medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20067777; this version posted April 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Updated April 16, 2020

Spatial variability of nitrogen dioxide, formaldehyde, and BTEX and residential exposure of children in the industrial area of Viadana, Northern Italy

Authors

Alessandro Marcon[1], Silvia Panunzi[1], Massimo Stafoggia[2], Chiara Badaloni[2], Kees de Hoogh[3,4], Linda Guarda[5], Francesca Locatelli[1], Caterina Silocchi[6], Paolo Ricci[5], Pierpaolo Marchetti[1]

Affiliations

1. Unit of Epidemiology and Medical Statistics, Department of Diagnostics and Public Health, University of Verona, Verona, Italy
2. Department of Epidemiology, Lazio Region Health Service ASL Roma 1, Rome, Italy
3. Swiss Tropical and Public Health Institute, Basel, Switzerland
4. University of Basel, Basel, Switzerland
5. UOC Osservatorio Epidemiologico, Agenzia di Tutela della Salute della Val Padana, Mantova, Italy
6. UOS Salute e Ambiente, Agenzia di Tutela della Salute della Val Padana, Mantova, Italy

Corresponding author

Alessandro Marcon, PhD, Unit of Epidemiology and Medical Statistics, Department of Diagnostics and Public Health, University of Verona, c/o Istituti Biologici II, strada Le Grazie 8, 37134 Verona, Italy; phone: +39-045-8027668; e-mail: alessandro.marcon@univr.it; ORCID: 0000-0002-2778-658X

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20067777; this version posted April 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Contributorship statement

AM and PR are co-principal investigators of the Viadana III study. AM conceived the present analysis and drafted the first version of the manuscript. SP performed the geographical and statistical analysis. PM supervised the geographical and statistical analyses. CS and PR obtained data from the environmental protection agency (ARPA Lombardia). LG reconstructed the residential history of participants. CB and MS provided exposure estimates from EPISAT. KdH provided exposure estimates from ELAPSE. All the authors contributed in the discussion of the analysis plan, the interpretation of results, and commented on the first draft, they critically reviewed and approved the final version of the manuscript.

Funding

The Viadana III study was funded by the Agenzia Tutela della Salute della Val Padana, Mantova (Decree n. 278, 17/07/2018). A Research Scholarship was co-funded by the Department of Diagnostics and Public Health, University of Verona. Geocoding of participants' addresses was supported by EGON Solutions, via Enrico Fermi 13/C, 37135 Verona.

Disclosure of potential conflicts of interest

The authors declare no conflicts of interest related to this study.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20067777; this version posted April 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Abbreviations

BC, black carbon

BTEX, benzene, toluene, ethylbenzene, o-xylene, mp-xylene

ELAPSE, Effects of Low-Level Air Pollution: A study in Europe

EPISAT, Dati satellitari ed uso del territorio per la stima delle esposizioni a livello nazionale

LoQ, limit of quantification

LUR, land use regression

NO_2, nitrogen dioxide

PM, particulate matter

PM_{10}, particulate matter with an aerodynamic diameter of 10 μm or less

$PM_{2.5}$, particulate matter with an aerodynamic diameter of 2.5 μm or less

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20067777; this version posted April 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

ABSTRACT

Chipboard production is a source of ambient air pollution. We assessed the spatial variability of outdoor pollutants and residential exposure of the paediatric population living in Viadana, in proximity to the largest industry for chipboard production in Italy.

We used data on NO_2, formaldehyde, and BTEX collected by the regional environmental agency at 25 passive sampling sites during 10 monitoring weeks in 2017-18, and compared passive sampling measurements for NO_2 with data from local continuous monitoring stations. We also compared interpolated surfaces of NO_2 and formaldehyde with previous maps obtained for 2010. Finally, we compared exposures attributed using these maps to estimates from land use regression models based on routine monitoring data (ELAPSE and EPISAT studies).

In 2017-18, the median annual concentration of NO_2 in Viadana was 16.0 (Q1–Q3: 12.4–16.6) μg/m^3. When recalibrated to data from continuous monitors, this figure changed to 26.4 (17.6–27.8) μg/m^3, suggesting a marked underestimation by passive samplers. The median annual concentration of formaldehyde was 1.6 (1.5–2.0) μg/m^3. Winter concentrations of benzene and toluene were 1.6 (1.5–1.6) and 2.8 (2.6–3.1) μg/m^3, respectively. For both 2010 and 2017-18, we observed higher concentrations of NO_2 and formaldehyde in the southern part of Viadana, with hot-spots in close proximity to the industry. All the models consistently detected a gradient of higher exposure for children living in proximity to the industry.

Consistent findings from a variety of exposure metrics, obtained using different methods and input data, highlight that people living in proximity to the industry in Viadana are exposed to an environmental risk that is persisting over time.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20067777; this version posted April 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Keywords

Aldehydes; BTEX; environmental monitoring; exposure assessment; geographic information system; nitrogen oxides; passive sampling; particulate matter; wood industry.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20067777; this version posted April 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

1. INTRODUCTION

Industrial wood manufacturing and chipboard production are a source of ambient air pollutants. These include dust from mechanical woodworking, formaldehyde from the resins used to bond wood particles, and a variety of combustion by-products (Dahlgren et al. 2003; Marcon et al. 2014).

The largest industrial park for the production of chipboard in Italy is located in the health district of Viadana (labelled “district” for short). It is located in the Po Plain, in the north of the country, one of the most polluted regions in Europe (Larsen et al. 2012; Stafoggia et al. 2017). The district extends over an area of 363 km^2 and includes Viadana and other 9 smaller municipalities, counting 47,701 inhabitants overall in 2018 (http://demo.istat.it). The main emission sources of air pollution are two big industries in the south of the district, which include chemical plants for the synthesis of urea-formaldehyde resins, chipboard production and storage facilities, and small incinerators (Marcon et al. 2014). Both industries are under Directive 2010/75/UE for Integrated Pollution Prevention and Control.

We previously showed that proximity to the industries is associated with several adverse outcomes in the paediatric population living in the district, including respiratory and irritation symptoms (de Marco et al. 2010; Girardi et al. 2012; Rava et al. 2012) and hospital admissions for respiratory diseases (Marchetti et al. 2014; Rava et al. 2011). Using data from *ad hoc* monitoring conducted in 2010, it was found that proximity to the factories is linked to a higher outdoor exposure to nitrogen dioxide (NO_2) and formaldehyde, and that exposure is associated with increased biomarkers of genotoxicity in mouth mucosa cells among children (Marcon et al. 2014).

