Atherosclerosis and Air pollution: understanding traffic-related exposure and its effects

Marcela Rivera

DOCTORAL THESIS – POMPEU FABRA UNIVERSITY 2012

DIRECTOR

Dr. Nino Künzli

Department of Epidemiology and Public Health, Swiss Tropical and Public Health Institute (SwissTPH) Basel, and Chair Social and Preventive Medicine, University Basel, Basel Switzerland

CODIRECTOR

Dr. Xavier Basagaña

Center for Research in Environmental Epidemiology (CREAL), Universitat Pompeu Fabra (UPF), Barcelona, Spain.

TUTOR

Dr. Jordi Sunyer

Center for Research in Environmental Epidemiology (CREAL), Universitat Pompeu Fabra (UPF), Barcelona, Spain.

DEPARTMENT OF EXPERIMENTAL AND HEALTH SCIENCES

UNIVERSITAT POMPEU FABRA

A mis padres,
Quienes han estado siempre tan cerca en la distancia.

ACKNOWLEDGEMENTS

I would like to thank all the people who contributed to my thesis project both with scientific and personal support.

Thanks to my thesis director Nino Künzli, for giving me the opportunity to redirect my career towards Epidemiology and Public Health. His guidance during the project development and his confidence has been determinant for the success of this thesis. It was a privilege to work with him; I learned a lot and enjoyed it very much too.

I also had the luck of having a senior statistician and young researcher as a co-director, Xavier Basagaña. He taught me a good deal of the very hard to grasp concepts in statistics. Thanks for your advice Xavi!

I thank my tutor, Jordi Sunyer who has guided me in all the administrative issues of research and who has supported me in taking some of the core courses in my training.

Mercè Medina, who was key in order to start working in a very productive research group; she also helped me extensively with the papers' first drafts.

Many thanks to my paper coauthors and my research group: Maria Foraster, Inma Aguilera, Laura Bouso and David Agis. Of course, this project would not have been done without their hard work.

Thanks to the REGICOR team, especially to Jaume Marrugat, Roberto Elosua, Pascual Solanas, Rafael Ramos, Martina Sidera, Núria Pagès and Dave MacFarlane.

Gerard Hoek, for his support during my stay in the IRAS institute, Utrecht, and for making that stay not only very productive, but also very nice. I would come back any time!

A special thanks to Agustin Lobo, who really looks for the interests of his students and helped me finding an opportunity in a field merging the two topics I wanted to work with, but on which I had the narrow experience of a Chemical Engineer: Air pollution and Public health.

Tom Burke and Alvaro Muñoz (Profesor, me quito el sombrero después de cada seminario suyo) your classes and the talks we had came in moments when I really needed an extra motivation to go on with a process that is not exciting at every stage.

Very special thanks to my professors from the Chemical Engineering faculty at the Universidad Nacional de Colombia. I ended up in a complete

different world but I would not have devoted my work to research if it weren't for the inspiring work and classes of Carlos Ariel, Jose Hernán Jaramillo, and Miguel Ángel Gómez. I want to thank my dearest friends and classmates from whom I learned just as much: Lida Marcela, Yeny, Juan Felipe, Julian, Leo (Leandro), Natalia, Wilmar…

I would also like to acknowledge the people I have run into and who, in one way or another, have helped me in this path: Gerry Poster; Laurel Blair; Charlotte and Rudy Braun-Fahrländer for making the stays in Basel much more enjoyable.

All the people at CREAL, especially Talita, Lidia (I would repeat a congress in Dublin whenever you want), David (así los abarco a todos, no?), Eileen, Irene, Solymar, Laurel Kincl, Kyriaki, Anna Schembari, Iris, Nadine, Dania, Joan, Alicia, Raquel, Anne Ellie, Anna Espinosa. Maria for scientific support, and for sharing with me much of the pre-doctoral experiences, hard work, and also fun -and Basel and Rotterdam.

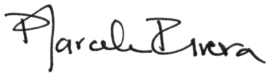
Mark Nieuwenhuijsen, Audrey de Nazelle, Marta Cirac, Michelle Mendez, Josep Maria Antó and Manolis Kogevinas for scientific advice; Anna Sillero, Natàlia Ras and Nora Bauer for administrative support.

Thanks to the REGICOR participants. There's no epidemiology without them. This project tries to give back to society a little of what society has given to us as scientists. I will always remember the warmth of the Gerundenses who welcomed us sometimes at their own homes. Thanks to the beautiful Girona. I hope Air Quality continues being a priority there.

Thanks to Barcelona and its people. After a long time abroad, I have felt like I am at home again.

Mi mayor gratitud es para mi Mamá y mi Papá quienes me han apoyado siempre y han creído en mi incluso antes que yo misma. Mis hermanos quienes los han cuidado mientras estoy fuera y me dan tantos momentos de alegría. Vanessa, quien me animo a perseguir este nuevo sueño. Antes de llegar al final, quiero darle las gracias a Bilbo, te voy a compensar por los paseos tan cortos que te di últimamente ;)

Y por supuesto a David, quien ha hecho por momentos de estadístico, editor, experto y coach en mi proyecto de tesis, y en lo que sigue de aquí en adelante, espero que por mucho tiempo –eso si, sin deadlines-. Gracias, totales.



Barcelona. December, 2011

PROLOGUE

This thesis project focused on the assessment of long-term exposure to traffic-related air pollution and its association with atherosclerosis. This is the first study that evaluated such association among the Mediterranean population.

This doctoral thesis contributes to the characterization of the spatial distribution of nitrogen dioxide, heavy metals and ultrafine particles in Girona, to the methodology in air pollution exposure assessment studying different markers of traffic-related air pollution, measurement instruments and protocols, as well as to the development of the LUR technique, identifying potential bias in the resulting coefficients of the health model and practical solutions to avoid them. Finally, this thesis added to the evidence that long-term exposure to traffic-related air pollution is associated with atherosclerosis and identified potential susceptibility factors for the Mediterranean population.

This doctoral thesis is based on the following original publications:

I. Monitoring of heavy metal concentrations in home outdoor air using moss bags
II. Spatial distribution of ultrafine particles in urban settings: a land use regression model
III. Effect of the number of measurement sites on Land Use Regression Models in estimating local air pollution
IV. Long-term exposure to traffic-related air pollution and subclinical atherosclerosis

Based on original research and guided by her thesis directors, the doctoral candidate's contribution included the exposure assessment design; participation and coordination of fieldwork and data collection; data management; training and supervising five technicians on data collection, data entry and quality control for the traffic and ultrafine particle monitoring; linkage of traffic intensity data to the digital road network; statistical analysis and reports writing; communications of results in internal meetings and international conferences; and writing the four scientific articles that constitute this doctoral thesis.

ABSTRACT

Epidemiological evidence on the effects of long-term exposure to air pollution on the chronic processes of atherogenesis is limited. Given the widespread exposure to traffic-related pollutants and the large and rising global burden of cardiovascular diseases, elucidating and quantifying the association of exposure to air pollution with atherosclerosis is of great relevance for public health. The aims of this thesis were to characterize people's exposure to long-term traffic-related pollution and evaluate its association with subclinical atherosclerosis in the adult population of the Mediterranean Spanish region.

We used data of the re-examination (2007-2010) of 2780 participants, 32 to 86 years of age, from the REGICOR study. REGICOR is a population-based prospective cohort from Girona, Spain. The main health outcome was the carotid artery intima media thickness (IMT), a validated marker of subclinical atherosclerosis, added to the study protocol at the 2007-2010 re-examination. To characterize local-scale pollution we conducted measurements of heavy metals for eight weeks -using mosses as biomonitors-, nitrogen dioxide (NO_2) for four weeks -using passive samplers- and simultaneous measures of traffic counts and ultrafine particles (UFP) -using mobile condensation particle counters- for 15-minute periods. To assess the spatial variability of these markers of traffic-related pollution, we developed land use regression (LUR) models. We also conducted a methodological study on the effects of the number of measurements on the performance of LUR models. From the characterized markers, NO_2 and traffic intensity were used to assign exposure to participants. Long-term exposure was calculated as the time-weighed average of NO_2 estimates across all residences of each participant in the last 10 years. The cross-sectional association between air pollution and IMT was investigated using multivariate linear regression analyses.

(i) The best predictor of heavy metals was the number of bus lines in the nearest street. Metals were not highly correlated with NO_2 and showed higher spatial variability than NO_2. LUR models explained between 40 to 85% of metals and 72% of NO_2 variability. Given the higher toxicity, stronger association with local traffic and higher spatial variability of heavy metals compared to NO_2, monitoring with mosses is an appealing alternative for long-term exposure assessment. **(ii)** The best predictors of UFP were traffic intensity,

distance to nearest major crossroad, area of high density residential land and household density. The LUR models of UFP explained between 36 to 51% of UFP total variation. For a subset of sites, using the mean of two repeated measurements improved the LUR model R^2 to 72%. The distribution of UFP can be modeled with fair performance based on short-term mobile monitoring. **(iii)** Models based on a small number of sampling sites, e.g. 20-30, result in artificially high adjusted R^2 and leave-one-out cross-validation R^2. The differences between these estimators and the "true" validation R^2 could be as high as 60% for small samples sizes ($\leq$20) and of around 10% for sample sizes above 100. The selection of predictor variables for the LUR model introduces classical measurement error resulting in bias of the health effect estimates. The magnitude of the error depends on the number of sampling sites and variables offered to the model. LUR models for complex urban settings should be based on a large number of measurement sites (>80) and the number of potential predictor variables should be restricted. **(iv)** Average residential traffic (load and intensity) was associated with 2% thicker IMT. Associations of residential NO_2 with IMT were weak and reached statistical significance in the subgroups of people with high education level and men above 60 years (a difference of 25$\mu g/m^3$ in NO_2 was associated with a 4.6 and 4.3% increase in IMT respectively). Long-term exposure to traffic-related air pollution is associated with carotid subclinical atherosclerosis in a random healthy population.

This thesis contributes to the characterization of the spatial distribution of NO_2, heavy metals and UFP in Girona, to the methodology in air pollution exposure assessment studying different markers of traffic-related air pollution, measurement instruments and protocols, as well as to the development of the LUR technique, identifying potential bias in the resulting coefficients of the health model and practical solutions to avoid them. Finally, this thesis added to the evidence that long-term exposure to traffic-related air pollution is associated with atherosclerosis and identified potential susceptibility factors for the Mediterranean population.

RESUMEN

La evidencia epidemiológica de los efectos de la exposición a largo plazo a la contaminación atmosférica sobre el proceso aterogénico es limitada. Dada la exposición generalizada de la población a contaminantes procedentes del tráfico y la elevada y creciente carga de morbilidad que suponen las enfermedades cardiovasculares a nivel mundial, entender y medir la asociación entre la exposición a contaminación atmosférica y la aterosclerosis es de gran relevancia para la salud pública. Los objetivos de esta tesis son caracterizar la exposición a largo plazo a contaminantes derivados del tráfico y evaluar su asociación con aterosclerosis subclínica en la población adulta de la región Mediterránea de España.

Utilizamos datos del re-examen (2007-2010) de 2780 participantes con edades entre 32 y 86 años procedentes del estudio REGICOR. REGICOR es una cohorte poblacional prospectiva de Girona, España. El principal indicador de salud fue el grosor de la íntima media de la arteria carótida (IMT), un marcador validado de aterosclerosis subclínica que se adicionó al protocolo del estudio para el re-examen de 2007-2010. Para caraterizar la contaminación a escala local medimos las concentraciones de metales pesados durante ocho semanas –usando musgos como biomonitores-, NO_2 durante cuatro semanas –mediante muestreadores pasivos- y medimos simultáneamente la intensidad de tráfico y la concentración de partículas ultrafinas (UFP) –usando contadores de partículas por condensación portátiles- durante periodos de 15 minutos. Para evaluar la variabilidad espacial de estos marcadores de contaminación por tráfico, aplicamos modelos de regresión "land use" (LUR). Además, desarrollamos un estudio metodológico sobre el efecto del número de medidas sobre la bondad de ajuste de los modelos LUR. Entre estos marcadores, el NO_2 y la intensidad de tráfico se utilizaron para calcular la exposición. La exposición a largo plazo se calculó mediante el promedio ponderado en el tiempo de las predicciones de NO_2 en todas las residencias de cada participante en los últimos diez años. La asociación transversal entre contaminación atmosférica e IMT se investigó mediante análisis de regresión lineal múltiple.

(i) El mejor determinante de los metales pesados fue el número de líneas de autobús en la calle más cercana. Los metales no estaban altamente correlacionados con el NO_2 y mostraron mayor

variabilidad espacial que este último. Los modelos LUR explicaron entre el 40 y el 85% de la variabilidad de los metales y el 72% de la variabilidad del NO_2. Dada la elevada toxicidad, la fuerte asociación con el tráfico local y la mayor variabilidad espacial de los metales pesados en comparación con el NO_2, el monitoreo con musgos es una alternativa atractiva para la evaluación de la exposición a largo plazo. **(ii)** Los mejores predictores de UFP fueron la intensidad de tráfico, distancia al cruce más cercano y el área de alta densidad residencial. Los modelos LUR de UFP explicaron entre el 36 y el 51% de la variabilidad total de UFP. Para un subconjunto de los puntos de muestreo en que se utilizó el promedio de dos medidas repetidas, la R^2 alcanzó el 72%. La distribución de las partículas ultrafinas puede modelarse con una bondad de ajuste aceptable basándose en monitoreos móviles de corto tiempo. **(iii)** Los modelos basados en pocos puntos de muestreo (de 20 a 30 puntos), resultan en R^2 ajustados y R^2 de "leave-one-out cross-validation" artificialmente elevados. Las diferencias entre estos estimadores y el R^2 verdadero puede llegar a ser del 60% para tamaños de muestra pequeños ($\leq$20) y de alrededor del 10% para tamaños de muestra mayores a 100. La selección de las variables predictoras para el modelo LUR introduce un error de medida clásico que resulta en estimaciones sesgadas del efecto de la exposición sobre la salud. La magnitud de dicho error depende del número de puntos de muestreo y del número de variables ofrecidas al modelo. Los modelos LUR para áreas urbanas complejas deberían basarse en un tamaño de muestra grande (>80) y el número de variables independientes iniciales debería ser limitado. **(iv)** La intensidad de tráfico residencial promedio en los últimos 10 años se asocia con un engrosamiento del 2% de la íntima media. Las asociaciones entre el NO_2 residencial con la IMT fueron débiles y alcanzaron la significación estadística en los subgrupos de población con alto nivel educativo y hombres de más de 60 años (una diferencia de 25$\mu g/m^3$ en NO_2 se asoció a un incremento del 4.6 y el 4.3% de IMT respectivamente). La exposición a largo plazo a contaminantes atmosféricos provenientes del tráfico se asocia a un incremento del riesgo de aterosclerosis subclínica en una muestra aleatoria de individuos sanos.

Esta tesis contribuye a la caracterización de la distribución espacial del NO_2, los metales pesados y las UFP en Girona, a la metodología de la evaluación de la exposición a contaminación atmosférica habiendo estudiado diferentes marcadores de contaminación por tráfico, instrumentos de medida y protocolos, así

como al desarrollo de la técnica de modelos LUR para los que se identificaron sesgos potenciales sobre la estimación del efecto sobre la salud de dicha exposición y aporta soluciones prácticas para evitarlos. Finalmente, esta tesis añade a la evidencia de que la exposición a largo plazo a la contaminación procedente del tráfico se asocia a la aterosclerosis e identifica factores de susceptibilidad potenciales para la población mediterránea.

RESUM

L'evidència epidemiològica de l'efecte de la contaminació atmosfèrica sobre el procés crònic de l'aterogènesi és limitada. Donada l'exposició generalitzada de la població a contaminants procedents del trànsit i a l'elevada i creixent càrrega de morbiditat que suposen les malalties cardiovasculars a nivell mundial, entendre i quantificar l'associació entre l'exposició a contaminació atmosfèrica i l'aterosclerosi és de gran rellevància per la salut pública. Els objectius d'aquesta tesi són caracteritzar l'exposició a llarg termini a contaminants derivats del trànsit i avaluar la seva associació amb l'aterosclerosi subclínica en la població adulta de la regió mediterrània d'Espanya.

Vàrem utilitzar dades de la re-avaluació (2007-2010) de 2780 participants amb edats entre 32 i 86 anys procedents del estudi REGICOR. REGICOR és una cohort poblacional prospectiva de Girona, Espanya. El principal indicador de salut va ser el gruix de la íntima mitjana de l'arteria caròtida (IMT), un marcador validat d'aterosclerosi subclínica que es va afegir al protocol de l'estudi en la re-avaluació del 2007-2010. Per a caracteritzar la contaminació a nivell local vàrem mesurar les concentracions de metalls pesats durant vuit setmanes -mitjançant molses com a biomonitors-, diòxid de nitrogen (NO_2) durant quatre setmanes -mitjançant mostreig passiu- i vàrem mesurar simultàniament la intensitat de trànsit i la concentració de partícules ultra-fines (UFP) -mitjançant comptadors portàtils de partícules per condensació- durant períodes de 15 minuts. Per avaluar la variabilitat espacial d'aquests marcadors de contaminació per trànsit, vàrem aplicar models de regressió "land use" (LUR). Així mateix, vàrem desenvolupar un estudi metodològic sobre l'efecte del nombre de mesures sobre la robustesa dels models LUR. D'entre aquests marcadors, el NO_2 i la intensitat de trànsit es van utilitzar per calcular l'exposició. L'exposició a llarg termini es va calcular com la mitjana de NO_2 ponderada pel temps viscut en totes les residències en els darrers deu anys. L'associació transversal entre contaminació atmosfèrica i IMT es va investigar mitjançant anàlisi de regressió lineal múltiple.

