc Cancer Res Educ Book* 2010; 2010: 51–55.

- 45 Schönland SO, Lopez C, Widmann T, *et al.* Premature telomeric loss in rheumatoid arthritis is genetically determined and involves both myeloid and lymphoid cell lineages. *Proc Natl Acad Sci USA* 2003; 100: 13471–13476.
- 46 Hou L, Wang S, Dou C, *et al.* Air pollution exposure and telomere length in highly exposed subjects in Beijing, China: a repeated-measure study. *Environ Int* 2012; 48: 71–77.
- 47 McCracken J, Baccarelli A, Hoxha M, *et al.* Annual ambient black carbon associated with shorter telomeres in elderly men: Veterans Affairs Normative Aging Study. *Environ Health Perspect* 2010; 118: 1564–1570.
- 48 Hoxha M, Dioni L, Bonzini M, *et al.* Association between leukocyte telomere shortening and exposure to traffic pollution: a cross-sectional study on traffic officers and indoor office workers. *Environ Health* 2009; 8: 41.
- 49 Chuang KJ, Chan CC, Su TC, *et al.* The effect of urban air pollution on inflammation, oxidative stress, coagulation, and autonomic dysfunction in young adults. *Am J Respir Crit Care Med* 2007; 176: 370–376.
- 50 Wuyts WA, Agostini C, Antoniou KM, *et al.* The pathogenesis of pulmonary fibrosis: a moving target. *Eur Respir J* 2013; 41: 1207–1218.
- 51 Li Y-J, Shimizu T, Hirata Y, *et al.* Diesel exhaust particle induce epithelial-to-mesenchymal transition by oxidative stress in human bronchial epithelial cell. *Eur Respir J* 2013; 42: Suppl. 57, P3896.
- 52 Líbalová H, Uhlířová K, Kléma J, *et al.* Global gene expression changes in human embryonic lung fibroblasts induced by organic extracts from respirable air particles. *Part Fibre Toxicol* 2012; 9: 1.
- 53 Caminati A, Madotto F, Cesana G, *et al.* Epidemiological studies in idiopathic pulmonary fibrosis: pitfalls in methodologies and data interpretation. *Eur Respir Rev* 2015; 24: 436–444.
- 54 Raghu G, Chen SY, Hou Q, *et al.* Incidence and prevalence of idiopathic pulmonary fibrosis in US adults 18–64 years old. *Eur Respir J* 2016; 48: 179–186.
- 55 Raghu G, Chen SY, Yeh WS, *et al.* Idiopathic pulmonary fibrosis in US Medicare beneficiaries aged 65 years and older: incidence, prevalence, and survival, 2001–11. *Lancet Respir Med* 2014; 2: 566–572.
- 56 Raghu G, Weycker D, Edelsberg J, *et al.* Incidence and prevalence of idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2006; 174: 810–816.
- 57 Bigi A, Ghermandi G, Harrison RM. Analysis of the air pollution climate at a background site in the Po valley. *J Environ Monit* 2012; 14: 552–563.