NO_2 has been related to a range of adverse effects, including mortality, exacerbations of obstructive airway diseases, and childhood asthma (Curtis et al. 2006). Formaldehyde is

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20067777; this version posted April 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

known to cause irritation and neuro-vegetative symptoms, and has mutagenic and carcinogenic properties (IARC 2006, update 2018). In urban environments, NO_2 and formaldehyde mostly derive from combustions linked with road traffic, industrial activities, and domestic heating; formaldehyde is also generated through photo-oxidation of anthropogenic hydrocarbons (Kheirbek et al. 2012). In the district, wood waste incineration and power generation at the industrial premises are additional sources of nitrogen oxides, whereas formaldehyde is emitted during chipboard production and storage (Marcon et al. 2014).

In the "Viadana III" study (http://biometria.univr.it/viadanastudy/) we are conducting new cohort studies of the resident population aged 0–20 years at baseline (2013), aimed to assess the relationship between exposure to air pollutants and further health outcomes during follow-up (2013–2017). The present analysis is restricted to the municipality of Viadana, where new air pollution monitoring campaigns were conducted in 2017-18. We aimed to understand whether living next to the industrial plant is still related to a higher air pollution exposure. We also discuss whether the exposure estimates obtained for 2010 can be reliably used in the new follow-up study.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20067777; this version posted April 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

2. METHODS

2.1. Passive sampling monitoring campaigns

We obtained data from monitoring campaigns conducted using Radiello® passive samplers (Fondazione Salvatore Maugeri, Padova, Italy) both in 2010 and in 2017–2018.

The monitoring campaigns in 2010 were planned and conducted in collaboration between the regional environmental agency and our research team for epidemiological purposes (Marcon et al. 2014). Briefly, measurements of formaldehyde and NO_2 were conducted during two weeks in the cold season and two weeks in the warm season at 62 sites in the district. Site selection overrepresented densely populated areas close to the two industries. Thirty-three of these sites were located in the eponymous municipality (labelled "Viadana" for short).

The monitoring campaigns in 2017-18 where conducted by the environmental agency. Twenty-five sites were identified over a 10 km^2 area around the centre of Viadana. NO_2 and formaldehyde, but also BTEX (benzene, toluene, ethylbenzene, o-xylene, mp-xylene), acetaldehyde, and benzaldehyde were monitored during five weeks in the cold season (from 21/11/2017 to 27/12/2017) and five weeks in the warm season (from 8/5/2018 to 12/6/2018). When measurements were below the limit of quantification (LoQ), we assigned half the LoQ value, except when this occurred during all/most of the weeks (**Table S1**).

2.2. Data from continuous air quality stations

For 2017-18, we obtained hourly concentrations of NO_2 from two air quality stations in Viadana (see maps in **Figure S1**) during the same passive sampling periods. The first station, located 500 m away from the industry (via Gianmarco Cavalli, Viadana), was a routine urban background station. The second one was a temporary station located in a meadow close to a graveyard (via Cesare Airoldi, Viadana). The two stations were

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20067777; this version posted April 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

equipped with similar detectors in agreement with current legislation (D. Lgs. 155/2010), and measured NO_2 by continuous chemo-luminescence detectors. There were no traffic monitoring stations operating in the district area.

2.3. Geocoding of residential addresses

We used data from the Viadana III study to derive a list of residential addresses in the study area. Viadana III includes prospective studies of two paediatric cohorts followed up between 2013 (baseline) and 2017. The first cohort is made of all the 3,854 children and adolescents born in 1992–2003 who were attending the schools in the district in 2006, when their parents were surveyed (de Marco et al. 2010). Of these, 3,711 (96%) subjects were traced through the healthcare provider records; 3,688 were still living in the district area in 2013. The second cohort consisted of all the 4,221 children born in the district in 2004–2012 who lived in the district in 2013. Overall, the age range of the two cohorts was 0–20 years at baseline.

We geocoded residential addresses at baseline in two stages (see **Appendix S1** for details). We first used the World Geocoding Service locator integrated into ArcGIS PRO software (ESRI, Redlands, California). Geocoding was successful if all these conditions were met: matching accuracy score was >90 (range: 0–100, https://pro.arcgis.com/en/pro-app/help/data/geocoding/what-is-included-in-the-geocoded-results-.htm); input and output municipality matched; input street number was recognised, i.e. not removed from the output. Poorly located addresses were geocoded again via outsourcing by EGON Solutions, via Enrico Fermi 13/C, Verona. See a comparison between geocoding methods in **Table S2**. Finally, we excluded 206 and 178 addresses that failed to be geocoded from the first and second cohort, respectively, which resulted in 3,482 (94%) available geocodes for the first cohort, and 4,043 (96%) for the second. After combining coordinates

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20067777; this version posted April 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

from the two cohorts and removing duplicates (i.e. children living at the same address), we obtained a list of 4,390 unique locations in the district; of these, 1,814 were in Viadana.

2.4. Population and street density indicators

With the aim of characterising non-industrial pollutant sources in the study area, we obtained the number of inhabitants, buildings, and households per census tract by national statistics census data (ISTAT, 14° Censimento generale della popolazione e delle Abitazioni (http://dawinci.istat.it/), and derived population density indicators within 500 m buffers around the passive sampling locations by area-weighting. We also obtained the road network through Open Street Map (https://www.openstreetmap.org/), and derived total length of all types of roads in 500 m buffers around samplers as an indicator of street density.

2.5. Exposure assignment

We obtained estimates of exposure to outdoor air pollutants at the 1,814 geocoded addresses by applying models available from three projects for baseline (or closest year), and new exposure models for 2017-18.