(i) El millor determinant dels nivells de metalls pesats va ser el nombre de línies d'autobús al carrer més proper. Els metalls no estaven altament correlacionats amb NO_2 i varen mostrar major variabilitat espacial que el NO_2. Els models LUR varen explicar entre el 40 i el 85% de la variabilitat dels metalls i el 72% de la variabilitat del NO_2. Donada la seva elevada toxicitat, la forta associació amb el trànsit local i la major variabilitat espacial dels

metalls pesats en comparació amb el NO_2, el monitoratge amb molses és una alternativa atractiva per l'avaluació de la exposició a llarg termini. **(ii)** Els millors predictors de UFP varen ser intensitat del trànsit, distància a l'encreuament més proper, àrea de gran densitat residencial i densitat d'habitatges. Els models LUR de UFP varen explicar entre el 36 i el 51% de la variabilitat total de UFP. Quan per a una mostra de punts de mostreig amb dues mesures repetides es va utilitzar la mitjana de les dues mesures, la R^2 va arribar al 72%. La distribució de partícules ultra-fines pot modelar-se amb una bondat d'ajust acceptable basant-se en un monitoratge mòbil de curta durada. **(iii)** Els models basats en un nombre petit de punts de mostreig (de 20 a 30 punts), resulten en R^2 ajustats i R^2 de validació creuada artificialment elevats. Les diferencies entre aquestes estimacions i el verdader R^2 pot arribar a ser del 60% per grandàries de mostra petites ($\leq$20) i del voltant del 10% per grandàries de mostra per sobre de 100. La selecció de les variables predictores per al model LUR introdueix un error de mesura clàssic que resulta en estimacions esbiaixades de l'efecte. La magnitud d'aquest error depèn del nombre de punts de mostreig i de les variables ofertes al model. Els models LUR per a espais urbans complexos haurien de basar-se en una gran mostra de punts (>80) i el nombre de variables independents inicials hauria de ser petit. **(iv)** La mesura d'intensitat i volum de trànsit a les residències dels últims 10 anys s'associa amb un engruiximent del 2% de la íntima mitjana. Les associacions entre el NO_2 residencial dels últims 10 anys amb IMT varen ser febles i varen assolir la significació estadística en els sub-grups de població amb alt nivell educatiu i homes de més de 60 anys (una diferència de 25$\mu g/m^3$ en NO_2 es va associar a un increment del 4.6 i el 4.3% de IMT respectivament). L'exposició a llarg termini a contaminants atmosfèrics provinents del trànsit es va associar a un increment del risc d'aterosclerosi subclínica en una mostra aleatòria d'individus sans.

Aquesta tesi contribueix a la caracterització de la distribució espacial del NO_2, els metalls pesats i les UFP a Girona, a la metodologia de l'avaluació de l'exposició a contaminació atmosfèrica utilitzant diferents marcadors de contaminació per trànsit, instruments de mesura i protocols, així com al desenvolupament de la tècnica LUR per identificar biaixos potencials sobre l'estimació de l'efecte en salut d'aquesta exposició i solucions pràctiques per a evitar-los. Finalment, aquesta tesi contribueix a l'evidència que l'exposició a llarg termini a la contaminació procedent del trànsit s'associa a l'aterosclerosi i identifica factors de susceptibilitat potencials per a la població mediterrània.

INDEX

1 ATHEROSCLEROSIS AND CARDIOVASCULAR DISEASES

Cardiovascular diseases (CVD) are a group of disorders that affect the cardiovascular system, including heart diseases, and diseases of the blood vessels. Although they can be due to several causes, the more prevalent are related to atherosclerosis, a progressive disease characterized by the accumulation of lipids and fibrous elements in the large arteries (Lusis, 2000) (**Figure 1**). CVD not related to atherosclerosis include congenital heart disease, rheumatic heart disease, cardiomyopathies, cardiac arrhythmias and inflammatory heart diseases (myocarditis, pericarditis, endocarditis and cardiomyopathy). Those related to atherosclerosis include ischaemic heart disease (heart attack), cerebrovascular disease (stroke or ictus) and diseases of the arteries (hypertensive heart disease and peripheral vascular disease). The latter cause 83% and 86% of total mortality due to CVD in women and men respectively.

The process of atherosclerosis begins in childhood (McGill et al., 1998) and there is a long induction period until evolving to clinical stages of the disease in middle and advanced ages. By measuring atherosclerosis, instead of cardiovascular events, the subclinical stages of the disease can be studied.

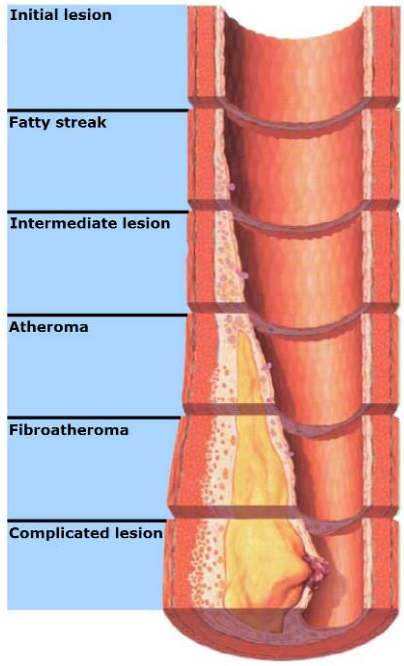


Figure 1. Illustration of the progression of atherosclerosis in the coronary arteries. Adapted from (Wikimedia Commons, 2007)

1.1 Burden of disease

Cardiovascular diseases (CVD) accounted for 17.3 million deaths worldwide in 2008 (World Health Organization, 2011). This represents 31% of all-cause mortality (32% in women and 27% in men), making CVD the leading causes of death worldwide. The elderly population is the most affected by CVD. It is however estimated that more than 3 million of the global deaths due to CVD occurred before the age of 60 (2008). CVD were responsible for 151 million Disability Adjusted Life Years (DALY), accounting for 10% of the total DALY estimated on a global scale for the year 2004. More than half of this burden involved people aged less than 60 years (World Health Organization, 2008). In Spain, CVD are the main cause of death accounting for 120.053 deaths in 2009, which represents 44% of all-cause death in women and 32% in men (INE, 2011), i.e. 31.2% of overall total deaths.

In the recent decades, the incidence of CVD has doubled in low- and middle-income countries. This and the aging of population will lead

to an increase in the global cardiovascular deaths in 2030 to a projected estimate of 23.4 million (World Health Organization, 2008).

1.2 Mechanisms and Risk factors

Atherosclerosis is a complex disease and its mechanisms are not fully elucidated. Biological pathways include systemic inflammation, endothelial dysfunction, alterations of extracellular matrix metabolism and hyper-reactivity of smooth muscle cells. Inflammation is triggered after injuries in the arterial wall. Such injuries are caused mainly by LDL molecules that are deposited on the wall and then oxidized by free radicals, becoming toxic. In response, inflammatory cells produce cytokines that activate the endothelium, shifting completely its antiadhesive and anticoagulant properties into adhesive and procoagulant properties (Angiolillo et al., 2004). Monocytes and T lymphocytes are then allowed to migrate from the blood stream into the arterial wall. Monocytes acquire characteristics of macrophages and foam cells and secrete reactive oxygen species, cytokines, metalloproteinases growth factor and tissue factor increasing the local inflammatory response (**Figure 2**). Over time, accumulation of macrophages, foam cells and cellular necrotic debris form atherosclerotic plaques that are covered by a fibrous cap. The plaque development results in arterial wall thickening. In late stages of the disease arterial stenosis (abnormal narrowing) may occur. Also, macrophages can secrete metalloproteinases that degrade the fibrous cap until its rupture. Blood enters in contact with tissue factor, a pro-coagulant protein, and thromboses are formed. Procoagulant properties of activated endothelial cells may alternatively lead to thrombi formation, which eventually leads to acute clinical manifestations, such as myocardial infarction, stroke or vascular dementia.

The evolution of atherosclerosis follows a life-long pathway and may or may not lead to an acute event. Many patients with atherosclerosis remain stable all through their life, while others, even with less severe atherosclerosis, develop an acute event as the first symptom of CVD. Thus, the mechanisms and also the risk factors for the process of atherogenesis and those for the development of an acute event are not necessarily the same (Angiolillo et al., 2004; Künzli et al., 2011).

Current concepts assume that the degree of atherosclerosis depends both on genetic factors and the total cumulative exposure to endogenous and exogenous (environmental) risks factors. Risk factors for acute events are the presence of atherosclerosis and circumstances occurring within one hour and up to ten days before the event, such as, a transient thrombotic change due to an infection (Künzli et al., 2011), use of cocaine, heavy meals, physical exertion, stressful situations or high levels of air pollution (Nawrot et al., 2011). Established risk factors for atherosclerosis development are high cholesterol levels, hypertension, diabetes, overweight and obesity, metabolic syndrome (Iglseder et al., 2005), sedentary lifestyle, unhealthy diet, smoking, high intake of alcohol, family history of atherosclerosis or CVD, poverty, psychological factors (e.g. hopelessness) (Whipple et al., 2009) and long-term radiation therapy (close to the heart) (Wittig et al., 2011). Unmodifiable risk factors are sex –with men being at higher risk for development of clinically relevant atherosclerosis- and age. In fact atherogenesis is inherently related to aging. Artery walls get thicker continuously and disease progression depends on the overall level and duration of exposure to the mentioned risks factors.

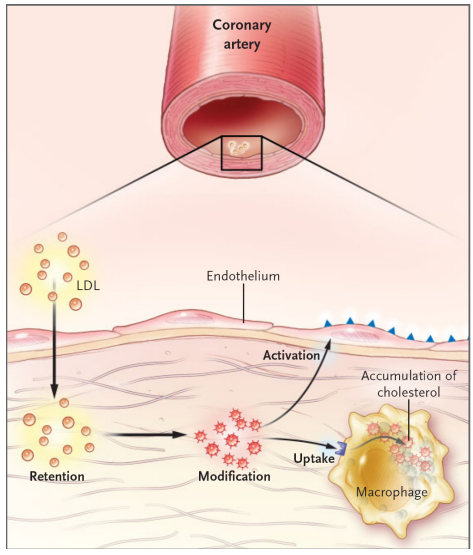


Figure 2. Illustration of the dynamic pathophysiology of atherosclerosis showing relationships between potential metabolic, cellular, and inflammatory biomarkers. Reproduced from Hansson (2005).

1.3 Measurements of atherosclerosis

A review of the surrogate measures of atherosclerosis in air pollution research was presented by Künzli et al. (2011). A brief description of the most common surrogates and a detailed description of the ones used in this thesis, namely carotid intima media thickness and ankle brachial index, are provided.

Carotid intima media thickness (IMT) measures the thickness of the artery walls, which is used as a marker of subclinical atherosclerosis. IMT is widely used in clinical trials and epidemiological studies and relies on non-invasive high-resolution ultrasound imaging coupled with automatic data processing systems (see **Figures 4** and **5**). It is associated with increased risks of CVD and with risk factors correlated with development and progression of carotid atherosclerosis in both population-based samples and high-risk populations (de Groot et al., 2008; Mancini et al., 2004). IMT predicts coronary heart disease, cardiovascular (MI, coronary death) and cerebrovascular events, such as stroke (Revkin et al., 2007), in both patients and asymptomatic individuals (Holewijn et al., 2010).

Several studies report that IMT measurements are reliable and reproducible, although there is not a standardized measurement protocol and the optimal site of measurement remains controversial (common carotid, bifurcation, internal carotid, posterior (far) or anterior wall). Comparison between studies can thus be difficult. Longitudinal studies to assess the effect of clinical interventions have found that a reduction in cholesterol (total and LDL) is associated with regression of IMT in most cases and with inhibition of IMT progression in other cases (de Groot et al., 2008). Regarding prediction of clinical outcomes, it has been suggested that the reduction in CV risk factors is not always associated with reduction in IMT progression (Mancini et al., 2004), however, research in this topic is not conclusive. For example, angiotensin-converting enzyme (ACE) inhibitors reduce CV events in high-risk patients. However, several but not all clinical trials suggested that ACE inhibitors reduce IMT progression (Lonn et al., 2009). This could be explained by differences in measurements protocols, observational period, or by a pathway that is not mediated by IMT or atherosclerosis reduction (lower blood pressure, reduced left ventricular mass or reversal of endothelial dysfunction has been suggested (MacMahon et al., 2000)).

Arterial stiffness is a measure of the vessel elasticity during systole after applying pressure. This measurement relies on non-invasive ultrasound imaging or tonometry and measurement of blood pressure simultaneously. It predicts cardiovascular events, cardiovascular morbidity and all-cause mortality (Cavalcante et al., 2011).

Ankle brachial index measures the ratio of blood pressure in the peripheral arteries (in the legs) to that in the proximal segments (in the arms). Low blood pressure in the legs, compared to the arms, indicates narrowing or blockage of the peripheral arteries. ABI is thus a marker of the degree of subclinical peripheral atherosclerosis and the severity of stenosis. Low ABI at rest predicts cardiovascular outcomes both in persons with underlying CVD and low-risk population. An index lower than 0.9 is associated with higher morbidity and mortality risks and is used as a cutoff point to diagnose peripheral arterial disease (PAD). An index higher than 1.3 has also been associated to higher mortality, calcification of the arterial wall and CVD risk factors. It relies on blood pressure measurements using Doppler technique. It is a simple, non-invasive and inexpensive test (Holewijn et al., 2010).

Angiography is the former "gold standard" measurement of atherosclerosis (Revkin et al., 2007). It consisted in injecting a radiopaque contrast media in the blood vessels followed by x-ray imaging. The lumen of the vessels could then be measured from the images. The invasive nature of this technique with its inherent risks and limitation to symptomatic patients, and the exposure to radiation drove innovations toward non-invasive imaging techniques.

Coronary artery calcification and aortic artery calcification are measures of coronary and systemic atherosclerosis, respectively, that predict clinical cardiovascular events (Pletcher et al., 2004). They rely on electron-beam or multi-detector computed tomography, non-invasive techniques that, however, expose to radiation.

Retinal vessel diameter measure the micro-vascular atherosclerosis, which is related to cerebrovascular diseases, coronary calcifications, coronary morbidity and mortality, and risk of heart failure, and stroke (Mimoun et al., 2009). It relies on non-invasive retinal photography that poses no risks.

The use of these and other measurements in air pollution research has been discussed by Künzli et al. (2011). A summary table is provided in **Table 1**.

Table 1. Scoring with regard to 5 criteria determining the usefulness of the outcome measure to investigate long-term atherogenic effects of ambient air pollution in epidemiological studies. Reproduced from (Künzli et al., 2011).

	Continuous Measure	Marker for Long-Term Effects	Specific for Artery Wall Pathology	Predict CVD	Practical Issues (Low Costs; Low Invasiveness)
Morphological characteristics of the arterial wall					
CIMT	++	++	++	+	++
CAC	+	++	++	++	+
AOC	+	++	++	+	+
Retinal vessel diameter	++	+	+	0	++
Measures based on functional performance of the vessel					
ABI	++	+	+	+	++
Orbital blood flow velocity	++	+	+	0	++
Arterial stiffness, compliance, elasticity	++	+	++	+?	++
Measures based on end-organ involvement					
Renal function	++	+	0	+	++
Ventricular mass	++	++	0	+	+
MI	0	0*	++	N/A	++
Cardiovascular death/mortality	0	0*	+	N/A	++

Scores are assigned relative to other outcome measures. Most appropriate features are scored with ++. A 0 score means that the outcome is not a good choice for the respective criterion. For a description of the outcome measures, see text.

* Marker for short-term effects of air pollution.

2 EXPOSURE TO AIR POLLUTION ASSOCIATED TO ATHEROSCLEROSIS

2.1 Previous evidence

The American Heart Association recently concluded that there is a causal relationship between traffic and combustion related air pollution and cardiovascular mortality and morbidity (Brook et al., 2010). Most of the evidence is on the associations between short-term exposure to air pollution and clinical cardiovascular outcomes, such as acute myocardial infarction (AMI), stroke, all cause- and cardiovascular-mortality (Dominici et al., 2003; Katsouyanni et al., 2001; Ostro et al., 2011), and with subclinical measures or surrogates of CVD, such as increase in oxidative stress and systemic inflammation. There is epidemiological evidence on the association of long-term exposure with cardiovascular mortality and ischaemic heart disease, but more investigation on its association with subclinical measures or surrogates of CVD is needed. Particularly, it is not yet clear to what extent long-term exposure to air pollution contributes to the chronic processes of atherogenesis and how the observed acute effects due to short-term exposures may contribute to the chronic effect (Brook et al., 2010).

There is evidence that long-term exposure to ambient particulate matter (PM) (Sun et al., 2005), ultrafine particles (UFP) (Araujo et al., 2008), and diesel and gasoline exhaust cause atherosclerosis and enhance plaque vulnerability in animals (Brook et al., 2010). But there are only a few human exposure studies. The first study was carried out in Los Angeles in a heterogeneous sample of 798 volunteers participating in two clinical trials (Künzli et al., 2005). A cross-sectional difference of 10 $\mu g/m^3$annual residential $PM_{2.5}$ was associated with a 4.2% increase in IMT (95%CI -0.2 to 8.9). This association was confirmed afterwards in a longitudinal study (Künzli et al., 2010; Künzli et al., 2005a) on 1503 participants of five clinical trials including those mentioned before. An increase of 10 $\mu g/m^3$ $PM_{2.5}$ was associated with an IMT progression rate of 2.53um per year (95%CI -0.31 to 5.38). Annual cross-sectional differences of 4.2 $\mu g/m^3$ $PM_{2.5}$ were associated with 4.3% increase (95%CI 1.9 to 6.7%) in IMT while no association was found for PM_{10} in 3380 population-based participants of the Heinz Nixdorf Recall study in Germany (Bauer et al., 2010). In a related study from the same

cohort, $PM_{2.5}$ was nonsignificantly associated with higher coronary artery calcification (Hoffmann et al., 2007). A population-based study in 750 young adults in the Netherlands found no association of annual residential NO_2, $PM_{2.5}$, SO_2, black smoke, proximity to traffic and traffic intensity with IMT. Among the same measured pollutants, NO_2 and SO_2 were associated with arterial stiffness (Lenters et al., 2010). In a population-based sample of 5172 participants from the US Multi-Ethnic Study of Atherosclerosis (MESA) 20-yr average contrast in 2.5 $\mu g/m^3$ $PM_{2.5}$ and 21 $\mu g/m^3$ PM_{10} were associated with 1% increase (95%CI 0 to 2%) and 3% increase (95%CI 1 to 5%) in IMT, respectively, while no association with carotid artery calcification was found (Diez Roux et al., 2008). In a cross-sectional study in 1147 participants of the same cohort, a difference of 10 $\mu g/m^3$ of $PM_{2.5}$ was associated with higher prevalence of abdominal aortic calcium (OR 1.06, 95%CI 0.96 to 1.16). In an occupational health study in Turkey, 61 highway toll collectors had 0.8mm thicker IMT (SD 0.2) on average, while 41 controls had IMT of 0.6mm (SD 0.1) (Erdogmus, 2006).