For baseline, we used the ordinary kriging models of NO_2 and formaldehyde concentrations that we had previously devised using data from the passive sampling campaigns in 2010 (Marcon et al. 2014). For the same year, we also used the land-use regression (LUR) models for $PM_{2.5}$, NO_2, and black carbon (BC) concentrations developed for Western Europe within the ELAPSE study (Effects of Low-Level Air Pollution: A study in Europe) (de Hoogh et al. 2018). ELAPSE models are based on routine air quality monitoring data for $PM_{2.5}$ and NO_2, and *ad hoc* monitoring data for BC, in combination with satellite data, dispersion model estimates, land cover and traffic indicators. Moreover, we obtained PM_{10} and $PM_{2.5}$ concentrations for 2012 and 2013, respectively, using the spatiotemporal models developed for Italy within the EPISAT project (Dati satellitari ed uso

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20067777; this version posted April 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

del territorio per la stima delle esposizioni a livello nazionale) (Badaloni et al. 2018; Stafoggia et al. 2017). EPISAT models are based on routine air quality data and incorporate both spatial (e.g. population density and emission data), and spatiotemporal predictors (e.g. satellite data and meteorology). Exposure estimates were optimised for the study area, aimed to capture PM variation due to very local sources, see stage 4 modelling in Stafoggia et al. (2017).

For 2017-18, we devised new ordinary kriging models for NO_2 and formaldehyde using passive sampling data. The best-fitting models were chosen by minimizing the root mean squared error (RMSE) by leave-one-out cross-validation (LOOCV) (Pebesma and Wesseling 1998). The spatial variogram providing the best model fit was a Gaussian class model for both pollutants. The variogram parameters were: partial sill = 13.4, range = 3.5 km, nugget = 2.5, and direction in plane = 135° (east) for NO_2; and partial sill = 0.09, range = 2.0 km, nugget = 0.07 (no anisotropy) for formaldehyde. The estimated LOOCV-RMSEs were 1.95 and 0.28, respectively.

2.6. Statistical analysis

For the ten monitoring weeks in 2017-18, we calculated daily NO_2 concentrations by averaging hourly data from each of the two continuous monitoring stations, setting to missing the days with less than 25% of hourly measurements available, and then imputing missing data by k-nearest neighbour algorithm using measurements from the other station as we did in Marchetti et al. (2017). We then obtained a “recalibration equation” for each of the two monitoring sites as follows: 1) we calculated weekly average “measured concentrations” using the imputed time-series; 2) we calculated weekly weighted-average “estimated concentrations” of NO_2 at the same location using passive sampling measurements as input data and sampler-to-station distance as a weighting factor; 3) we regressed estimated concentrations on measured concentrations. The two resulting

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20067777; this version posted April 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

equations were used to predict NO_2 concentrations at the 25 passive sampling sites as if they were obtained by continuous monitors. Finally, by averaging the two predicted concentrations, we obtained one recalibrated estimate of annual NO_2 concentration at each site.

We reported statistics on air pollution measured concentrations, as well as estimated exposures at the residential addresses, as a function of distance to the industry. Overall comparisons of exposures across groups of children were tested by the Kruskall-Wallis rank test.

The analyses were carried out using Stata Statistical Software, Release 16.1 (College Station, TX: StataCorp LLC) and R version 3.5.1 (http://www.R-project.org/).

2.7. Ethical approval

The Viadana III study has been approved by the Comitato Etico Val Padana (Prot. n. 4813, 12/02/2019).

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20067777; this version posted April 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

3. RESULTS AND DISCUSSION

3.1. Performance of passive samplers

For the present study, we used available monitoring data collected using Radiello® tubes by the environmental protection agency of Lombardia region. The working principle of passive samplers is uptake of gaseous pollutants from the air by an adsorbing material at a constant uptake rate controlled by diffusion, followed by desorption and titration (Buczynska et al. 2009). As opposed to continuous monitors, passive samplers are cheap and easy to use since they do not require forced air movement through the device and thus a pump operated using electricity (Santana et al. 2017). However, diffusion rates can vary largely according to environmental conditions thus affecting their performance (Mason et al. 2011).

For this reason, we compared passive sampling measurements of NO_2 with data from continuous monitors. We found that all the Radiello® measurements obtained during the cold season in Viadana, and most of those obtained in the warm season, were below the average concentrations derived by background and industrial stations for the corresponding weeks (**Figure 1**). **Figure S2** reports a comparison of weekly concentrations measured by continuous monitoring stations with the corresponding measurements by a co-located (or the closest) passive sampler (panels A and B), or with weighted-average concentrations estimated using data from all passive samplers (panels C and D). It is evident that passive samplers strikingly underestimated NO_2 concentrations in Viadana, especially in the upper range of concentrations.

Our conclusion form these findings is that recalibration may be necessary when a study aims at obtaining quantitative estimates of health effects related to NO_2 exposure derived by passive sampling. However, due to their high correlation with concentrations obtained

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20067777; this version posted April 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

by continuous monitoring stations, passive samplers can be reasonably employed when the purpose is to detect exposure-response relationships, even when no data are available for recalibration.

We did not have reference data to assess the performance of passive samplers for the other pollutants. Using Radiello® tubes, Evans and Stuart (2011) observed quite similar winter concentrations of formaldehyde between active monitors and co-located passive samplers (2.1 vs 2.2 μg/m^3) in Hillsborough County, Florida. Mason et al. (2011) documented an excellent performance in controlled experiments (2% difference from a reference method) but also substantial (40%) under- and over-estimation in real-world settings.

3.2. Air pollution concentrations in Viadana

In 2017-18, the mean concentrations of NO_2 during the cold and warm seasons were 19.8 and 9.9 μg/m^3, respectively (**Table 1**). Spatial variability was higher for NO_2 during the warm season (interquartile range/median ratio = 0.64), compared to measurements during the cold season as well as to the other pollutants (**Table 2**). The median annual concentration was 16.0 (Q1–Q3: 12.4–16.6) μg/m^3. When recalibrated using data from continuous stations, it was 26.4 (Q1–Q3: 17.6–27.8) μg/m^3. All the recalibrated annual concentrations were within the EU air quality standard of 40 μg/m^3 on average (**Figure 2**). For comparison, average NO_2 concentrations measured by routine stations in 2017-18 were around 30 μg/m^3 for the urban area of Mantova, and 20 μg/m^3 for the larger province also encompassing suburban/rural areas (ARPA Lombardia Rapporto qualità aria 2018). Using a different type of passive sampler (Ogawa badges), Cyrys et al. (2012) found median annual concentrations between 26.6 and 54.3 μg/m^3 in three towns located in the Po Plain in 2010.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20067777; this version posted April 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