Among the mentioned studies, four investigated the associations of subclinical atherosclerosis with proximity to traffic. Additionally, a cross-sectional study on children in Salerno, Italy (Iannuzzi et al., 2010) also evaluated this association. Only one of these studies found an association between residential proximity to traffic and IMT. In the study on volunteers of clinical trials, for participants living within 100m of a highway, representing 1.55% of the total sample, the progression of atherosclerosis was 5.46µm/yr (95%CI 0.13 to 10.8) faster than among those living further away.

No association was found between exposure to PM and ABI. Neither 20-yr residential exposure to $PM_{2.5}$ nor PM_{10} (Diez Roux et al., 2008) in the MESA study and one year exposure to $PM_{2.5}$ (Hoffmann et al., 2009b) in 4302 participants from the Recall study, were associated with ABI. The Recall study also investigated associations with residential proximity to high traffic. This variable was associated with both ABI and prevalence of PAD. An odds ratio for PAD of 1.77 (95%CI 1.01 to 3.10) was found for participants living within 50m of a major road (10.000 to 130.000 veh/day) compared to those living more than 200m away (Hoffmann et al., 2009b).

The characteristics of the study population, exposure assessment, main results and effect modifications of the mentioned studies, as

well as published studies on other subclinical markers of CVD are summarized in **Table 2.**

Table 2. Population characteristics, exposure assessment, main results and effect modifications of the published epidemiological studies on subclinical markers of CVD. Adapted from Künzli et al. (2011)

Study	Population	Mean age (range) (yr.)	Endpoint	Exposure
Künzli et al. 2005	798 participants from 2 clinical trials, USA	59 (≥40)	IMT	Annual residential $PM_{2.5}$
Künzli et al. 2010	1503 participants from 5 clinical trials, USA	59 (30-89)	IMT progression (longitudinal)	Annual residential $PM_{2.5}$
				Proximity to traffic
Bauer et al. 2010	3380 participants, Heinz Nixdorf Recall Study, Germany	60 (45-75)	IMT	Annual residential $PM_{2.5}$, PM_{10}
				residential proximity to major road (traffic intensity in higher quintile)
Hoffmann et al. 2007	4494 participants, Heinz Nixdorf Recall Study, Germany	60 (45-75)	Coronary artery calcification (CAC)	Annual residential $PM_{2.5}$
				Residential proximity to high traffic (10.000 to 130.000 veh/day)
Hoffmann et al. 2009	4302 subjects, Heinz Nixdorf Recall Study, Germany	60 (45-75)	ABI	Annual residential $PM_{2.5}$
				Residential proximity to high traffic
			Prevalence of PAD	Annual residential $PM_{2.5}$
				Residential proximity to high traffic (10.000 to 130.000 veh/day)
Lenters et al, 2010	750 subjects of the Atherosclerosis Risk in Young Adults Study	28.4 (27-30)	IMT	Annual residential NO_2; $PM_{2.5}$; SO_2; Black smoke; proximity to traffic; traffic intensity
			Arterial Stiffness	Annual residential NO_2
				SO_2
				$PM_{2.5}$, Black smoke, proximity to traffic, traffic intensity
Diez Roux et al. 2008	5172 subjects, Multi-Ethnic Study of Atherosclerosis	62 (45-84)	IMT	20-yr average residential PM_{10};
				20-yr average residential $PM_{2.5}$
			ABI	20-yr average residential PM_{10}, $PM_{2.5}$
			Prevalence of any CAC	20-yr average residential PM_{10}, $PM_{2.5}$
Allen et al. 2009	1147 subjects, Multi-Ethnic Study of Atherosclerosis, USA	62 (45-84)	Prevalence of Abdominal aortic calcium	2-year annual residential $PM_{2.5}$
				Residential proximity to traffic
O'Neill et al. 2007	4933 subjects, Multi-Ethnic Study of Atherosclerosis, USA		Arterial stiffness	20 year average residential PM_{10}
Iannuzzi et al, 2010	52 children, Vietri sul Mare, Salerno, Italy	Range, 6-14	IMT Arterial stiffness	Proximity to traffic
Erdogmus et al. 2006	61 highway toll collectors, 48 controls, Turkey	36.2 (24-56)	IMT	Occupational traffic exposure
Erdogmus et al. 2007	20 highway toll collectors, 20 controls, Turkey	36.2 (24-56)	Peak systolic flow in ophthalmic artery	Occupational traffic exposure
Wu et al, 2010	17 mail carriers, Sinjhuang, Taipei County, Taiwan	32.4 (22-46)	Arterial stiffness: cardio-ankle vascular index (CAVI)	Occupational traffic exposure

Method of exposure assessment; scale of estimates/measurements	Unit of exposure	Main result[1]	Interactions (Direction of interaction)
Geostatistical model, 23 monitoring stations; 25m grid cells	10 µg/m^3 $PM_{2.5}$	4.2% increase (95%CI -0.2 to 8.9)	≥60 (+); women(+); never smoker(+); lipid low.medication(+)
Geostatistical model, 23 monitoring stations; 25m grid cells	10 µg/m3 PM2.5;	2.53µm/yr progression (95%CI -0.31 to 5.38)	Income>$20K/yr(0); treatment arm (+); sex(0); lipid low.medication(0); ethnicity(0); education(0)
Road type classification	100 m to highway	5.46µm/yr progression (95%CI 0.13 to 10.8)	Income>$20K/yr(+); treatment arm (+); sex(0); lipid low.medication(0); ethnicity(0); medium education level (+) (compared to high)
Chemistry transport model, 1km grid cell	4.2 µg/m3 $PM_{2.5}$	4.3% increase (95%CI 1.9 to 6.7%); No association with PM_{10}	<60(+); bmi≥27(+); no diabetes(+); statins(0); full-time employer(+); current, former smoker(+); sex(0); coronary artery disease(0)
Traffic intensity	1,939 m (interdecile range)	1.2% increase (95%CI −0.2% to 2.6%)	Not reported
Dispersion model; 5km grid cell	3.91 µg/m^3 $PM_{2.5}$	17.2% increase (95%CI -6 to 46)	Women(+); ≥60(0); low education(0); current smoker(0)
Traffic intensity	≤50 m vs. >200 m	OR 1.63 (95%CI 1.14 to 2.33)	Men(+); <60(+); low education(+); current smoker(0)
Dispersion model; 5km grid cell	3.91 µg/m^3 $PM_{2.5}$	No association	Women(0), Smoker(0), BMI≥30(0), statin(0); ≥60(0); diabetes(0)
Traffic intensity	≤50 m vs. >200 m	-0.024 (95%CI -0.047 to -0.001)	Women(+); Smoker(+); BMI≥30(+); statin(+); ≥60(0); diabetes(0)
Dispersion model; 5km grid cell	3.91 µg/m^3 $PM_{2.5}$	OR 0.87 (95%CI 0.57 to 1.34)	Women(0), Smoker(0), BMI≥30(0), statin(0); ≥60(0); diabetes(0)
Traffic intensity	≤50 m vs. >200 m	OR 1.77 (95%CI 1.01 to 3.10)	Women(+), Smoker(+), BMI≥30(0), statin(0); ≥60(+); no diabetes(+)
LUR for Regional background, urban background & local component.	25 µg/m3 NO_2	No association	Women(0); low education(0); current smoker(+)
	5 µg/m3 SO_2	4.05 (95%CI 0.13 to 7.97) pulse wave vel. 37.5 (95%CI 2.2 to 72.9) Augmentation index 5.26 (95%CI 0.09 to 10.43) pulse wave vel. No association	Women (+) (p=0.003 for augmentation index); education (0); never, former smoker (+) for pulse wave vel; never smoker (+) for augmentation index
Space-time model; mean distance from address to nearest monitor 13Km (SD 33Km)	21 µg/m^3 PM_{10}	3% increase (95%CI 1 to 5%)	Age(0); sex(0); lipid status(0); smoking(0); diabetes(0); bmi(0); education(0)
	12.5 µg/m^3 $PM_{2.5}$	1% increase (95%CI 0 to 2%)	
		No association	
		No association	
Universal krigging if >10 monitors per city, else: inverse distance weighting; mean distance from address to nearest monitor per city ranged between 2.2km (SD 1.1km) to 7.1km (SD 4.5km)	10 µg/m^3 $PM_{2.5}$	OR 1.06 (95%CI 0.96 to 1.16)	Women (+), >65(+), lipid low.medication(+); Hispanics(+), diabetes(+); BMI≥30(0), education(0)
Road type classification	≤100 m from highway or ≤50 m from major road	No association	Income>$50K/yr(+); sex (0); >65(0); lipid low.medication(0); ethnicity(0); diabetes(0); BMI≥30(0), high education(+)
	10 µg/m^3 PM_{10}	No association	Age(0); sex(0); race(+); waist/hip(0); former smoking(-); diabetes(0); bmi(0); education(0); phys.act(0)
Distance from only road with high traffic in town	Within 300 m to major road	No association Absolute difference 1.58 (P = .007)	NA
	Exposure vs. no exposure	0.8mm (SD 0.2) vs. 0.6mm (SD 0.1) (unadjusted)	NA
	Exposure vs. no exposure	36.8 (SD 2.2) cm/s vs. 33.2 (2.4) cm/s (unadjusted)	NA
$PM_{<0.25}$; $PM_{0.25-0.5}$; $PM_{0.5-1}$; $PM_{1-2.5}$; $PM1$ (sum of 4 last); $PM_{>2.5}$ (coarse); ozone	7.2 µg/m3 $PM_{1-2.5}$; 17.6 ppb ozone	2.5% increase in CAVI (95%CI 0.3 to 4.8) for $PM_{1-2.5}$; 4.8% (95% CI: 1.6 to 8.2) for ozone; no asociation for the rest	NA

2.2 Potential mechanisms involved

Exposure to air-pollutants can mediate CVD through direct and indirect effects. Direct effects include systemic transmission of signals and translocation of gaseous pollutants, particle soluble constituents and possibly ultrafine particles into the blood and cardiovascular system, mediating acute effects within hours of pollutants inhalation. The direct pathway is thoroughly explained by Simkhovich et al. (2008) and Brook et al. (2010). The indirect effects include perturbations to the autonomic nervous system, which increases the low- and high-frequency heart rate variability, and pulmonary and systemic oxidative stress and inflammation, which is relevant to the atherogenic process.

Pulmonary oxidative stress causes the release of pro-inflammatory mediators (e.g., cytokines, activated immune cells, or platelets) or vasculoactive molecules (e.g., circulating endothelin , possibly histamine, or microparticles) (Brook et al., 2010). But the inflammatory responses go beyond the lungs: exposure to particulate matter increases the serum levels of IL-6, which determines the synthesis of C-reactive protein, a marker of systemic inflammation directly associated with higher risk of CVD. The latter is also positively associated with exposure to total suspended particles (TSP) and PM. Inflammation through cytokines and C-reactive protein promotes atherosclerosis via the mechanisms previously described (Section 1.2).

3 MEASUREMENT OF TRAFFIC-RELATED AIR POLLUTION

3.1 Surrogates of traffic-related exposure

The effects of traffic-related pollution on health started to be studied in the 1920's as part of more general studies of the effects of air pollution (Henderson Y, 1921; Sayers et al., 1924). In 1931, the New York academy of medicine reported on the deleterious effects of prolonged exposure to heavy traffic presumably due to carbon monoxide levels (The New York Academy of Medicine, 1931). Exposure to traffic-related pollutants is, most of the times, still measured indirectly, for example, using atmospheric concentrations of one or a few components of the traffic-emitted mixture used as surrogates. The complex mixture of pollutants emitted by traffic is mainly generated in the fuel combustion process, but also in the resuspension of dust particles and brake and tire wearing. The most commonly used surrogates are NO_2, NO_x, particulate matter and SO_2, among others.

The most widely used surrogate in health effect studies is **nitrogen dioxide (NO_2)**. This is because affordable samplers have been developed for its passive measurement (based on diffusion without the need for air pumping); NO_2 is regulated with standards available worldwide (See **Appendix 1**) and standardized measurement protocols; most NO emitted in combustion processes – traffic exhaust and industry- is converted to NO_2; and it is the main contributor to the formation of tropospheric ozone (Brook et al., 2004).

Particulate matter (PM), a mixture of solid and liquid particles generated by direct emissions (primary PM) and particle formation (secondary PM), is also widely used as a marker of traffic emissions. Primary particles are mainly composed of elemental carbon (EC, also called soot and black carbon), polycyclic aromatic hydrocarbons and metals. Secondary particulate matter is produced by the oxidation in the atmosphere of precursor pollutants such as NO_x (53%), SO_2 (22%) and NH_3. Its composition depends on time (hour, day, season), climate variations, geographical position and emission sources. Sources of particulate matter include vehicle emissions, tire fragmentation and resuspension of road dust, fossil

fuel combustion, metal processing, construction agriculture, domestic heating, wood burning, forest fires, cigarette smoke, natural trees, windblown soil, pollens and molds, forest fires, volcanic emissions, and sea spray. PM is classified depending on the diameter of the particles in PM_{10} the fraction with diameter less than 10 μm, coarse particles ($PM_{2.5}$–PM_{10}, diameter 2.5–10 μm), fine particles ($PM_{2.5}$, diameter <2.5 μm), and ultrafine particles (UFP or $PM_{0.1}$, diameter <0.1 μm). Nowadays, traffic contributes to approximately 15% (with a range of 4% to 25%) of $PM_{2.5}$ at seven U.S. EPA supersites (Watson et al. 2008). In Barcelona, traffic was the main source of $PM_{2.5}$ (vehicle exhaust contributed 30% of total mass) and PM_{10} (vehicle exhaust contributed18%, and road dust 17% of total mass) in weekly measurements, since 2003 to 2007 in an urban background monitoring station (Ostro et al., 2011). Measurements of particulate matter are obtained using a pump to draw ambient air at a constant flow rate into a specially shaped inlet, where particulate matter is separated into size fractions, and through a filter. Airborne particulate matter is then collected on the filter. The concentration of particles in the air is calculated dividing the particulate matter weight by the total volume of air filtered.

Air quality guidelines of the European Union for these pollutants are presented in **Appendix 1**.

Ultrafine particles (UFP) acquired relevancy in the last years due to their high proportion in the number of particles within PM, their high toxicity and capacity to reach the alveolar area in the lungs and, potentially, the circulatory system (Brook et al., 2010). Main sources of UFP include fuel combustion, wood smoke, industrial sources, and formation of new particles (nucleation) by chemical reactions with sulfate, nitrate and organic PM as precursors. The atmospheric lifetimes of UFP are short, and their concentrations decrease rapidly with distance from the source –such as roads- (Zhu et al., 2006). Between 52 to 86% of UFP (size range 20-800 nm) annual average (2003-2004) was emitted by vehicles at an urban background monitoring station in Barcelona (Pey et al., 2009). UFP can be measured as mass concentration in a similar way as described above for PM measurements or as number concentration by optical instruments that count the number of particles in a given volume of air pumped. These instruments have been used for short-term personal monitoring. Technology is, however, not yet ready for large-scale applications in epidemiologic studies because of the cost

of the monitors and the burden for the subjects (HEI Panel on the Health Effects of Traffic-Related Air Pollution, 2010).

Specific constituents of particulate matter have also been used as traffic markers due to their high chemical reactivity and toxicity. **Heavy metals** are highly biorreactive elements –some of them, such as chromium, lead and arsenic can induce carcinogenic effects- that get bound to particulate matter of all size fractions. More toxic metals, however, show higher concentrations in the finer fractions of PM, for example lead, nickel, cadmium, copper, zinc and arsenic (Moreno et al., 2006). Due to specificity of metals as constituents of e.g. catalytic converters, brake pads or fuel additives, some are used as markers of emissions from those sources, for example, vanadium, chromium and nickel are related to fuel combustion; copper, antimony and molybdenum are markers of brake emissions. Metal concentrations for air pollution studies have been measured in tailpipe emissions, brake and tire wear emissions, roadway dust, roadway tunnel and ambient air. Only the latter have been used in epidemiological studies in humans. Heavy metal measurements consist in collecting particulate matter followed by chemical elemental analysis to identify the concentrations of metals in the PM collected.

Exposure to traffic-related pollutants can also be assessed indirectly by measuring nearby **traffic** as a surrogate for pollutants from vehicle exhausts and/or resuspended dust (dust deposited on the road that is resuspended by passing traffic, wind and turbulences). Traffic measures include vehicle mix (i.e. diesel and gasoline-fueled or light and heavy-duty vehicle intensity), traffic intensity also called traffic density/volume/count (daily number of vehicles) in a given street or within a given buffer (e.g. 100m), traffic load (sum of traffic intensity multiplied by length of road segment in all segments of a given buffer), length of street segments in a given buffer and self-reported traffic exposures (e.g. nuisance of road traffic). All these approaches attempt to characterize exposure to near-road pollutants, thus, the term 'proximity models' is often used. **Figure 3** shows the GIS maps and information used to calculate these measurements.

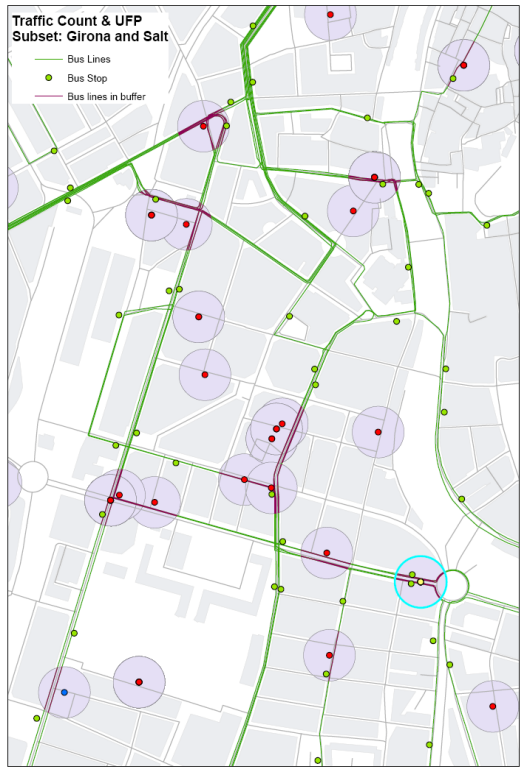


Figure 3. Example of GIS maps and information layers used to measure traffic proximity markers as surrogates of air pollution exposure. Red dots are monitoring sites. Purple circles are buffers around the monitoring sites.