As opposed to the campaigns in 2010 (Marcon et al. 2014), we found higher formaldehyde concentrations in the cold season than in the warm season (2.2 vs 1.1 μg/m^3, respectively, **Table 1**), which was surprising because lower concentrations are expected in winter (Kheirbek et al. 2012; Villanueva et al. 2014). During the warm season, temperatures ranged between 12.8 and 25.1°C, which may not completely represent summertime in Italy, where temperature can reach 40°C. The median annual concentration of formaldehyde was 1.6 μg/m^3 (Q1–Q3: 1.5–2.0 μg/m^3), which is similar to other studies that used Radiello® tubes (Evans and Stuart 2011; Kheirbek et al. 2012). There are no ambient air quality standards for aldehydes, and studies on the health effects of outdoor formaldehyde are still few (Dahlgren et al. 2003; Marcon et al. 2014; Morello-Frosch et al. 2000). For long-term exposure, Effects Screening Levels (calculated so that lower ambient levels are unlikely to be of concern for health and the environment) are set to 3.3 μg/m^3 from the Texas Air Monitoring Information System (USA – Texas Commission on Environmental Quality, https://www.tceq.texas.gov/ accessed 11 March 2020) (Santana et al. 2017). However, formaldehyde concentrations in this study are above 0.8 μg/m^3, which is the benchmark set by the US Environmental Protection Agency for a 1 in 100,000 lifetime cancer risk for inhaled exposures (https://cfpub.epa.gov/ accessed 12 March 2020).

In the cold season, mean concentrations of benzene, toluene, and acetaldehyde were 1.5, 2.8, and 0.7 μg/m^3 respectively (**Table S3**), whereas they were mostly below the LoQ during the warm season (**Table S1**). Benzene and toluene concentrations were comparable or lower to concentrations observed at background or semirural sites in other studies using Radiello® tubes (Buczynska et al. 2009; Gaeta et al. 2016; Kerchich and Kerbachi 2012). According to EU air quality standards, annual average concentrations of benzene should be below 5 μg/m^3 (European directive 2000/69/EC for 2010), although

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20067777; this version posted April 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

there is no evidence of a safe exposure threshold (Amini et al. 2017). Ethylbenzene, o-xylene, mp-xylene, and benzaldehyde concentrations were below the LoQ during both seasons (**Table S1**), although it should be noted that LoQs in our study were higher compared to other studies (Buczynska et al. 2009).

3.3. Air pollution concentrations and proximity to the industry

For both NO_2 (**Figure 3**) and formaldehyde (**Figure 4**), we observed higher concentrations in the southern part of Viadana, with concentrations above the 90th percentile in close proximity to the industry. Maps for 2017-18 were broadly consistent with the distributions obtained for 2010, despite methodological differences in the overall duration of monitoring, and in the number and location of samplers. In particular, no samplers were located in the area at the East of the industry in 2017-18, which might have affected precision of estimated concentrations in that zone.

For NO_2, a clear relationship with distance was also observed when looking at observed concentrations (**Table 2**). The median recalibrated concentration was 11 μg/m^3 higher within 1 km from the industry than in the area beyond 1.8 km (**Figure 2**). As regards formaldehyde, observed concentrations did not vary appreciably according to distance (**Table 2**), and spatial contrasts were small, as we observed during the previous campaigns (Marcon et al. 2014). A relatively large nugget effect (0.07) compared to the partial sill (0.09) in our model for formaldehyde, indicating either measurement error or variability that kriging was not able to account for, could be partly attributed to the low spatial variability for this pollutant. Using Radiello® tubes, Kheirbek et al. (2012) reported a coefficient of variation of 22% for formaldehyde measurements in New York City at springtime, whereas even lower coefficients of 12-13% were reported in other studies (Evans and Stuart 2011; Marcon et al. 2014).

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20067777; this version posted April 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

BTEX are mainly emitted through vehicular exhaust fumes and fuel evaporative losses, so their concentrations are expected to reflect traffic levels (Buczynska et al. 2009; Kerchich and Kerbachi 2012; Kheirbek et al. 2012). In Viadana, areas closer to the industry were characterised by a higher density of roads, population and buildings (**Table 3**). For sampling sites at increasing distance to the factory, the median total road length in 500-m buffers decreased from 8.3 km to 4.6 km (sites at <1.2 and >1.8 km, respectively). Nonetheless, BTEX concentrations did not vary as a function of distance to the industry. This may be due to the small number of samplers, or to the fact that measurements were too close to the LoQ to appreciate spatial contrasts.

3.4. Comparison between air pollution exposure models

Estimates of residential exposure attributed to children's addresses are reported in **Table 4**. Despite the fact that they were quite different in methodology (kriging vs LUR), input data (routine vs passive sampling data), extension (municipality vs national/European), and period (2010-13 vs 2017-18), all the exposure models consistently detected a gradient of concentrations with distance to the industry, whatever the pollutant considered. Interestingly, the three models for NO_2 detected a maximum concentration at a distance of 1–1.7 km to the industry, followed by a decline at a larger distance, which could either reflect the pattern of dispersion of nitrogen oxides from the 70-m high chimney, or the local distribution of streets, traffic type, and built spaces. Differences in median exposures between children living <1 km and children living ≥3.5 km to the industry were approximately 2 and 5 μg/m^3 for $PM_{2.5}$ and PM_{10}, respectively.

We assigned estimates of exposure to NO_2 and formaldehyde using ordinary kriging. We were not able to develop LUR models because reliable data to derive road traffic/density indicators were unavailable. A drawback of the present study is that we had no direct measurements of PM in the study area to develop exposure models. PM are one of the

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20067777; this version posted April 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

most important emissions from the industry in Viadana, and they convey most of formaldehyde mass that is released during production. They are the most critical component of air pollution in the Po Plain due to their regional distribution and long-term atmospheric persistence (Larsen et al. 2012). For these reasons, we obtained estimates of exposure to PM_{10} and $PM_{2.5}$ from LUR models using routine input data, as well as exposure to BC, which is deemed to be one of the most toxic PM components and are responsible for 40-70% of the roadside increment in $PM_{2.5}$ mass (Janssen et al. 2011).

LUR models based on routine monitors from country-wide or larger areas perform well in the Po Plain (de Hoogh et al. 2018; Stafoggia et al. 2017) but they may not be ideal to capture variability in industrial pollution. This can in part be attributed to the relatively small number of industrial monitoring sites that can be included at the modelling stage (Stafoggia et al. 2017), but also to the fact that spatial LUR predictors are unable to accurately reflect specific types of industrial productions and facilities.

4. CONCLUSIONS

Using data from passive monitoring campaigns, we found that residential exposure to NO_2 and formaldehyde was higher in proximity to the chipboard industry in Viadana during 2017-18. The spatial distribution of the pollutants was consistent with what we previously observed in 2010 (Marcon et al. 2014), suggesting that the increased environmental risk for residents in proximity to the industry is persisting over time. Exposure estimates derived from available large-scale LUR models of NO_2, $PM_{2.5}$ and PM_{10} from routine stations showed comparable associations with distance. This suggests that people living close to the industry are exposed to a mixture of air pollutants, and it also supports the reliability of exposure estimates derived for the Viadana III study. Finally, we conclude that

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20067777; this version posted April 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

attention should be paid in epidemiological studies aimed at quantifying air pollution effects when a reference measurement method is not available.

TABLES

Table 1. Distribution of weekly concentrations of air pollutants measured by 25 passive samplers in Viadana; daily temperature and rain precipitations from the nearest meteorological station for the same weeks.

Week n.	Starting date (DD/MM/YY)	NO_2, µg/m^3 (mean, SD)	Formaldehyde, µg/m^3 (mean, SD)	Temperature, °C (mean, min–max)	Rain, mm (mean, min–max)
1	21/11/17	19.5 (4.3)	1.6 (1.0)	6.8 (2.9–9.5)	2.6 (0.0–11.4)
2	28/11/17	24.1 (3.1)	3.2 (0.6)	2.8 (0.9–5.0)	0.0 (0.0–0.0)
3	05/12/18	22.3 (5.5)	2.0 (1.5)	1.7 (-0.8–4.3)	1.7 (0.0–9.4)
4	12/12/18	12.2 (5.6)	2.6 (1.2)	2.2 (-0.9–3.4)	0.1 (0.0–0.4)
5	18/12/18	20.8 (4.5)	1.8 (0.9)	1.3 (-0.6–5.2)	0.0 (0.0–0.0)
Overall (cold season)		19.8 (3.5)	2.2 (0.5)	2.9 (-0.9–9.5)	0.8 (0.0–11.4)
6	08/05/18	13.2 (5.9)	0.6 (0.2)	18.7 (16.1–20.4)	2.2 (0.0–7.4)
7	15/05/18	8.5 (2.7)	0.7 (0.4)	17.5 (12.8–19.7)	0.7 (0.0–5.2)
8	22/05/18	10.1 (4.3)	0.6 (0.4)	21.8 (17.3–23.9)	7.8 (0.0–50.6)
9	29/05/18	9.5 (3.9)	0.8 (0.5)	23.2 (22.4–23.8)	0.4 (0.0–2.0)
10	05/06/18	8.3 (2.9)	2.6 (0.8)	23.6 (20.5–25.1)	5.0 (0.0–32.2)
Overall (warm season)		9.9 (3.1)	1.1 (0.3)	21.0 (12.8–25.1)	3.2 (0.0–50.6)

Table 2. Distribution of measured concentrations of air pollutants in 2017-18 at the 25 passive sampling locations, by groups based on tertiles of distance to the industry in Viadana. Median with interquartile range reported.

	Distance to the industrial plant				
	<1.2 km (n=8)	1.2–1.7 km (n=9)	1.8–15.5 km (n=8)	Overall (n=25)	Interquartile range/median (n=25)
NO_2 (cold season), μg/m^3	21.5 (19.8–23.6)	19.4 (18.8–21.2)	16.4 (14.7–20.4)	19.9 (17.5–22.0)	0.22
NO_2 (warm season), μg/m^3	11.6 (9.5–13.4)	9.0 (6.6–13.0)	8.8 (6.5–11.0)	9.9 (6.7–13.0)	0.64
NO_2 (annual), μg/m^3	16.4 (15.8–16.6)	14.4 (13.0–17.1)	11.9 (10.9–16.1)	16.0 (12.4–16.6)	0.26
Formaldehyde (cold season), μg/m^3	2.3 (2.1–2.4)	2.0 (1.71–2.7)	2.2 (2.0–2.3)	2.2 (2.0–2.6)	0.28
Formaldehyde (warm season), μg/m^3	1.1 (1.1–1.4)	1.1 (0.9–1.2)	1.0 (0.9–1.0)	1.0 (0.9–1.2)	0.20
Formaldehyde (annual), μg/m^3	1.7 (1.5–2.0)	1.5 (1.4–2.0)	1.6 (1.4–1.7)	1.6 (1.5–2.0)	0.34
Acetaldehyde (cold season), μg/m^3	0.7 (0.6–0.7)	0.7 (0.6–0.9)	0.7 (0.6–0.8)	0.7 (0.6–0.8)	0.23
Benzene (cold season), μg/m^3	1.6 (1.5–1.6)	1.6 (1.3–1.6)	1.6 (1.5–1.6)	1.6 (1.5–1.6)	0.12
Toluene (cold season), μg/m^3	2.9 (2.8–3.0)	2.9 (2.5–3.1)	2.8 (2.6–3.2)	2.8 (2.6–3.1)	0.17

Table 3. Distribution of population and street density indicators in 500 m buffers around the 25 passive sampling sites, by groups based on tertiles of distance to the industry in Viadana. Median with interquartile range reported.

	Distance to the industrial plant			
	<1.2 km (n=8)	1.2–1.7 km (n=9)	1.8–15.5 km (n=9)	Overall (n=25)
Number of inhabitants	1506 (1048–2701)	713 (536–2702)	527 (337–939)	993 (536–2383)
Number of buildings	332 (236–508)	255 (151–501)	198 (107–255)	254 (185–426)
Number of households	631 (419–1189)	294 (229–1202)	224 (126–381)	393 (229–1053)
Total road length (km)	8.3 (6.9–10.7)	5.6 (3.2–9.9)	4.6 (3.1–6.9)	6.9 (3.8–9.7)

Table 4. Distribution of outdoor concentrations of air pollutants estimated at children's residential addresses, by groups based on quartiles of distance to the industry in Viadana. Median with interquartile range reported. [a]