3.2 Exposure assessment

As explained above, atherogenesis is a long-term process that develops over life time. The degree of atherosclerosis reached at any point in time is the cumulative sum of all pro- and anti-atherogenic exposures. Accordingly, from an air pollution perspective, the relevant time-window for exposures leading to atherosclerosis development is as well the accumulated life-time exposure. Whereas life-time exposure is ultimately the sum of exposures over time – e.g. the sum of all daily exposures – the degree of atherosclerosis measured at some point in time point is not dependent on 'yesterdays' pollution'. The latter is informative to investigate the role of short-term exposure in triggering acute events. Instead, studies investigating the atherogenic role of pollution

need to assign some measures of long-term exposure. Being practically impossible to measure exposure during lifetime, studies have used estimates of pollution over shorter 'long-term windows', such as the average ambient concentrations of one or a few years, as a proxy. The detailed interplay between long- and short-term exposure in the development of atherosclerosis remains to be explored. However, there is no evidence of short-term pollution events such as isolated peaks (e.g. due to peaks in emission levels– e.g. traffic rush hours, unusual meteorological conditions, travelling to places with very contrasting pollution levels or compositions, or a smog episode etc.) and/or independent relevance for the chronic health effects. Therefore, studies on long-term effects of air pollution assume that the long-term mean concentrations (which include those peak conditions as well) reflect well the accumulated exposure.

From a study design perspective, it is important to distinguish temporal and spatial variability of ambient air pollutants. Temporal variability is mostly driven by meteorological factors, which cause simultaneous variations in air pollution components, including those from different sources (Brauer, 2010). Since general meteorological conditions usually influence larger regions simultaneously, the temporal variability of air pollution within one urban area (or air shed) is typically rather homogeneous across space. In other words, the daily fluctuations are usually paralleled across larger areas. Spatial variability is instead due to the change in pollution levels from one place to another. Within a city or town the spatial variability is mainly driven by local sources of pollution such as industry, residential heating and in particular road traffic with its major influence at a very local spatial scale. The spatial contrast between different locations within an area as usually considered to be rather stable over several years, as empirically shown in a few studies (Hoek et al., 2008a). In the long run, the spatial variability within an urban area is often larger than the temporal variability (Hoek et al., 2008a). The differences in long-term exposure between individuals are thus well characterized by spatial variability in the average pollutant concentrations. The exposure assessment methods that are described next are focused on characterizing spatial variability in air pollution, although some could be extended or modified to capture temporal variability.

Each of the surrogates described in the previous section was historically measured at specific locations, such as, near industries or at routine monitoring stations with the purpose of controlling levels to

comply with limits set to protect humans and the environment from adverse effects. Environmental epidemiology studies on the effects of air-pollution made use of such measurements to compare populations exposed to different levels. Ecological estimates constituted one of the early exposure assessment methods of environmental epidemiology (Morris, 1981).

Some pollutants -in particular primary pollutants such as NO and ultrafine particles- are unequally distributed across different areas within one town. Assignment of the same concentration, measured at a single monitor, to all people living within a city or community (Dockery et al., 1993), does not characterize differences in personal exposure to such pollutants (Carnow et al., 1969).

Thus, the development of personal exposure monitoring within the environmental monitoring field (it has already been extensively used in industrial hygiene) was recommended by the National Academy of Sciences (Study Group on Environmental Monitoring, 1977). Personal monitoring is based on measuring air pollution levels with devices that are worn or carried by the participants. It is still considered today the method providing the most accurate estimates of actual personal exposure to specific pollutants *per se*. However, it has several practical disadvantages that limit its use particularly in long-term and also in large studies: it is expensive, labor intensive, the measurement period is limited to a few hours or days, and carrying the equipment and completing the time-activity diary, that is usually required, lead to respondent burden and the difficulty of ensuring compliance. Moreover, from a policy perspective, one is interested in the effects and burden related to pollution from a specific source, e.g. traffic, rather than a single pollutant *per se* (e.g. NO_2). As long as there is no unique source-specific marker of pollution available, personal exposure measurements of a pollutant may not necessarily be more informative then some ambient monitoring or modeling approaches. E.g. a life-time measurement of personal exposure to NO_2 will not enhance our understanding of the effects of traffic-related pollution given that the total personal exposure to NO_2 is dependent on exposure to gas cooking, traffic, tobacco smoke and other sources of combustion. Thus, proper characterization of the ambient conditions, combined with information on the time people spend in these conditions remains a highly important approach to estimate personal exposure to outdoor air pollution.

Accordingly, refinements to the use of the pollution concentration at one or a few monitoring station started to get developed. Modeling to estimate pollution at unmeasured sites based on measurements at fewer sites was applied (Horie et al., 1976; Namekata et al., 1979; Sidorenko et al., 1972) and time activity patterns were incorporated to weight the pollution levels at different locations by the time spent at each location (Dockery et al., 1981; Moschandreas et al., 1979; Ott, 1982). Consequently, to characterize exposure at different locations, monitoring at indoor and outdoor residential environments, microenvironments and while commuting (Holland, 1983; Lambert et al., 1992; Ott et al., 1973; Ott, 1982) were implemented. Exposure assessment based on residential or personal monitoring alone is not efficient and even unrealistic for long-term studies. Although in the last decades the monitoring equipment have evolved and more accurate portable devices have become available nowadays, the biggest innovations in the exposure assessment field have been modeling techniques and software, which have been largely benefited from the development of GIS tools.

3.3 Modeling approaches for the assessment of exposure to local pollutants.

Economic and logistic constraints make it unfeasible to collect measurements from every participant (or his/her residence) of epidemiological studies. Thus several modeling techniques have been applied to estimate concentrations of pollutants based on information that is easier or less expensive to collect for all the study areas on a relevant scale (resolution). Models can be developed at every scale, from the very local (10's of meters) to the regional and even country wide areas. For pollutants with high concentration variability on the intra-urban scale, such as NO_x –including NO_2 and in particular NO– and UFP, methods providing enough resolution are needed. The methods more widely used at the local-scale will be described. More detailed evaluation and comparison of methods, is given by Jerrett et al. (2004), Briggs (2005) and the Health Effects Institute (HEI Panel on the Health Effects of Traffic-Related Air Pollution, 2010). The latter two references describe the modeling on the regional scale.

Proximity models estimate exposure as the concentration near by sources weighted by the distance from each source –e.g. distance to

traffic). For example distance to a road. Since emissions or concentrations levels at the roads are not available most of the time, the traffic intensity might be used as a surrogate. In cases where traffic intensity is neither available, the type of road (major road, highway, etc.) might be used instead. Distance cut-offs are also used to define exposure as a categorical variable; those living in the nearest buffer are more exposed than those living further away. Since these measures are based on source's location and are unspecific to any single pollutant, they are advantageous when exposure to the source, rather than to an individual pollutant, is of interest (Briggs, 2005).

Dispersion models are based on Gaussian dispersion equations, physicochemical laws and mass balance laws to estimate the concentration at a point for a specified time period given the background pollution levels, emissions, meteorological conditions and topography in the surrounds. They require thus an exhaustive quantity of input data that may not be available for many locations and are/or may be expensive. Emissions are usually specified as daily or annual means and thus, unmeasured temporal variability can introduce exposure measurement error. On the other hand, they incorporate spatial and temporal variability and can be applied at different spatial scales ranging from the regional (100–1000 km) to the micro (10–100 m) scale (Jerrett et al., 2004).

(Geostatistical) Interpolation models provide predictions estimated by interpolating the concentrations at the nearest locations where the concentration is known (monitors). They rely on geostatistical techniques, ranging from the most simple such as inverse distance weighting to more complex such as, splines (fit n-order equations as a function of coordinates x and y) (Wood, 2003), kriging, universal kriging. A major advantage of kriging is that they provide the standard error of every predicted value, an indicator of its reliability. Universal kriging takes into account a global trend in the spatial variability in addition to the local variability. These models require a relatively dense monitoring network and a level of expertise to develop the analyses (Jerrett et al., 2004) but are attractive tools to apply e.g. to existing health data to assign exposure to individuals Künzli et al. (2005a). They assume a homogeneous spatial variability determined by distance, and do not take into account that dispersion is also determined by topography and buildings. It is also possible to include covariates measuring such determinants. The method is called co-kriging.

Land use regression (LUR) models also called GIS-based pollution mapping, regression mapping or regression modeling are stochastic models that explain the spatial variability of air pollutants. In such equations, the measured concentration of a pollutant is the dependent variable and several predictors are tested as independent variables. Regression analysis is applied to find the set of independent variables that best describe the variability in the pollutant levels. The independent variables are potential determinants of the pollution levels, i.e. selected based on knowledge of physicochemical processes that govern pollutants formation and distribution. They could go from proximity variables, like distance to the nearest major road or distance to the nearest bus stop, to a description of the surroundings in terms of emission sources e.g. traffic intensity or area of ports, population density and built environment e.g. area of buildings or area of green spaces in a 100m radius. These variables are commonly derived using geographical information systems (GIS). LUR models do not require emission inventories and are a practical approach for the assessment of exposure to traffic-related pollution (Jerrett et al., 2004). Where sufficient measurements exists, LUR area-specific models can be developed locally offering an extremely high resolution (Briggs, 2005). This method also requires a dense monitoring network to obtain robust models (Basagaña et al., In press.). A detailed revision of LUR model performance and application is given by Hoek et al. (Hoek et al., 2008a).

Space-time models also called dynamic modeling integrate models measuring spatial distribution of pollution with temporal patterns of pollution and/or time-activity patterns of people. Examples of models with time varying patterns linked to residential history have been used in epidemiological studies by Oosterlee et al. (Oosterlee et al., 1996) based on dispersion models and by Diez Roux et al. (2008) based on trend, cyclic and autoregressive components to derive the time variability and thin-plate splines to characterize the spatial variability. Space-time models integrating time-activity patterns have been implemented by Jensen et al. (2001) and Gulliver and Briggs (2005), who derived a dispersion model for different microenvironments coupled with people's aggregated time-activity data and individual-level data from activity diary surveys, respectively. Künzli et al (1997) adopted similar approaches to characterize long-term exposure to ozone where a Californian ozone surface was combined with time-activity information. Effect

estimates of ozone on lung function in college students were however not sensitive to the inclusion or exclusion of personal time-activity information.

Hybrid models combine different modeling techniques to determine local, urban and regional contributions/components of air pollution, and may also couple them with personal monitoring. Hoek et al. (2002) and Beelen et al. (2007a) derived hybrid models for application in epidemiology. They made separate estimates of local traffic-related, urban background and regional background concentrations of elemental carbon, NO_2 and SO_2 using inverse-distance weighting and regression modeling. These models could also integrate time-activity patterns as in the so-called hybrid individual-exposure models (HEI Panel on the Health Effects of Traffic-Related Air Pollution, 2010).

There is nothing that can be considered a generalized best approach or one that serves all purposes. Choosing which model or measure of exposure to use in an epidemiological study depends on 1) defining what is the exposure time-period relevant for the health effect under investigation, e.g. long-term or short-term; 2) the relevant pollutant, or the marker of a mix of pollutants to be characterized, which depends on the source or sources investigated; 3) The scale relevant to the exposure distribution, which should be determined based on the previous points and the geographical characteristics of the study area (see discussion in Section 9) and 4) the availability of information and resources (time, money and expertise) to be invested in the exposure assessment.

Finally, it should be borne in mind that "*All models are wrong*" (Sterman, 2002), i.e. they are an estimate of a measure limited by what they were designed for. For example, routine monitoring is designed with regulatory purposes; models based on such measurements are thus not necessarily representative of the population's exposure. Models account for only part of the total exposures (e.g. estimates for known and inventoried sources, individual pollutants rather than mixes, pollutants mass rather than components, average exposure rather than integrated exposure (Ott, 1982)) and part of the total variability of such exposures (e.g. unmeasured temporal variability from scattered events such as dust episodes or spatial variability not taken as a continuum but averaging pollution in microenvironments). Models are tools to guide decisions, to inform decisions, but it should be the researcher who

makes those decisions taking into account the uncertainties inherent to the exposure model.

Sant Joan les Fonts Photo: Pere Crosas

4 RATIONALE

Since the 1990's, numerous studies investigated the short-term effects of air-pollution on cardiovascular health. However, the evidence on the effects of long-term exposure on cardiovascular chronic processes, pathologies or diseases in humans is based on far fewer studies. The association of exposure to air pollution with markers of the degree of atherosclerosis has been investigated in 13 studies (see **Table 2**). Most, but not all of them, have found positive associations of air pollution with atherosclerosis (see Section 2.1).

All except one study (Diez Roux et al., 2008) relied on 1-year estimates of residential pollution used as a proxy for long-term exposure. Besides, all studies relied on residential outdoor pollution levels derived from models based on measurements at a few monitoring locations. One study used a chemical dispersion model at a 1Km scale. Model-derived estimates of pollutant markers for longer term, e.g. 10-years, based on a dense network at the residential level would provide more precise exposure measurements, of added value to the assessment of their association with atherosclerosis.

The REGICOR population-based cohort study conducted in the Girona region is a unique resource to investigate the atherosclerosis-air pollution hypothesis in a Mediterranean region. The Girona region is particularly interesting given that the cardiovascular mortality rates are among the lowest in Europe while the established risk factors are among the highest. This is known as the French paradox (Masiá et al., 1998). This thesis is the first study on atherosclerosis and air pollution in the Mediterranean region.

Considering the widespread exposure to traffic-related pollutants and the large, and rising, global burden of CVD, measuring the risks associated with long-term exposure to air-pollution is of great relevance for public health.

Olot Photo: Pere Crosas

5 OBJECTIVES

General Objectives

I. Assess the long-term exposure to traffic-related air pollution in Mediterranean towns.

II. Evaluate the association of long-term exposure to traffic-related air pollution with subclinical atherosclerosis.

Specific Objectives

1. Develop spatial models to characterize the spatial distribution of NO_2 based on long-term measurements at participant's residences.

2. Characterize the long-term local distribution of heavy metals using mosses. Identify the main determinants of the spatial distribution of heavy metals in Girona, Spain.

3. Develop a traffic intensity road network database for the REGICOR-Air towns based on a short-term traffic count protocol.

4. Evaluate a monitoring strategy based on short-term UFP measurements in a highly dense monitoring network using mobile condensation particle counters to characterize the spatial variability of UFP. Test the performance of LUR models for ultrafine particles based on such short-term measurements in the region of Girona, Spain.

5. Investigate the effect of the number of measurement sites on the LUR model performance.

6. Investigate the association between long-term exposure to traffic-related air pollution and subclinical atherosclerosis in Spain.

Llagostera Photo:Z. Cochrane

6 HYPOTHESIS

This thesis developed and evaluated methods and tools for exposure assessment to ultimately investigate the following main hypothesis:

Long-term exposure to traffic-related air-pollution is associated with subclinical atherosclerosis.

Historical quarter and Onyar River, Girona city. Photo: Albert Marin (www.graph.cat)

7 METHODS

This section provides a brief summary integrating the methods used for the different papers included in this thesis. Further methodological details can be found in the results section (Section 8).

7.1 Study population and health assessment

The REGICOR study is a series of population-based cohorts of randomly selected adults living in Girona province. This study's main objective is to determine the prevalence, incidence and risk factors of the ischaemic heart disease in Spain (www.regicor.org). Within its framework, the REGICOR-Air project is an ongoing study that aims to determine the long-term effects of local traffic-related air pollution on cardiovascular health, including atherosclerosis. (Künzli et al., 2005a). This thesis is part of the REGICOR-Air project and is based on the follow-up of the cohorts originally enrolled in 1995, 2000 and 2005.

The follow-up of such cohorts was conducted in 2007-2010 and included a complete reassessment of the health status and collection of address history linked to time period during the last 10 years. It also included measurements of the IMT of the carotid artery by ultrasound (**Figures 4** and **5**), which constitute the main health outcome in this thesis. A detailed description of the methodology is presented in Paper IV (page172).

The study area of REGICOR-Air includes 12 towns of the province of Girona, Spain. These towns include settings like the average mid-size Mediterranean city of Girona and adjacent Salt, tourism getaways on the beach, industrial centers, maritime ports, historic quarters and small villages in the valleys and mountains. Most of these towns have a historical quarter used for residences and businesses with narrow streets and less than three-story buildings. The architecture in the downtown is also compact –high buildings and narrow streets– especially in the biggest towns. The façade of buildings and houses is separated from the road only by a narrow sidewalk and in a few cases also by a hallway or garden. **Figures 6** and **7** are pictures of different areas of these towns. Other pictures are provided at the beginning of each section.

A map of the study area is provided in **Appendix 2** (KLM files can be viewed in Google Earth).

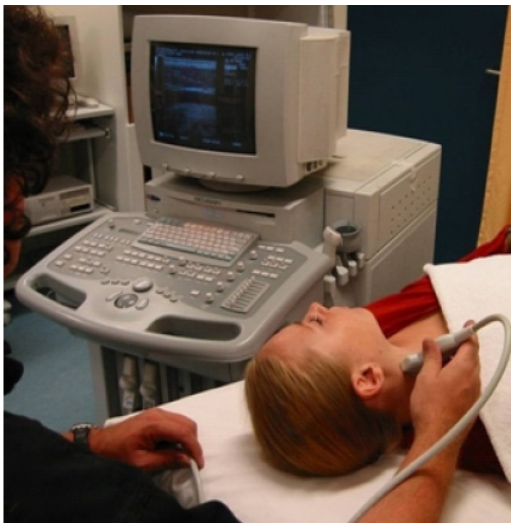


Figure 4. Measurement of intima media thickness of the aortic carotid artery by ultrasound.