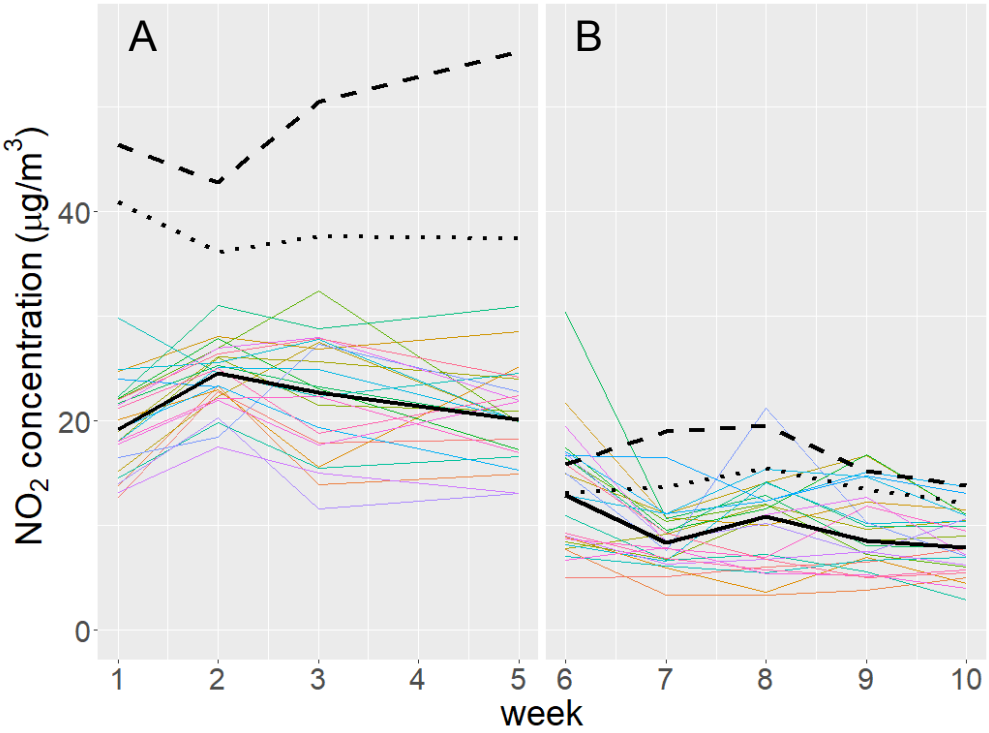
	Distance to the industrial plant				
	<1 km (n=454)	1–1.7 km (n=455)	1.8–3.4 km (n=452)	3.5–15.7 km (n=453)	Overall (n=1,814)
Viadana II study (2010)					
NO_2, μg/m^3	17.9 (16.8–18.4)	19.3 (17.8–20.0)	18.1 (17.4–18.6)	16.9 (14.9–17.4)	17.8 (16.8–18.7)
Formaldehyde, μg/m^3	2.7 (2.6–2.9)	2.7 (2.6–2.7)	2.5 (2.5–2.6)	2.3 (2.2–2.5)	2.6 (2.5–2.7)
ELAPSE study (2010)					
NO_2, μg/m^3	28.7 (27.4–29.9)	29.0 (27.9–29.8)	28.4 (27.0–29.5)	24.6 (23.1–26.7)	28.1 (26.2–29.5)
$PM_{2.5}$, μg/m^3	26.9 (26.4–27.0)	26.8 (26.6–26.9)	26.6 (26.2–26.9)	24.5 (22.9–26.4)	26.6 (26.0–26.9)
BC, 10^{-5} m^{-1}	2.0 (2.0–2.0)	2.0 (2.0–2.0)	2.0 (1.9–2.0)	1.9 (1.8–1.9)	2.0 (1.9–2.0)
EPISAT study (2012-13)					
$PM_{2.5}$ (2013), μg/m^3	25.8 (25.0–26.6)	25.5 (25.2–26.0)	25.9 (25.1–26.8)	23.6 (22.5–24.8)	25.4 (24.6–26.3)
PM_{10} (2012), μg/m^3	38.3 (36.7–40.5)	37.2 (36.1–39.0)	36.6 (34.7–38.8)	33.3 (31.8–34.6)	36.5 (34.0–38.6)
Viadana III study (2017-18)					
NO_2, μg/m^3	16.5 (15.4–16.8)	17.1 (16.7–17.2)	16.0 (13.9–16.8)	11.2 (10.8–12.1)	16.2 (13.6–16.9)
Formaldehyde, μg/m^3	1.7 (1.6–1.8)	1.8 (1.7–1.8)	1.7 (1.6–1.8)	1.6 (1.5 1.6)	1.7 (1.6–1.8)

[a] p-values for overall comparisons of concentrations across groups were all <0.001 (Kruskall-Wallis rank test)

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20067777; this version posted April 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

FIGURES

Figure 1. Weekly concentrations of NO_2 measured by continuous monitors and passive samplers during the cold season (panel A) and during the warm season (panel B). Municipality of Viadana, 2017-18. [a]



[a] the dashed and dotted lines represent weekly average concentrations by the routine and temporary stations, respectively; the solid lines represent measurements by passive samplers, with the median concentration highlighted as a solid thick line

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20067777; this version posted April 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Figure 2. Box & whiskers plot of recalibrated NO_2 concentrations at passive sampling locations, by groups based on tertiles of distance to the industry in Viadana.