Photo: Eric de Groot. Amsterdam Medical Centre

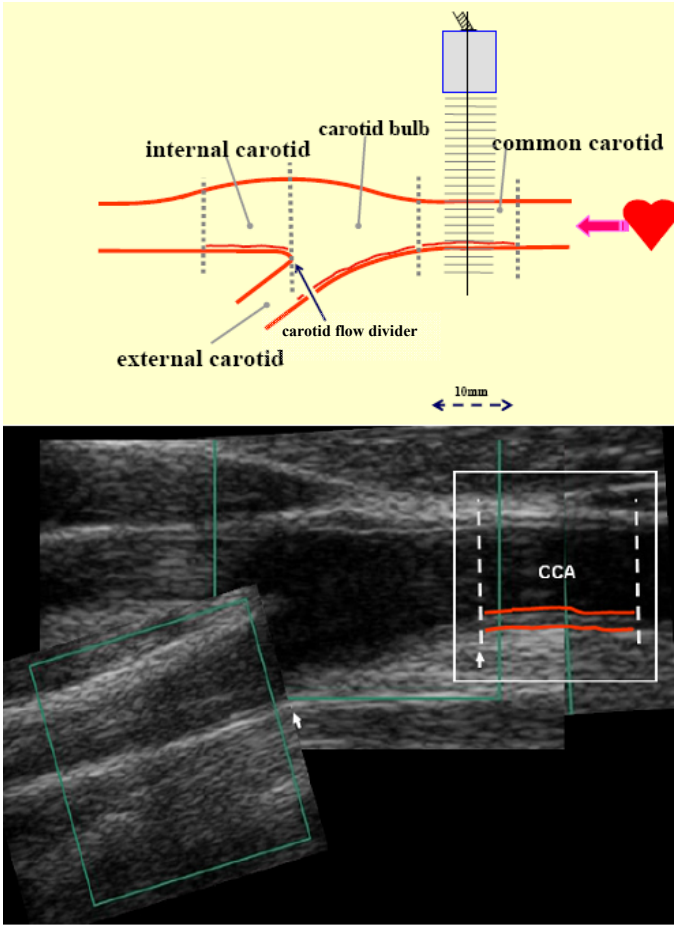


Figure 5. Resulting image and schematic drawing from intima media thickness measurement. CCA: common carotid artery.

Images: Eric de Groot. Amsterdam Medical Centre

Figure 6. *Downtown Girona city.* Photo: Jesus (http://www.dzoom.org.es)

Figure 7. *Residential area. Girona city.* Photo: Marcela Rivera

7.2 Exposure Assessment for the REGICOR-Air study

For the proper assessment of exposure to traffic-related local air pollutants, a good characterization of the spatial dispersion of air pollutants is required. Characterization of the distribution of air pollutants in the REGICOR-Air towns, with the exception of data on NO_2 from a report on air quality published by the local authorities (Targa, 2008), was lacking.

To achieve the first general objective we conducted monitoring campaigns based on different measurement techniques, subsequently used to build land-use regression models for several air pollutants.

NO_2 was measured using Palmes diffusion tubes in the balcony of 562 participants' homes, for one month between June of 2007 and July of 2009 (**Figure 8**). The methodology for NO_2 characterization is in Paper IV (page 178).

Heavy metals were measured using moss samplers (**Figure 9**) indoors and outdoors of 20 participant's residences and 3 outdoor locations during 2 months in 2008. In parallel, NO_2 was measured by diffusion tubes. In this thesis only the data of outdoor pollutants was used. The complete methodology for heavy metals measurements is written in Paper I (page 73).

Given that official data on **traffic intensity** was very limited or absent, and particularly unavailable for low intensity streets, we counted the number of vehicles driving in both directions for 15 minutes in 675 streets of the REGICOR-Air towns using a manual tally counter and a chronometer (**Figure 10**). Traffic counts were performed during non-rush hour period. Vehicles were distinguished as heavy or light vehicles. A similar traffic counting protocol was used before by Van Roosbroeck et al. (2007). As they proposed, we scaled the short-term counts to 24-hrs to estimate daily traffic intensity. The traffic intensity data was inputted to a road network in GIS format and the data were extrapolated to other segments of the same street until the street intersected with a medium or major street (see roadnetwork in **Figure 12,** page 207). The complete methodology for traffic intensity measurements is written in Paper II (page 107).

Simultaneously to the traffic count, **UFP** measurements were collected. UFP were measured for 15 minutes on the sidewalk of 644 participants' homes in 2009 using mobile condensation particle counters (P-Trak) (**Figure 10**). The complete methodology for UFP measurements is written in Paper II (page 106).

Because of the differences in urban structure in the Mediterranean cities compared to cities of northern Europe and the United States and given the not well known impact of narrow street canyons on near-road pollution, characteristic of some Mediterranean cities, we performed the monitoring of all pollutant components in high number of sites per unit of area i.e. a dense monitoring network (**Figure 11**). All these measurements and other measurements of potential predictors made out in the field, plus geographical information was used to study the distribution of pollutants across the towns, determine their main predictors and build models to estimate the pollutants concentrations at the participant's residences.

To evaluate the performance of the LUR model and the effect of the number of measurement sites on such performance, we conducted a methodological study using the NO_2 measurements. The complete methodology of this study is written in Paper III (page 140).

Finally, to achieve the second general objective we measured the participants' long-term exposure to air pollution using NO_2 and traffic proximity (intensity) markers and studied its association to subclinical atherosclerosis. The complete methodology to estimate long-term exposure and to assess its relationship with atherosclerosis is written in Paper IV (page 174). The other markers of exposure – UFP and heavy metals- will be studied in future analyses of the REGICOR-Air study (see Section 9.4.2).

Figure 8. Picture of the NO_2 passive sampler set outdoor of participant's homes

Figure 9. Picture of the moss monitor set outdoors of participants' homes.

Figure 10. Picture of the mobile condensation particle counter (P-Trak) and tally counter and setting used to monitor outdoors of participants' homes.

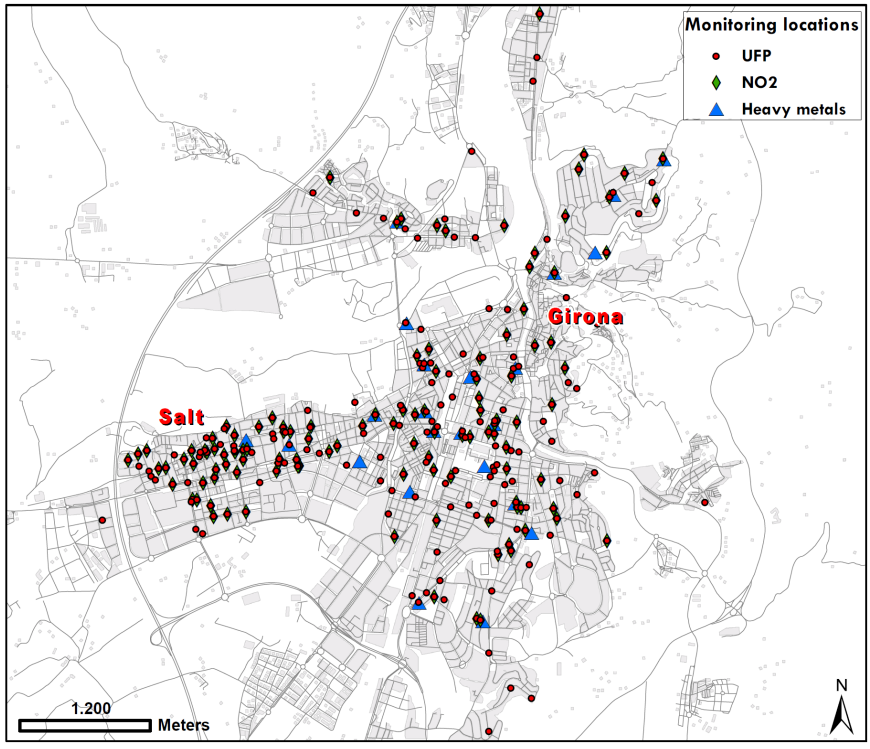


Figure 11. Locations of UFP, NO_2 and heavy metals measurements in the cities of Girona and Salt.

8 RESULTS

8.1 Monitoring of heavy metal concentrations in home outdoor air using moss bags (Paper I)

Rivera, M., Zechmeister, H., Medina-Ramón, M., Basagaña, X., Foraster, M., Bouso, L., Moreno, T., Solanas, P., Ramos, R., Köllensperger, G., Deltell, A., Vizcaya, D., Künzli, N., 2010. Environmental Pollution 159(4), 954-962.*

* *This paper is reproduced according to the original print version. References of this paper are included in Section 10.*

Supplementary Data

Chemical Analysis

The concentrations of Al, As, Cd, Cr, Cu, Mo, Pb, Sb, Sn and Zn were measured by inductively coupled plasma sector field mass spectrometry (ICP-SFMS) carried out on an Element 2 ICP-SFMS (ThermoFisher, Bremen, Germany). When the monitoring was finished, the mosses were shipped to the University of Natural Resources and Applied Life Sciences, Department of Chemistry and the Umweltbundesamt GmbH laboratories in Vienna (Austria), where the chemical analysis was performed.

Moss samples were dried at 30°C (for 72 hours) and grinded under liquid nitrogen in porcelain vessels, and then submitted to microwave digestion with H2O2/HNO3. Sample aliquots of 200 mg plus 4mL of ultra-pure nitric acid and 1 mL 30% H2O2 solution (ultra-pure, Merck) were filled into acid-steam-cleaned PFA (perfluoroalkoxy polymer)-microwave digestion vessels and closed with pre-cleaned PTFE (polytetrafluoroethylene) caps. The digestion was performed applying a microwave program employing maximum microwave power of 450 W. After cooling, the digested samples were transferred to PP (polypropylene) vials and filled up to 10 mL with ultra-pure water. Microwave digestion blanks were prepared using ultra-pure water. For quantification multi-element standards were prepared from Merck ICP-Single element standard in PFA bottles and vials through dilution in ultra-pure water and addition of 1% nitric acid. The acid content was matched to the content in the respective samples. Prior to ICP-SFMS measurement, Indium was added to all samples as an internal standard at a final concentration of 1 µg/L. The certified reference material (CRM) TM 27.2 (low level fortified sample for trace elements) was used for calibration quality control. The certified values and the obtained values agreed within measurement uncertainty for the tested CRM.

Elemental analysis was carried out on an Element 2 ICP-SFMS (ThermoFisher, Bremen, Germany). As sample introduction system a PFA micro-flow nebulizer (Elemental Scientific Inc., Cuming, Omaha, USA) with an internal diameter of 45 µm was used during the study. The self aspirating PFA micro-flow nebulizer (ESI) at a flow of 100 µL/min was combined with a PFA double pass Scott-type spray chamber, a sapphire injector pipe, a quartz torch and platinum

sampler and skimmer cones (all parts Thermo Fisher). The following ICP-SFMS operating conditions were applied in this study: RF power of 1300 W and plasma gas flow of 16 L/min. Sample gas and auxiliary gas flows were set to 1.06 L/min and 0.86 L/min, respectively.

Selected isotopes for interference- free ICP-SFMS measurement were ^{195}Pt and ^{208}Pb at low resolution (LR), ^{27}Al, ^{52}Cr, ^{65}Cu, ^{66}Zn, ^{95}Mo, ^{111}Cd, ^{118}Sn, ^{121}Sb, at medium resolution (MR) and ^{75}As at high resolution. During all measurements ^{115}Indium was used as internal standard at all resolutions. Nominal mass resolutions of the Element 2 ICP-SFMS for low resolution (LR), medium resolution (MR) and high resolution (HR) are 350, 4 500 and 10 000, respectively.

Table 1. Spearman correlation coefficients across pollutants. (N=21 locations)[a].

	Al	As	Sb	Cd	Cr	Cu	Pb	Mo	Sn	Zn	NO_2
Al	1.0000										
As	0.4039	1.0000									
Sb	0.6974*	0.3195	1.0000								
Cd	0.0299	-0.1299	0.2247	1.0000							
Cr	0.9195*	0.3623	0.8351*	0.1442	1.0000						
Cu	0.8597*	0.4000	0.7273*	-0.0429	0.8662*	1.0000					
Pb	0.5429*	0.1416	0.5156*	0.5013*	0.5870*	0.4623*	1.0000				
Mo	0.4662*	0.0831	0.7909*	0.3727	0.6416*	0.5273*	0.6143*	1.0000			
Sn	0.7377*	0.2779	0.8753*	0.2636	0.8701*	0.7636*	0.6416*	0.8545*	1.0000		
Zn	0.5260*	0.2143	0.6026*	0.5221*	0.5143*	0.5506*	0.7143*	0.5455*	0.5221*	1.0000	
NO_2	0.1870	0.2403	0.6143*	0.1481	0.3104	0.3221	0.2519	0.4104	0.3740	0.4532*	1.0000

[a] Two outdoor samplers located in streets are not included

* Significant at 0.05 (2-tailed);

[a] Two outdoor samplers located in streets are not included
Significant at 0.05 (2-tailed);

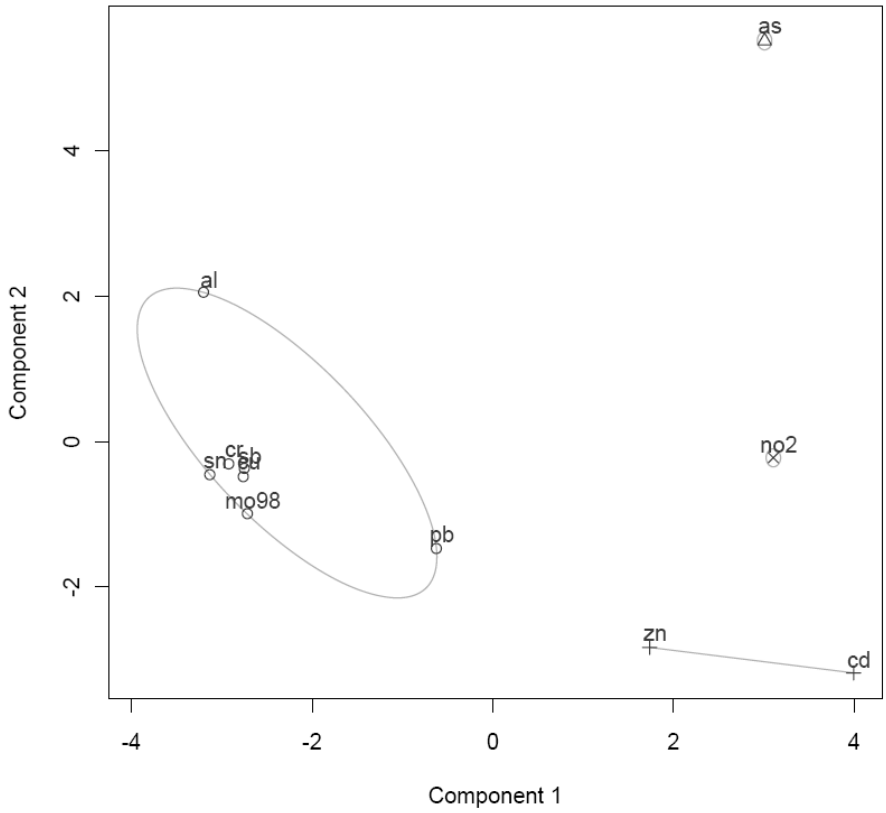


Figure 1. Cluster analysis (Partitioning Around Medoids method) for metals and NO_2; Numbers of clusters k=4. Components are derived from principal component analysis (data not shown). N=23

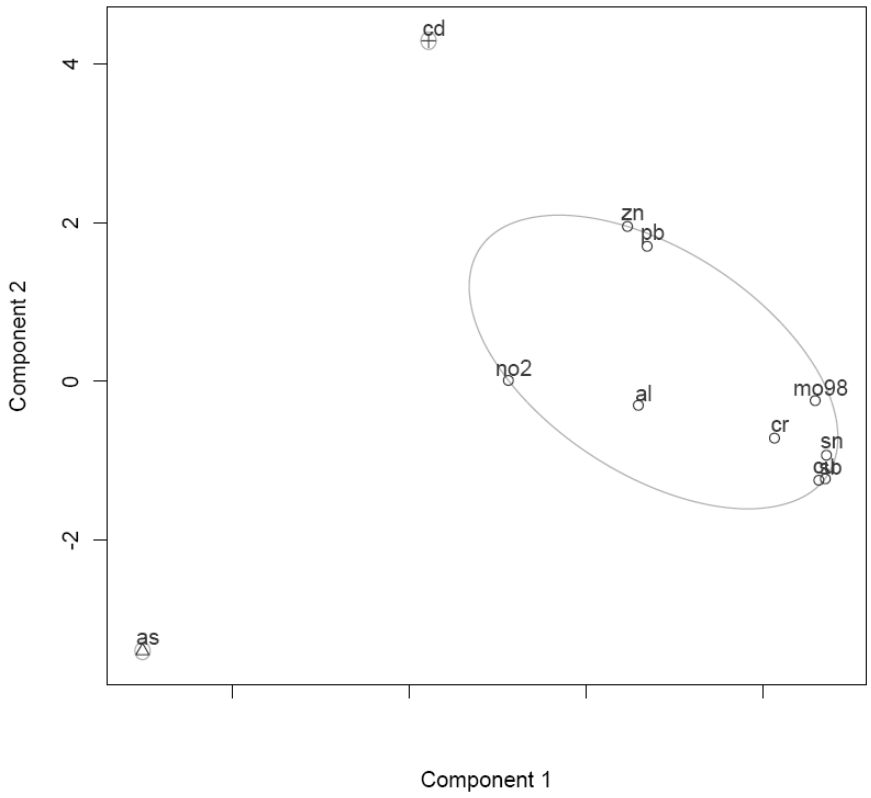


Figure 2. Cluster analysis (Partitioning Around Medoids method) for metals and NO_2; Numbers of clusters k=3. N=21 Two outdoor samplers located in non-residential sites are not included.