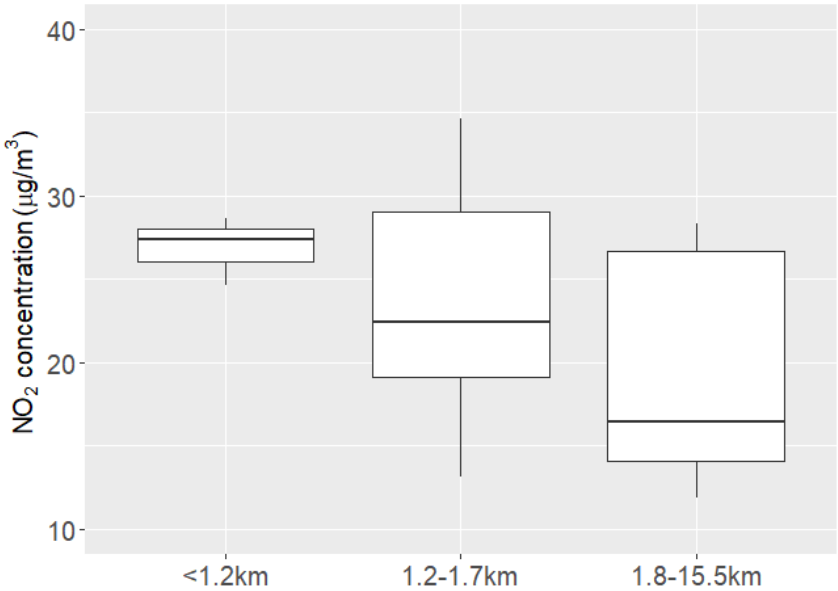
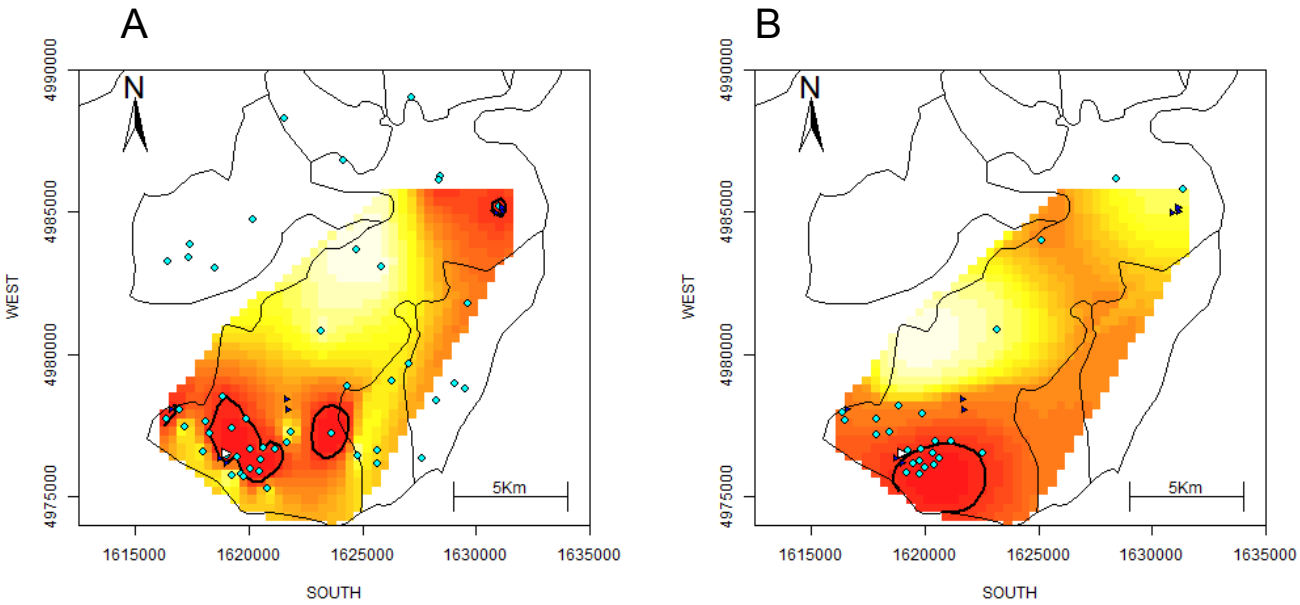


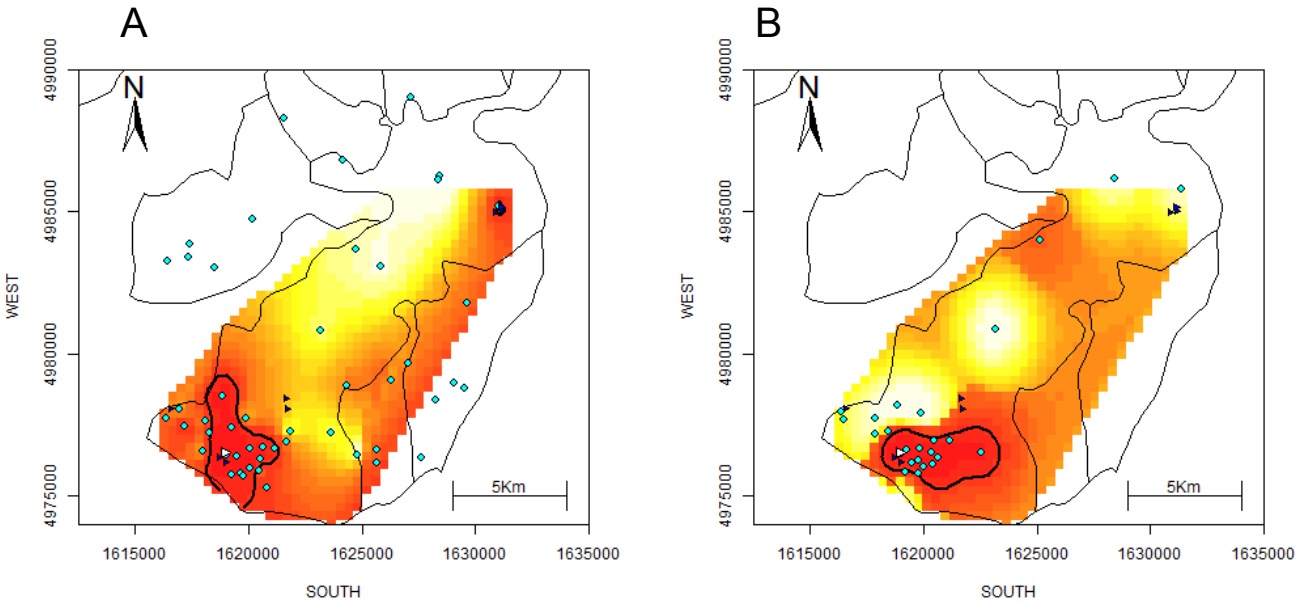
Figure 3. Spatial distribution of annual NO_2 concentrations in 2010 (panel A) and 2017-18 (panel B). [a]



[a] Black contour lines represent the 90th percentile. The white triangle represents the industry in Viadana. Blue triangles represent smaller wood factories. Cyan dots represent passive sampling sites.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20067777; this version posted April 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Figure 4. Spatial distribution of annual formaldehyde concentrations in 2010 (panel A) and 2017-18 (panel B). [a]



[a] Black contour lines represent the 90th percentile. The white triangle represents the industry in Viadana. Blue triangles represent smaller wood factories. Cyan dots represent passive sampling sites.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20067777; this version posted April 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

REFERENCES

Amini H, Yunesian M, Hosseini V, Schindler C, Henderson SB, Künzli N. 2017. A systematic review of land use regression models for volatile organic compounds. Atmos Environ 171:1-16.