8.2 Spatial distribution of ultrafine particles in urban settings: a land use regression model (Paper II)

Rivera, M., Basagaña, X., Aguilera A., Agis D., Bouso, L. , Foraster, M., Medina-Ramón, M., Pey, J., Künzli, N., Hoek, G., In press. Atmospheric Environment.
Submitted September 21, 2011, revised version submitted December 8, 2011, accepted January 20, 2012.*

* *This paper is reproduced according to the revised version submitted. References of this paper are included in Section 10.*

SPATIAL DISTRIBUTION OF ULTRAFINE PARTICLES IN URBAN SETTINGS: A LAND USE REGRESSION MODEL

Authors:

Marcela Rivera[a,b,c], Xavier Basagaña[a,b], Inmaculada Aguilera[a,b,d], David Agis[a,b], Laura Bouso[a,b], Maria Foraster[a,b,c,d], Mercedes Medina-Ramón[a], Jorge Pey[e], Nino Künzli[f,g], Gerard Hoek[h]

Authors' affiliations:

[a] Centre for Research in Environmental Epidemiology CREAL, Barcelona, Spain
[b] Hospital del Mar Research Institute (IMIM), Barcelona, Spain
[c] Universitat Pompeu Fabra, Barcelona, Spain
[d] CIBER Epidemiología y Salud Pública (CIBERESP), Spain
[e] Institute of Environmental Assessment and Water Research (IDÆA-CSIC), Barcelona, Spain
[f] Swiss Tropical and Public Health Institute, Basel, Switzerland
[g] University of Basel, Basel, Switzerland
[h] Institute for Risk Assessment Sciences (IRAS), Division of Environmental Epidemiology, Utrecht University, Utrecht, The Netherlands

Corresponding Author:

Marcela Rivera, CREAL - Centre for Research in Environmental Epidemiology. Barcelona Biomedical Research Park. Dr. Aiguader, 88. 08003 Barcelona. Spain. Tel. +34 932 147 342. Fax + 34 93 3 160 575.
arivera@creal.cat, marcelariverav@gmail.com

Abstract

Background: The toxic effects of ultrafine particles (UFP) are a public health concern. However, epidemiological studies on the long term effects of UFP are limited due to lacking exposure models. Given the high spatial variation of UFP, the assignment of exposure levels in epidemiological studies requires a fine spatial scale. The aim of this study was to assess the performance of a short-term measurement protocol used at a large number of locations to derive a land use regression (LUR) model of the spatial variation of UFP in Girona, Spain.

Methods: We measured UFP for 15 minutes on the sidewalk of 644 participants' homes in 12 towns of Girona province (Spain). The measurements were done during non-rush traffic hours 9:15-12:45 and 15:15-16:45 during 32 days between June 15 and July 31, 2009. In parallel, we counted the number of vehicles driving in both directions. Measurements were repeated on a different day for a subset of 25 sites in Girona city. Potential predictor variables such as building density, distance to bus lines and land cover were derived using geographic information systems. We adjusted for temporal variation using daily mean NO_x concentrations at a central monitor. Land use regression models for the entire area (Core model) and for individual towns were derived using a supervised forward selection algorithm.

Results: The best predictors of UFP were traffic intensity, distance to nearest major crossroad, area of high density residential land and household density. The LUR Core model explained 36% of UFP total variation. Adding sampling date and hour of the day to the Core model increased the R^2 to 51% without changing the regression slopes. Local models included predictor variables similar to those in the Core model, but performed better with an R^2 of 50% in Girona city. Independent LUR models for the first and second measurements at the subset of sites with repetitions had R^2's of about 47%. When the mean of the two measurements was used R^2 improved to 72%.

Conclusions: LUR models for UFP were developed, based on a highly cost-effective short-term monitoring campaign at a large number of sites, with fair performance. Complementing the approach with further strategies to address sources of temporal variation of UFP is likely to result in improved models as indicated by the good performance of the model based on the subset of sites with one

repeated measurement. Our approach is promising for UFP and possibly for other PM components requiring active sampling.

Keywords: Particle number concentration; PM0.1; LUR; exposure assessment; monitoring; outdoor home; traffic; Girona; Spain.

Background

There is increasing evidence that ultrafine particles (UFP) are of public health concern (Ibald-Mulli et al., 2002; Peters et al., 2011). UFP, also known as particulate matter $PM_{0.1}$, are airborne particulates with aerodynamic diameter smaller than 100 nm (0.1 µm). Some studies comparing the effects of UFP and $PM_{2.5}$ have reported stronger associations between respiratory outcomes and UFP in humans (Peters et al., 1997; Politis M, 2008). Studies on animals provide evidence of stronger cardiovascular effects and promotion of early atherosclerosis associated to UFP (Araujo et al., 2008). The epidemiological evidence is still insufficient, with only a few studies reporting on the short-term associations between mortality and UFP. According to a recent expert elicitation, daily all-cause mortality was reported to increase between 0.1 and 0.4% per 1000 particle/cm^3 increase in ambient UFP (Hoek et al., 2009). There are no studies on the effects of long-term exposure to UFP. However, the burden due to chronic long-term effects is usually much larger than that due to acute effects (Brook et al., 2010; Künzli et al., 2010), thus, studies on the long-term effects of UFP are necessary (Hoek et al., 2009; World Health Organization, 2006). Relying on proximity to busy roads as a marker for exposure to traffic-related pollutants, such as UFP, has several limitations, and transferring such results into policy is difficult (HEI Panel on the Health Effects of Traffic-Related Air Pollution, 2010). Thus, information on the spatial distribution of long-term mean UFP levels is needed.

The use of one or a few fixed-site monitors to represent the average concentration in a city or large urban area is usual for some pollutants as $PM_{2.5}$ (Pope et al., 2002a), but is not appropriate when studying UFP. Ultrafine particles are emitted by combustion processes, and tail-pipe exhaust from motor vehicles. Based on one-year measurements of UFP (2003-2004) in Barcelona, Pey et. al. found that between 52 to 86% of the particles with size range 20-800 nm was emitted by vehicles (2009). Because UFP are a primary pollutant that is rapidly transformed by physicochemical processes (dispersion, coagulation, deposition, etc.), and emitted mainly by mobile sources, they show a very high spatial variation. In open spaces, pollutant concentrations can drop to background levels at distances of 100-300 m from important sources such as highways (Zhou, 2007; Hagler, 2009). In the compact urban structures characteristic of European cities, spatial extent is even smaller (10's

of meters) because traffic intensity inside the city is typically lower than on highways and concentrations fall to background levels behind a row of uninterrupted buildings (Bloemen et al., 1993). Additionally, given the small distances from home to road and the narrow streets typical of these towns, even modest traffic intensities of a few thousands vehicles per day determine outdoor and personal exposure concentrations. Therefore, to properly characterize the UFP spatial distribution, estimates at a fine spatial scale are needed.

To assess exposure at a fine spatial scale, both deterministic and stochastic models have been developed. A thorough review of intraurban models is presented by Jerret et al. (2004). In general, the models that give the most reliable estimations of air pollution are dispersion models, integrated emission-meteorological models and land use regression models (LUR). The first two methods require an exhaustive quantity of input data such as the configuration of the setting, wind speed and direction, emission inventory and emission factors estimates. These data are not always available, particularly for small cities and low income countries. In the specific case of UFP, availability of emission data is more limited than for regulated pollutants such as NO_2 or PM_{10}. In contrast, LUR models use as input data potential determinants of pollution levels that are commonly available or can be derived through geographic information systems (GIS). Such variables typically include traffic counts, proximity variables (e.g. to roads) and street configuration variables.

To our knowledge, only one LUR model of UFP has been published. This model, developed for the city of Amsterdam (The Netherlands), captured 67% of the UFP variability using long-term measurements in 50 locations (Hoek et al., 2010). The model was based upon a 17-month monitoring period with weekly measurements at each location performed in the framework of a study aimed at evaluating temporal correlations across urban areas. Such an extensive campaign is not feasible when the aim is to develop a LUR model. Typical LUR models are based upon monitoring campaigns at 40-100 sites with one to four weeks of sampling (Brauer et al., 2003). Additionally, the use of integrated samplers left unattended at the sites for 1-2 weeks, as is common for PM campaigns, is not possible in UFP monitoring because of the cost of equipment (typically 6-10 instruments used in these studies) and the need to supervise it. More affordable mobile condensation particle counters allow monitoring

for short time periods and easier access to locations were the conventional instruments can not be deployed.

The objective of this study was thus to evaluate a monitoring strategy based on short UFP measurements in a highly dense monitoring network using mobile condensation particle counters. We tested the performance of LUR models for ultrafine particles based on such short-term measurements in the region of Girona, Spain.

Methods

Study design

Ultrafine particles were monitored outside the residences of participants of the REGICOR study (http://www.regicor.org/). REGICOR is a series of population-based cohorts of adults, randomly selected in Girona province in 1995 and 2000 (Masiá et al., 1998). Girona is the northernmost province of Spain. Its towns include settings like the average mid-size Mediterranean city of Girona and adjacent Salt, tourism getaways on the beach, industrial centers, maritime ports, historic quarters and small villages in the valleys and mountains. Most of these towns have a historical quarter used for residences and businesses with narrow streets and less than three-story buildings. The architecture in the downtown is also compact –relatively high buildings (usually 5-8 stories) and narrow streets– especially in the larger towns. The façade of buildings and houses is separated from the road only by a narrow sidewalk or, in a few cases, by a sidewalk and a hallway or garden. In the study sites the median distance from the front door to the curb was 2.5 m.

This study was designed to share efforts and resources with an assessment of traffic density. The most important local determinant of UFP is traffic. In the absence of reliable traffic data on local streets, we adapted a traffic counting protocol used before in the Netherlands (Van Roosbroeck et al., 2007), expanded by parallel measurements of UFP. Sites were selected to cover a broad range of traffic-related pollution. In the absence of UFP data, the selection was based on previous measurements of NO_2 assuming that this strategy would as well maximize the distribution of UFP. Additionally, sites covered a broad range of urban settings, such as low and high building-density areas, narrow and broad streets,

downtowns, historical quarters and peripheral residential areas. The sites selected in each town were broadly representative also of the residential locations of the cohort participants and were well distributed across the town. The number of sites and measurements was limited by the seven-week period in which logistics - including equipment - were available: 32 working days between June 15 and July 31, 2009.

UFP and traffic monitoring

We measured UFP for 15 minutes on the sidewalk of 644 participants' homes in 12 towns of Girona province (Spain) (**Table 1, Figure 1**) with the highest number of sites selected in the city of Girona (N=167). We performed repeated measurements in different days for 25 sites in Girona. The UFP were monitored with P-Trak counters (TSI model 8525) located perpendicularly to the street, 1.5 m above ground on the sidewalk and, whenever possible, at 1.5 m from the façade. Five P-Traks were used. The P-Trak counts the total number of particles in the size range of 20-1000 nm. Particles in this range both exclude the 10-20 nm and extend beyond the 100 nm limit of the UFP but represent well the ultrafine particles as has been found after comparing P-Trak measurements with Scanning Mobility Particle Sizer coupled to a Condensation Particle Counter measurements (Pearson correlation coefficient for the size range 10-70 nm ranged between 0.7-0.9) (Hagler et al., 2009). In urban areas, UFP typically account for about 80% of the total particle number count. For ease of terminology, we will refer to the P-Trak measurements as UFP. We took measurements for 6-hour periods before and after the monitoring period to evaluate the agreement of the P-Traks. Intra-class correlation between the five P-Trak measurements was 0.989. The maximum relative difference was on average 15.6.

In parallel, we counted the number of light, heavy and motorcycle vehicles driving in both directions. Traffic was counted manually using tally counters. Since 15-minute traffic measurements in the non-rush hour period have been shown to represent well the long-term traffic mean (Van Roosbroeck et al., 2007) and the within-day pattern of UFP concentrations closely resembles that of traffic load (Diapouli et al., 2007; Wang et al., 2010), we conducted all measurements in non-rush hour periods during daytime. Based on continuous 24-hour traffic measurements at 22 locations, available

from the city council Mobility plan of 2007 (www.ajgirona.cat/mobilitat/plamobilitat.html), we defined the non-rush hour periods as 9:15 to 12:45 and 15:15 to 16:45. Spearman correlation coefficient between the 15-minute traffic from our counts and the 24-hour traffic at these locations was 0.71 (p-value 0.0003). These measurements were also used to scale the 15-minute traffic averages to 24-hour averages. The scaling ratio was 1/69.48, very similar to that reported for The Netherlands, 1/61.92 (Van Roosbroeck et al., 2007).

Other predictor variables

Addresses of participants' homes were geocoded at the front door level by a private company (Arvato Services). Addresses geocoded with less than maximum precision (i.e. exact address) were again geocoded by entering the address to the web mapping application of the Cartographic Institute of Catalonia (http://mercuri.icc.cat) and capturing the geocode at the building's door. We compared the geocodes with maximum precision level provided by Arvato with those obtained from the mapping application for a subsample of 64 sites. The median distance was 13.6 meters.

We assigned the 24 hour scaled total traffic and the fractions of heavy duty traffic and motorcycles vehicles to the central road network used within ESCAPE (www.escapeproject.eu). We defined "major road" as having a traffic intensity higher than 7000 veh./day. Data on land-cover and geographic characteristics, e.g. area of residential land, distance to the sea and to continental water, altitude, were taken from the CORINE database (European Environmental Agency). Data on land used for industry was obtained from the Cartographic Institute of Catalonia. Meteorological data, at four stations in the study area, were available from the Catalonian Meteorology Service. Population and household density were available from the 2001 Spanish census (INE, 2004). Routes of bus lines and location of bus stops were provided by the city councils of Girona, Salt and Olot. Predictor variables referring to road length or to a given area around the monitoring location were derived for buffers of 25, 50, 100, 150, 300, 500 and 1000m. Detailed distribution of these variables is given in **Supplementary Table S1**.

Statistical Analysis

As sites were measured on different days and times, temporal variation may affect the modeling of spatial variability. In the Amsterdam study, UFP data at a central urban-background site were used for adjustment (Hoek et al., 2010). In the Girona area there was no fixed site with UFP data available. To account for the temporal variation, we calculated the ratio of the mean level of NO_x in the 7-week measurement period to the NO_x in the day of each measurement at an urban background fixed station in Girona city. Then, each measurement was multiplied by this ratio. This approach was built on the assumption that the relationship of the 15-minute UFP measurement with the 7-week average UFP at any location within the study area is equal to the relationship of the daily NO_2 to the 7-week average NO_x at the fixed monitoring station. Other approaches like the adjustment using hourly measurements of NO_2, hourly and daily measurements of NO_x, and the UFP measurements trend over time were also explored.

The models were derived by supervised forward linear regression following the methodology used by Hoek et al. (2010). Log-transformed values of temporally adjusted UFP minus a factor of 659.2 were used, because in exploratory analyses the residuals of the models did not follow a normal distribution. The factor of 659.2 is a small quantity compared to the monitored concentrations and was subtracted so that the skewness equals zero. The forward selection process started with an empty model. In every iteration, all the potential predictors were entered independently. The predictor variable producing the highest increase in the adjusted R^2 was retained provided the adjusted R^2 increase was higher than 1% and the direction of the association with UFP was as expected (i.e. for traffic or road length a positive association was conditioned, while for distance to traffic or area of nature a negative association was conditioned). We repeated this process until no additional variables could increase the adjusted R^2 more than 1%. Covariates with p-value higher than 0.1 were sequentially removed from the model.

We developed models for the full area and for separate groups of towns. Because of the heterogeneity in the study area we combined adjacent towns with similar urban characteristics and traffic intensity levels, resulting in seven geographic groups. For those locations where we had repeated measurements, we additionally fitted independent models for the first and second measurement as well as

for the average across the two repeated measurements. We investigated the influence of the short-term sampling variability by comparing the performance of the model for the average of two measurements with the models for a single measurement. To evaluate the suitability of the temporal adjustment approach, we added sampling date and hour to the core model. Significant effects would indicate residual temporal variation in the adjusted UFP levels.

To assess the predictive ability of the model to a new dataset, cross-validation R^2 was computed by the leave-one-out method (see detailed explanation in Supplementary Material).

The analyses were done using Stata 10.1 (StataCorp, College Station, TX) and ArcGIS 9.2 (ESRI, Redlands, CA).

Results

Descriptive statistics of the UFP for the study area and by groups of towns are presented in **Figure 2**. Median of UFP concentrations was 8.313 particles per cm^3 ranging from 1.840 to 53.103. The highest concentrations of UFP were found in Girona, Blanes and Salt the larger cities in the study area. The lowest were in Sta.Cristina/Llagostera, two small towns on the coast. UFP concentrations adjusted for temporal variations were very similar to unadjusted concentrations.

The weather in the monitoring days was relatively constant, with mean daily temperatures ranging from 19 to 27°C except for one day with temperature of 29.8°C in Girona city and 28.2C in the coast (**Figure 3**). Average daily temperature in the closest meteorological station was inversely associated with the number of ultrafine particles. Wind velocity on measurement days ranged from 0.6m/s in Girona to 3.2 in Banyoles/Porqueres and 3.5m/s in the coast. Relative humidity varied between 43 and 81%. A few showers occurred on 14 monitoring days. Only two of these days had rainfall of more than 2.5mm (equivalent to l/m^2). UFP monitoring was avoided during the rain but in nine measurements it started raining after the measurements were initiated. **Figure 3** illustrates temporal correlation in weather patterns across the areas (particularly temperature), but also illustrates differences in absolute values of especially wind speed and relative humidity on the same

day, supporting the decision to evaluate local LUR models in addition to a study-area overall model.

The best predictors of the UFP distribution were the 24h traffic intensity, high density residential land in a 1.000 m buffer, distance to the nearest intersection of two major crossroads and household density in a 100 m buffer (**Table 2**). The median traffic intensity was 695 vehicles per day and it ranged between 69 and 35.781. Especially in the smaller towns, the number of sites with moderate or high traffic counts was few. Traffic distribution by towns is shown in **Figure 4**. The median area of high density residential land in a 1.000 m buffer was 1.16km^2. The median distance to the nearest intersection of two major roads was 299 m. And the median number of households in a 100 m buffer was 42. This is equivalent to 1337 households per km^2.