Badaloni C, Cattani G, De' Donato F, Gaeta A, Leone G, Michelozzi P, et al. 2018. Big data in environmental epidemiology. Satellite and land use data for the estimation of environmental exposures at national level. Epidemiol Prev 42:46-59.

Buczynska AJ, Krata A, Stranger M, Godoi AFL, Kontozova-Deutsch V, Bencs L, et al. 2009. Atmospheric BTEX-concentrations in an area with intensive street traffic. Atmos Environ 43:311-318.

Curtis L, Rea W, Smith-Willis P, Fenyves E, Pan Y. 2006. Adverse health effects of outdoor air pollutants. Environ Int 32:815-830.

Cyrys J, Eeftens M, Heinrich J, Ampe C, Armengaud A, Beelen R, et al. 2012. Variation of NO2 and NOx concentrations between and within 36 European study areas: Results from the ESCAPE study. Atmos Environ 62:374-390.

Dahlgren J, Warshaw R, Thornton J, Anderson-Mahoney CP, Takhar H. 2003. Health effects on nearby residents of a wood treatment plant. Environ Res 92:92-98.

de Hoogh K, Chen J, Gulliver J, Hoffmann B, Hertel O, Ketzel M, et al. 2018. Spatial PM2.5, NO2, O3 and BC models for western europe - evaluation of spatiotemporal stability. Environ Int 120:81-92.

de Marco R, Marcon A, Rava M, Cazzoletti L, Pironi V, Silocchi C, et al. 2010. Proximity to chipboard industries increases the risk of respiratory and irritation symptoms in children: The Viadana study. Sci Total Environ 408:511-517.

Evans AM, Stuart AL. 2011. A passive sampling study of small-scale variations in ambient acetaldehyde and formaldehyde concentrations. Air, Soil and Water Research 4:ASWR. S7582.

Gaeta A, Cattani G, Di Menno di Bucchianico A, De Santis A, Cesaroni G, Badaloni C, et al. 2016. Development of nitrogen dioxide and volatile organic compounds land use regression models to estimate air pollution exposure near an Italian airport. Atmos Environ 131:254-262.

Girardi P, Marcon A, Rava M, Pironi V, Ricci P, de Marco R. 2012. Spatial analysis of binary health indicators with local smoothing techniques: The Viadana study. Sci Total Environ 414:380-386.

IARC. 2006, update 2018. Formaldehyde, 2-butoxyethanol and 1-tert-butoxypropan-2-ol. IARC Monogr Eval Carcinog Risks Hum 88:1-478.

Janssen NA, Hoek G, Simic-Lawson M, Fischer P, van Bree L, ten Brink H, et al. 2011. Black carbon as an additional indicator of the adverse health effects of airborne particles compared with PM10 and PM2.5. Environ Health Perspect 119:1691-1699.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20067777; this version posted April 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Kerchich Y, Kerbachi R. 2012. Measurement of BTEX (benzene, toluene, ethybenzene, and xylene) levels at urban and semirural areas of Algiers city using passive air samplers. Journal of the Air & Waste Management Association 62:1370-1379.

Kheirbek I, Johnson S, Ross Z, Pezeshki G, Ito K, Eisl H, et al. 2012. Spatial variability in levels of benzene, formaldehyde, and total benzene, toluene, ethylbenzene and xylenes in New York City: A land-use regression study. Environmental Health 11:51.

Larsen BR, Gilardoni S, Stenström K, Niedzialek J, Jimenez J, Belis CA. 2012. Sources for PM air pollution in the Po plain, Italy: Ii. Probabilistic uncertainty characterization and sensitivity analysis of secondary and primary sources. Atmos Environ 50:203-213.

Marchetti P, Marcon A, Pesce G, Girardi P, Guarda L, Pironi V, et al. 2014. Children living near chipboard and wood industries are at an increased risk of hospitalization for respiratory diseases: A prospective study. Int J Hyg Environ Health 217:95-101.

Marchetti P, Pesce G, Villani S, Antonicelli L, Ariano R, Attena F, et al. 2017. Pollen concentrations and prevalence of asthma and allergic rhinitis in italy: Evidence from the GEIRD study. Sci Total Environ 584:1093-1099.

Marcon A, Fracasso ME, Marchetti P, Doria D, Girardi P, Guarda L, et al. 2014. Outdoor formaldehyde and NO2 exposures and markers of genotoxicity in children living near chipboard industries. Environ Health Perspect 122:639-645.

Mason JB, Fujita EM, Campbell DE, Zielinska B. 2011. Evaluation of passive samplers for assessment of community exposure to toxic air contaminants and related pollutants. Environmental Science & Technology 45:2243-2249.

Morello-Frosch RA, Woodruff TJ, Axelrad DA, Caldwell JC. 2000. Air toxics and health risks in California: The public health implications of outdoor concentrations. Risk Anal 20:273-291.

Pebesma EJ, Wesseling CG. 1998. Gstat: A program for geostatistical modelling, prediction and simulation. Computers & Geosciences 24:17-31.

Rava M, Marcon A, Girardi P, Pironi V, Silocchi C, Ricci P, et al. 2011. Proximity to wood factories and hospitalizations for respiratory diseases in children. Sci Total Environ 410:80-86.

Rava M, Crainicianu C, Marcon A, Cazzoletti L, Pironi V, Silocchi C, et al. 2012. Proximity to wood industries and respiratory symptoms in children: A sensitivity analysis. Environ Int 38:37-44.

Santana FO, Campos VP, Cruz LP, Luz SR. 2017. Formaldehyde and acetaldehyde in the atmosphere of Salvador-Ba, Brazil, using passive sampling. Microchemical Journal 134:78-86.

Stafoggia M, Schwartz J, Badaloni C, Bellander T, Alessandrini E, Cattani G, et al. 2017. Estimation of daily PM10 concentrations in Italy (2006-2012) using finely resolved satellite data, land use variables and meteorology. Environ Int 99:234-244.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.16.20067777; this version posted April 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Villanueva F, Tapia A, Notario A, Albaladejo J, Martínez E. 2014. Ambient levels and temporal trends of VOCs, including carbonyl compounds, and ozone at Cabañeros national park border, Spain. Atmos Environ 85:256-265.