The LUR model for the entire study area is shown in **Table 2** (core model). It had a moderate performance explaining 36% of UFP total variation and the cross-validation R^2 was 0.35 (**Supplementary Figure S1**). The same model using the untransformed UFP, instead of its logarithm, showed very similar results with an R^2 of 0.33 and a cross-validation R^2 of 0.31

Because differences in the UFP dispersion may exist between towns we added town to the core model. The explained variability increased to 40%. Three towns significantly accounted for the UFP variability: Blanes, St.Joan/Olot and Porqueres/Banyoles. Since the predictors' effects could be different by town, we included interaction terms in the core model. The interactions of town and high density residential land and distance to the intersection of two major roads were statistically significant and the model explained 42% of the UFP variability (**Supplementary Table S2**).

Since part of the UFP variability was captured by the class variable town, we evaluated both the performance of the core model in each town and the performance of local models. Applying the core model (**Table 2**) to each group of towns showed a good performance in Sta.Cristina/Llagostera, similar performance than the core model in Girona, Blanes, and Palamos/La Bisbal/Palafrugell and poor performance in Salt, St.Joan/Olot, and Porqueres/Banyoles. The goodness of fit and cross-validation R^2 are shown in **Table 3**. The local models included predictor variables similar to those of the core model: Traffic, distance to traffic and proxies of building and

household density. Additionally, length of major roads in 50 m and 100 m buffers, length of bus lines, North-South orientation of the nearest major street and area of natural and industrial land were also selected as best predictors for the UFP distribution in the town-specific models (**Table 3**). Most predictor variables in these models had small spatial scales in the range of 100 m, and the areas with wider extension sampled (Sta.Cristina/Llagostera and Palamos/La Bisbal/ Palafrugell) also had variables with spatial ranges between 500 and 1000m. Overall, local models performed better – sometimes substantially better– than the model for the entire study area, except for the town Porqueres/Banyoles where performance remained poor. We investigated whether specific weather circumstances explained this low performance. Excluding measurements on days with mean wind velocity above 3 m/s did not improve the model.

When we added sampling date and hour of the day as categorical variables to the core model, they resulted as significant predictors of the UFP levels and the adjusted R^2 increased to 51% (**Supplementary Table S3**). Similar improvements were observed when adding these variables to the core models adjusted for town and town interactions (data not shown). The regression slopes of the spatial predictors were, however, almost identical to the slopes without temporal predictors (**Supplementary Table S3** and **Table 2**).

Repeated measurements

The Pearson correlation coefficient and intra-class correlation for repeated measurements of UFP levels were 0.24 and 0.32 respectively (**Figure 5**). This low correlation was not due to variability in total traffic levels during the two 15-minute periods. Correlation coefficient for the total traffic counts (heavy, light and motorcycles) was 0.95 (p-value <0.0001). Correlation coefficient for heavy vehicles in the first and second measurements was 0.59 (0.0018).

The LUR models for the first and second measurements on these 25 locations had R^2's of 0.47, but strongly improved to 0.72 when the mean of the two measurements was used (**Table 4**). The main predictors in these models were length of roads in buffers smaller than 150m, length of and distance to bus lines and characteristics of the urban space as building density in a 100 m buffer and the area urban green land in a 500 m buffer. Nevertheless, traffic intensity, a

very common variable in the core model and the town specific models, was only selected into the model for the second measurement. The main determinants of UFP in the subset of 25 locations with repeated measurements were very similar to those for the full set for Girona city (**Table 3**).

Costs

Average fieldwork time for one visit was 50 minutes. This time includes commuting time (average in our study was approximately 15 minutes) from one location to the next, completing datasheets, downloading data from the P-Trak, and emailing collected data to the study center (Barcelona) every two days.

Discussion

We derived LUR models that captured between 36 and 42% of the spatial variability of ultrafine particles in the study area. Traffic intensity, distance to nearest major crossroad, area of high density residential land and household density were the main UFP predictors.

The predictors of UFP in Girona are similar to those reported for Amsterdam, namely product of traffic and inverse distance to traffic squared, address density in 300 m buffer and area of port in 3000 m buffer (Hoek et al., 2010). The performance of the Amsterdam model was better, explaining 65% of variability. The performance of our model is however similar to a variation of the Amsterdam model (R^2=44%) in which the variables collected on the field, specifically, distance from sampling site to the road, were removed and the potential predictors were limited to GIS derived variables, as in our study. The lower R^2 of the GIS model compared to the GIS plus field observations model in Amsterdam was attributed to limited geographical precision of geocoding and GIS predictors (Hoek al. 2010). Our study area is more geographically complex whereas the Amsterdam study was based on inner city locations only. The model for Girona city, the biggest city in our area, explained 50% of the UFP variability. Finally, UFP data for Amsterdam were available for a longer time period (one week) than UFP data for Girona (15 minutes), but at fewer sites.

Model performance

Several reasons might have led to the moderate predictive ability of our models, specifically: the temporal variability in the UFP levels not captured by the temporal adjustment, the heavy traffic variability not properly captured by 15-minute traffic counts, the moderate UFP levels and the limited variance in the distribution of some covariates and its related measurement error.

Temporal variability

Despite monitoring in the non-rush hours in a consecutive 7-week monitoring period in one season and adjusting for temporal variation using NO_x at an urban background site, including sampling date and hour to the LUR models improved the R^2 of the core model from 36 to 51% (**Supplementary Table S3**). This illustrates that temporal variation remained in our measurements despite the adjustment. Some of the variability that the models were unable to capture is thus explained by remaining temporal variability. The variation of the UFP levels with the monitoring hour was consistent with the pattern described in several studies (Hoek et al., 2008b; Pey et al., 2009; Wang et al., 2010). After a morning peak simultaneous to that of traffic, the UFP levels steadily decrease, largely because of the increase in atmospheric turbulence and atmospheric instability that develops after sunshine. Other methods to adjust for temporal variability in the UFP levels were tested (NO_2 and NO_x hourly levels, NO_2 daily levels, splines of the temporal patterns in our measurements) as a sensitivity analysis. None of them increased the R^2 of the LUR models. Adjustment was likely insufficient because we did not have UFP measurements available from a monitoring station and UFP has only a moderately high correlation with NO_2 and NO_x (Paatero et al., 2005). Furthermore, the NO_x concentrations in Girona may not adequately capture the temporal variation in the other towns. An important observation was that the coefficients of the covariates of the spatial part of the LUR models did not change compared to the core model without the temporal part. Hence these coefficients can be used to model the spatial variation of UFP.

The repeated measurements showed a low correlation; this can be explained by differences in traffic composition, time of day and day to day variability in the processes that govern the UFP dispersion i.e. wind, solar radiation, turbulence, etc. Although the total traffic

showed a very high correlation across the two measurement periods, the correlation for the heavy traffic was lower. Variability of heavy traffic in a 15-minute period is higher than that of the total traffic. Since heavy duty vehicles emit more particles, the variability in the number of heavy vehicles explains, in part, the modest correlation of the UFP repeated measurements, and thus, the modest R^2 of the LUR models. The mean temperature, relative humidity and wind speed in the first and second measurement days were uncorrelated (Pearson correlation coefficient and p-values 0.06, 0.77; 0.13, 0.53; 0.24, 0.24, respectively). Wind away from the monitoring site may account for differences in the UFP observed levels. However, this is unlikely to play a role at most sites given the narrow streets and the short monitor to traffic distance.

One strategy to address the temporal variation is the use of repeated measurements. As shown, the use of two instead of a single 15-minute UFP measurement in 25 sites in Girona city increased substantially the performance of the related local LUR model (from 47 to 72% explained variance). Such rather simple expansion of the protocol may thus result in LUR models with higher performance, while still capitalizing on the major logistic advantages of short-term monitoring periods. Extending the measurement period would better capture the temporal variability, particularly from heavy traffic measurements. Longer measurement periods may though require different logistic approaches.

UFP levels

Compared to larger cities –e.g. Barcelona, Amsterdam – the UPF levels observed in our towns were relatively low. However, the UFP contrasts were larger across the Girona sites as compared to the Amsterdam sites (interdecile range/median was 1.86 compared to 1.09) from the previous UFP LUR study (Hoek et al., 2010). The mean UFP concentration in Girona was 10.523 part/cm^3, around 40% less than the mean levels for the period November 2003–December 2004 in an urban background monitoring site in Barcelona (mean level of particles with diameter 13–800 nm = 17.000 part/cm^3) and about 50% less than the UFP concentrations at the Amsterdam sites. Moderate levels are in correspondence to the size of the towns, and traffic levels in the study area. While the traffic intensity in the major road nearest to the station in Barcelona was 106.000 vehicles per day, the mean traffic in the nearest major road in our sites was 11.500 vehicles per day.

We measured during non-rush hours, a time of lower UFP levels. For traffic density it is known that the annual mean of daily counts are better represented by the traffic during non-rush hours. Since vehicle exhaust is the main source of UFP, a focus on the more representative non-rush hours appears to be an appropriate approach to estimate the mean daily levels. The observed UFP levels may also underestimate the long-term mean due to seasonal factors. Low temperatures and atmospheric stability presumably enhance the formation of ultrafine particles, which show lower levels in the summer (Wang et al., 2010), when the monitoring was done. However, spatial variation may still be similar across seasons, as has been documented for other traffic-related pollutants including NO_2 (Hoek et al., 2008a).

A disadvantage of our study relates to the use of the P-Trak with particles in the lowest size range (10-20 nm) being below detection limit. While the use of P-Trak may have affected the absolute levels of our UFP measurements, the spatial difference between locations and the performance of LUR model is hardly affected. The annual average seven-minute mode of UFP (particle diameter 13–800 nm) concentrations measured in Barcelona was approximately 35 nm (Pey et al., 2009).

Covariate range and measurement error

Model performance was particularly modest in small towns. Low absolute values and a low range in traffic intensity may in part explain this. Similar results were observed in a Danish study where the correlation between traffic intensity and measured NO_2 was much higher within Copenhagen city than in the smaller towns around Copenhagen (Raaschou-Nielsen et al., 2000). The measurement error of the covariates is inherent to the limited precision in the GIS based predictors. Given that UFP concentration drops to background level in a few tens to hundreds of meters from its source, the urban characteristics relevant to UFP dispersion are in this scale. Therefore the precision of the GIS-derived measurements may affect the performance of the predictive models. The exclusion of sites located less than 25 m from a major road intersection did not change the coefficients neither the predictability of our models.

Monitoring Strategy

A major advantage of our approach compared to more conventional campaigns with active samplers left in the field for weeks is the much higher flexibility in site selection as the technician supervises sampling. In most cities, it is a major challenge to find safe locations within major streets which is why investigators rely on first floor sampling such as balconies. Thus, the cost of finding appropriate sites will be diminished.

In light of limited resources, the choice of the monitoring strategy is always a trade-off between optimizing number of locations, number of measurements per location, and duration of measurements. Our study reduces the measurement duration from the usual several weeks to only 15 minutes, while drastically increases the number of sites to a density not seen in any other study (e.g. 69 sites within an area of 2 km^2 in the town of Salt). While this decision was primarily driven by our needs to count traffic in the absence of reliable data, the performance of the UFP model was good. Short-term measurements ranging from only 5 to 13 minutes have been used before to develop a LUR model for black carbon (Larson et al., 2009). In that study the concentrations were measured from a vehicle driving around the four blocks surrounding 39 monitoring locations and the monitoring period was 8 days. Models R^2's ranged from 0.51 to 0.72.

The monitoring strategy we used overcomes one of the main limitations of LUR modeling, allowing a monitoring of a large number of locations in a short study period. Specifically for UFP, current instruments are expensive and/or cannot be left unattended for weeks as is feasible for the integrated PM sampler. Depending on the number of sites and number of repetitions, costs are likely lower as well. It was previously estimated that a campaign with active PM sampling at 40 locations with four 1-week measurements per location required 64 person days of fieldwork (Hoek et al., 2008a, Table 7). Implementing our protocol for a UPF monitoring campaign at 40 locations with four 15-minute measurements in different days would require 17 person days of fieldwork.

The number of sites in LUR models largely depends on logistics and economic resources available. It has been suggested that the sampling size is less important for the performance of the models than the variability of the covariates (Ryan et al., 2007b).

Maximizing the covariate's variability is indeed important. The exposure predictions derived with a LUR model with high variability of the covariates in the monitoring locations would more likely result in more precise health effects estimates (Szpiro et al., 2011a). Nevertheless the sample size could also play a role in the models performance. In a recent study, training sets of different sample sizes (N= 20 to 140) were randomly selected from a dataset of NO_2 monthly mean concentrations in 159 locations in Girona city, while the remaining points constituted the validation sets. For each N the random selection was repeated 200 times. LUR models were derived using the training set, and then validated using the validation set. With decreasing sample size the adjusted R^2 and leave-one-out cross-validation R^2 increased while the validation R^2 (true R^2) decreased. It was concluded that the smaller the sample size the larger overestimation of the true R^2 and the lower predictability and stability of the models (Basagaña et al., 2011). Whether these results extend to other cities and areas remains to be explored.

Conclusion

A LUR model was developed for ultrafine particles based upon short-term monitoring at 644 locations in Girona province that explained up to 42% of the spatial variability. A single 15-minute repeat at a subset of 25 sites, conducted within a few weeks, improved the LUR substantially, thus we recommend the use of repeated measurements. Further efforts to better control factors affecting temporal variation of UFP, such as having continuous measurement of UFP at multiple reference sites, may improve these models as well. We conclude that far shorter monitoring protocols than currently assumed to be needed provide very promising models at lower costs. Such protocols are of particular relevance in resource constraint settings as low income countries and small towns. This approach is also useful to derive a first approximation of the UFP spatial distribution that helps both guiding the design of new epidemiological studies on UFP and deciding where to invest for LUR models.

Acknowledgements

Project funded by ADEME/ANSES (French Environment & Energy Management Agency/ French Agency for Food, Environmental & Occupational Health & Safety), as part of the TRITABS study. Ms. Rivera received funding from DIUE GC/FSE (2009FI_EX 00041), the "Comissionat per a Universitats i Recerca del Departament d'Innovació, Universitats i Empresa de la Generalitat de Catalunya" and the European Social Fund. We would like to acknowledge the field workers: Laura Batchelli, Marc Roca, Meritxell Portella, Joan Masdeu and Jordi Vicens. We are grateful to Lee Jane Sally Liu (2.11.1964-6.6.2011) from the Swiss TPH and her team at the School of Public Health and Community Medicine at the University of Washington, Seattle, WA (USA) for borrowing the P-Traks and technical support.

References

References of this paper are included in Section 10.

Tables and Figures

Table 1. Distribution of monitoring locations across towns. Population in 2010, approximate size of sampled area in km^2, number of sampling sites (N and %), number of days with monitoring activity, and monitoring period.

Town/Area	2010 Population (hab)	Approx. extension of sampled area (km2)	N	%	Number of monitoring days
Girona	96,236	20.6	167	26.1	29
Salt	30,304	2.0	69	11.0	11
Blanes	40,010	6.5	104	16.0	16
Sant Joan les Fonts, Olot	36,340	12.9	38	5.8	4
Santa Cristina d'Aro, Llagostera	12,982	22.6	45	6.9	7
Porqueres, Banyoles	23,180	8.7	76	11.7	10
Palamos, La Bisbal, Palafrugell	51,076	7.9	145	22.6	14
Total			644		32

Table 2. Land use regression model for ln (ultrafine particles - 659.2): Core model. N=644. UFP are expressed in (part/cm3). Predictor variables are divided by the difference between the10th and 90th percentile indicated in parenthesis. Coefficients are expressed as proportional change in UFP for an increase between the 10th and 90th percentile in each of the predictors.

	Coef.	P>t	[95%	CI]	R^2_A	R^2(cv)
Core model						
					0.36	**0.35**
heavy, light and motorcy. veh in 24 hours (veh/9726)	0.433	<0.001	0.35	0.52		
area of high density residential land within 1000m (m2/1930508)	0.355	<0.001	0.20	0.51		
distance to intersection of two major roads (m/904)	-0.21	<0.001	-0.28	-0.13		
household density within 100m (number/184)	0.144	0.008	0.04	0.25		
constant	8.679	<0.001	8.56	8.79		

R^2_A: adjusted coefficient of determination

R^2(cv): cross-validation coefficient of determination

Table 3. Town-specific LUR models and performance of Core model, model adjusted for town and model with effect modifications of town applied to each town.

Town/Area	N	Town specific LUR models			Core model applied to each town		Model adjusted for town applied to each town		Model with effect modifications of town applied to each town	
		Determinants	R^2_A	R^2(cv)	R^2_A	R^2(cv)	R^2_A	R^2(cv)	R^2_A	R^2(cv)
Girona	167	length of major roads within 50m; area of buildings within 100m; length of bus lines within 300m; heavy, light and motorcy. veh. in 24 hours; distance to highway AP-7	0.50	0.47	0.38	0.35	0.37	0.34	0.37	0.32
Salt	69	heavy, light and motorcy. veh in 24 hours; Product T.I on nearest major road*inverse distance to nearest major road; population density within 100m	0.32	0.26	0.17	0.13	0.18	0.12	0.31	0.21
Blanes	104	household density within 100m; heavy, light and motorcy. veh. in 24 hours; area of buildings within 100m; area of residential land within 100m	0.43	0.36	0.27	0.19	0.32	0.28	0.38	0.27
St. Joan, Olot	38	household density within 500m; 24 hour total traffic load of all roads in 150m; N-S orientation of nearest major road	0.55	0.45	0.08	0.00	0.07	0.00	0.13	0.00
Sta.Cristina, Llagostera	45	heavy, light and motorcy. veh in 24 hours; area of natural land within 500m; household density within 1000m; area of industrial land within 1000m	0.76	0.74	0.53	0.45	0.54	0.46	0.63	0.58
Porqueres, Banyoles	76	heavy, light and motorcy. veh in 24 hours; population density within 100m	0.13	0.09	0.10	0.00	0.11	0.07	0.19	0.12
Palamos, La Bisbal, Palafrugell	145	length of major roads within 100m; heavy, light and motorcy. veh. in 24 hours; area of residential land within 1000m; Product T.I on nearest road*inverse distance squared	0.35	0.32	0.33	0.22	0.32	0.27	0.31	0.28

R^2_A: adjusted coefficient of determination

R^2(cv): cross-validation coefficient of determination

Table 4. Land use regression models for ultrafine particles (part/cm3) in locations with repeated measurements. N=25

	Determinants	R^2_A	R^2(cv)
Measurement Day 1	length of major roads within 50m; area of buildings within 100m	0.47	0.39
Measurement Day 2	24 hour total traffic load of all roads in 150m; natural logarithm of distance to nearest regular bus line	0.47	0.40
Average Day 1 and Day 2	length of major roads within 150m; area of urban green land within 500m; length of all roads within 50m; length of bus lines within 25m	0.72	0.68

R^2_A: adjusted coefficient of determination
R^2(cv): cross-validation coefficient of determination

Figure 1. Monitoring locations. Points represent monitoring locations.

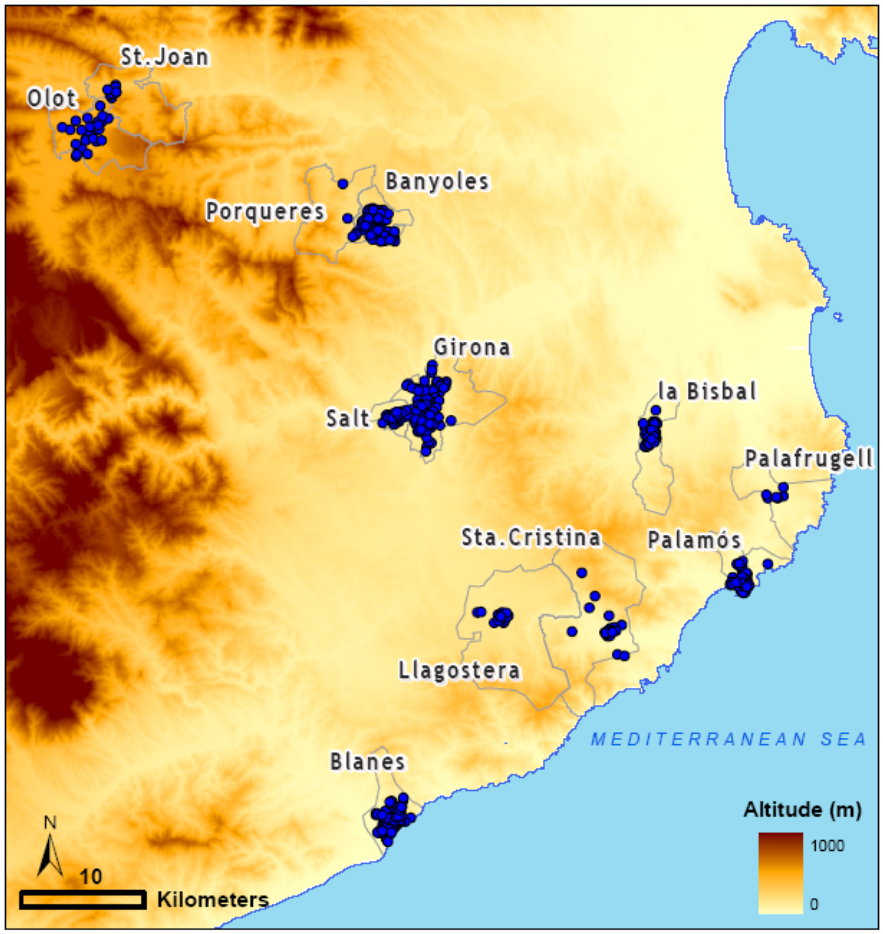


Figure 2. Distribution of ultrafine particles (part/cm3) across towns.

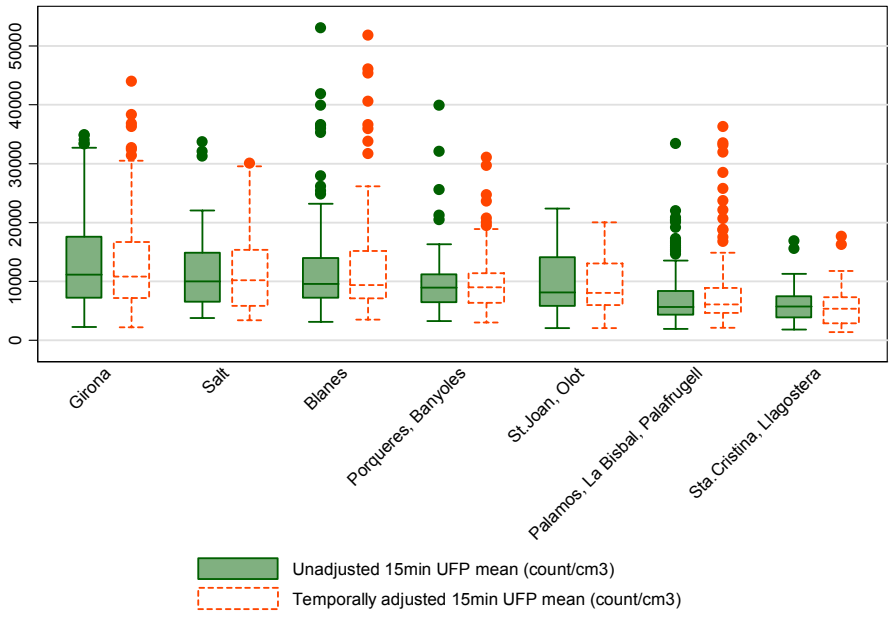


Figure 3. Daily average temperature, relative humidity and wind speed, on monitoring days, at four meteorological stations located in the study area.

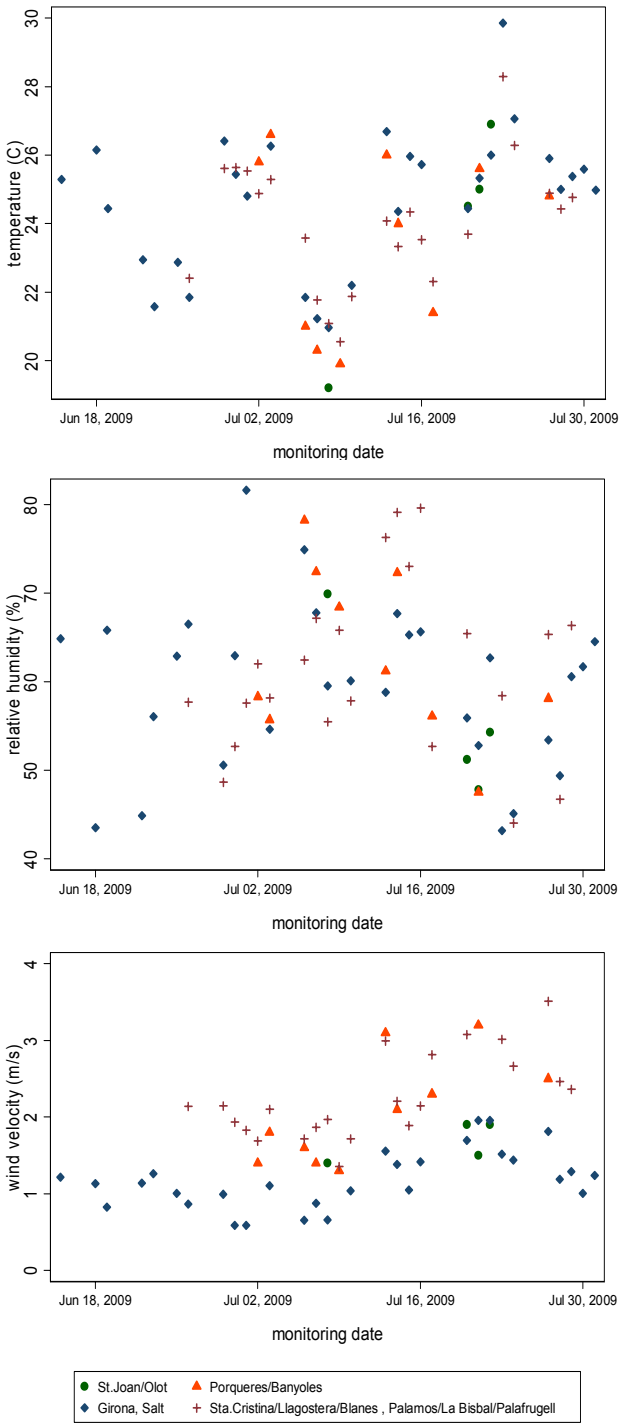
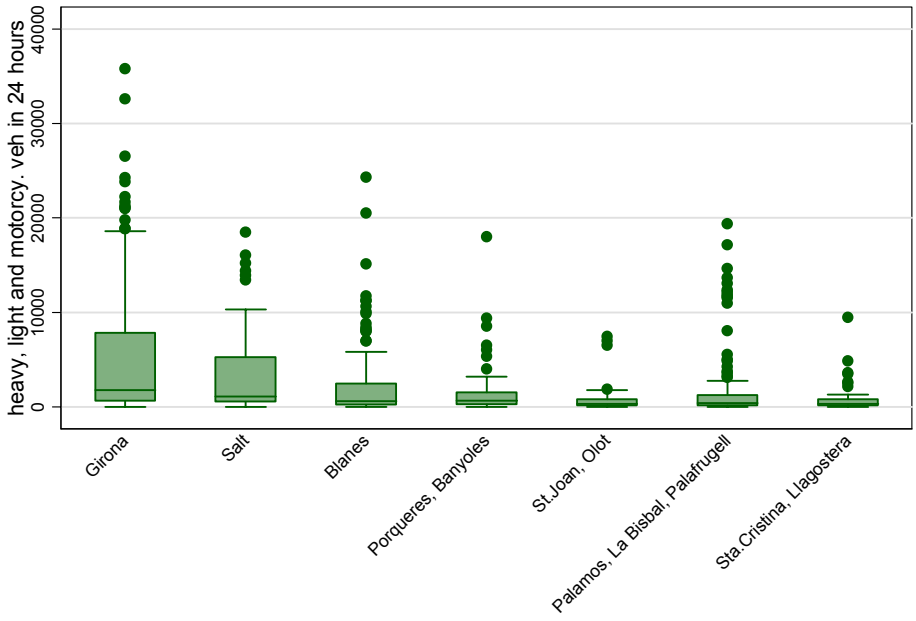
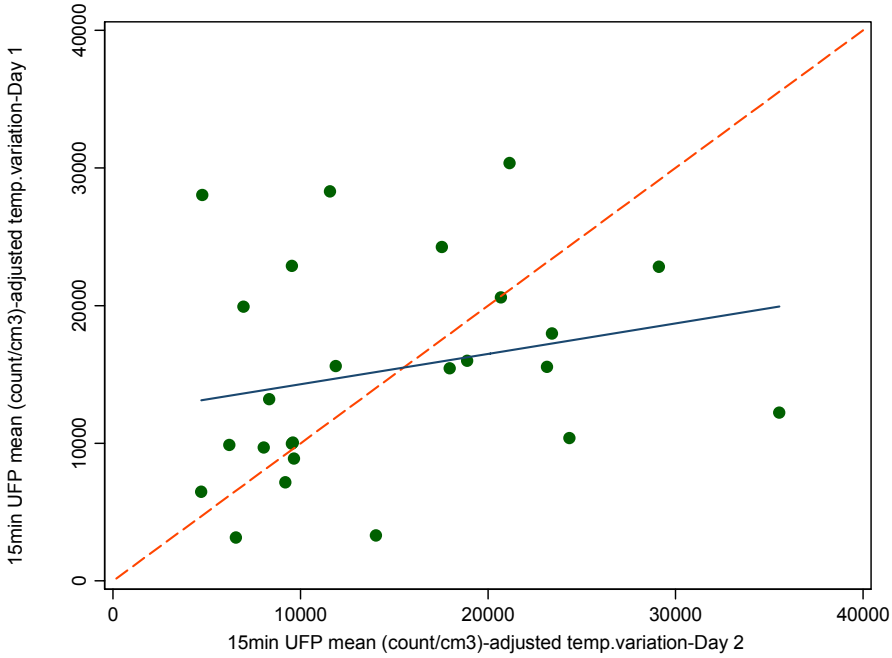


Figure 4. Distribution of heavy, light and motorcycle vehicles in 24 hours (number of vehicles) across towns.



In the box plots, the box represents the interquartile range, solid line in the box is the median. The whiskers extend to the upper and lower adjacent values. The upper adjacent value is defined as the largest data point less than or equal to the 75th percentile + 1.5 x interquartile range. The lower adjacent value is defined as the smallest data point greater than or equal to the 25th percentile + 1.5 x interquartile range. All values outside the adjacent values are represented with dots.

Figure 5. Scatter plot of UFP measured in Day 1 and Day 2. Dashed line: y=x. Continuous line: Fitted values. Pearson correlation coefficient 0.2351, significance 0.2579.



Supplementary Material

Leave one out cross-validation

The leave-one-out cross-validation procedure consists in using *n*-1 observations to fit the model. The resulting model equation is then used to predict the response for the observation that has been left out. This process is repeated *n* times so that each point has a prediction based on a model that did not use the observed value in the model fitting. The leave-one-out cross-validation R^2, R^2(cv), was then calculated as one minus the ratio of the mean squared prediction error (MSPE) to the sample variance of the response. Where MSPE is the average of the squared differences between the observed values y_i and the predicted values $\hat{y}_i$,

$$MSPE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 .$$

Table S1 Distribution of predictor variables that resulted in any LUR model. N=644 except when noted. Minimum, 10th percentile, median, 90th percentile and maximum. Distances are expressed in m, areas are in m^2, traffic intensities in number of vehicles/day. Major roads: with traffic intensity >7000 veh/day. T.I.: Traffic intensity

Predictor Variable	min	p10	p50	p90	max
heavy, light and motorcy. veh in 24 hours	0	69	695	9796	35781
24 hour total traffic load of all roads within 150m	0	731586	3810000	10300000	18500000
length of bus lines within 25m	0	0	0	140	413
length of bus lines within 300m *	0	1019	2219	5599	10202
length of major roads within 50m	0	0	0	106	313
length of major roads within 100m	0	0	0	301	775
length of major roads within 150m	0	0	13	550	1045
length of all roads within 50m	0	133	226	344	583
product T.I on nearest road and inverse distance squared	0	1	34	1602	180000000
distance to highway AP-7†	106.8	759	2249	3210	4738
product T.I on nearest major road*inverse distance to major road	3.92	17	80	1467	261308
distance to intersection of two major roads	1.23	69	299	974	4334
ln. distance to nearest regular bus line †	-2.99	1	3	5	6
population density within 100m	0.54	14	105	494	1522
household density within 100m	0.2	5	42	189	537
household density within 500m	5.04	200	1154	3481	4956
household density within 1000m	31.02	798	3638	10293	14029
area of high density residential land within 1000m	0	271981	1160000	2200000	2980000
area of buildings within 100m	0	3589	9449	16534	21361
area of residential land within 100m	0	7166	31413	31413	31413
area of residential land within 1000m	0	741270	1440000	2370000	3050000
area of natural land within 500m	0	0	0	161765	623566
area of urban green land within 500m	0	0	0	0	258550
area of industrial land within 1000m - source ICC	0	0	49801	435534	806331

* only available for Girona, Salt and Olot
† only available for Girona and Salt.

Table S2. Land use regression models for ln (ultrafine particles - 659.2) adjusted for town and with effect modification of town. N=644.UFP are expressed in (part/cm3). Predictor variables are divided by the difference between the10th and 90th percentile indicated in parenthesis. Coefficients are expressed as proportional change in UFP for an increase between the 10th and 90th percentile in each of the predictors.

	Coef.	P>t	[95%	CI]	R^2_A	R^2(cv)
LUR model adjusted for town						
					0.40	**0.38**
heavy, light and motorcy. veh in 24 hours (veh/9726)	0.456	<0.001	0.37	0.54		
area of high density residential land within 1000m (m2/1930508)	0.25	0.003	0.08	0.42		
distance to intersection of two major roads (m/904)	-0.2	<0.001	-0.30	-0.11		
household density within 100m (number/184)	0.222	<0.001	0.11	0.34		
Sta.Cristina, Llagostera (ref. category)						
Salt	0.072	0.554	-0.17	0.31		
Girona	0.133	0.237	-0.09	0.35		
Porqueres, Banyoles	0.339	0.004	0.11	0.57		
Blanes	0.337	0.003	0.11	0.56		
Palamos, La Bisbal, Palafrugell	0.008	0.942	-0.20	0.22		
St.Joan, Olot	0.429	<0.001	0.19	0.67		
constant	8.54	<0.001	8.30	8.78		
LUR model with effect modifications of town						
					0.42	**0.40**
heavy, light and motorcy. veh in 24 hours (veh/9726)	0.47	<0.001	0.38	0.55		
area of high density residential land within 1000m (m2/1930508)	2.29	<0.001	1.03	3.56		
interaction of area of high density residential land within 1km with:						
Salt	-3.08	<0.001	-4.78	-1.37		
Girona	-2.05	0.002	-3.33	-0.77		
Porqueres, Banyoles	-2.03	0.002	-3.33	-0.73		
Blanes	-2.30	0.001	-3.61	-0.98		
Palamos, La Bisbal, Palafrugell	-2.09	0.003	-3.47	-0.71		
St.Joan, Olot	-1.61	0.027	-3.04	-0.18		
distance to intersection of two major roads (m/904)	-0.04	0.663	-0.20	0.13		
interaction of distance to intersection of two major roads with:						
Salt	1.09	0.015	0.21	1.97		
Girona	-0.13	0.500	-0.52	0.25		
Porqueres, Banyoles	0.33	0.129	-0.10	0.75		
Blanes	-0.11	0.686	-0.65	0.43		
Palamos, La Bisbal, Palafrugell	-0.52	0.006	-0.90	-0.15		
St.Joan, Olot	0.14	0.441	-0.22	0.51		
household density within 100m (number/184)	0.26	<0.001	0.13	0.38		
Sta.Cristina, Llagostera (ref. category)						
Salt	1.24	0.028	0.13	2.35		
Girona	0.92	0.007	0.25	1.59		
Porqueres, Banyoles	0.91	0.007	0.25	1.57		
Blanes	1.30	<0.001	0.58	2.01		
Palamos, La Bisbal, Palafrugell	1.00	0.005	0.30	1.71		
St.Joan, Olot	0.62	0.187	-0.30	1.54		
constant	7.72	<0.001	7.12	8.32		

R^2_A: adjusted coefficient of determination

R^2(cv): cross-validation coefficient of determination

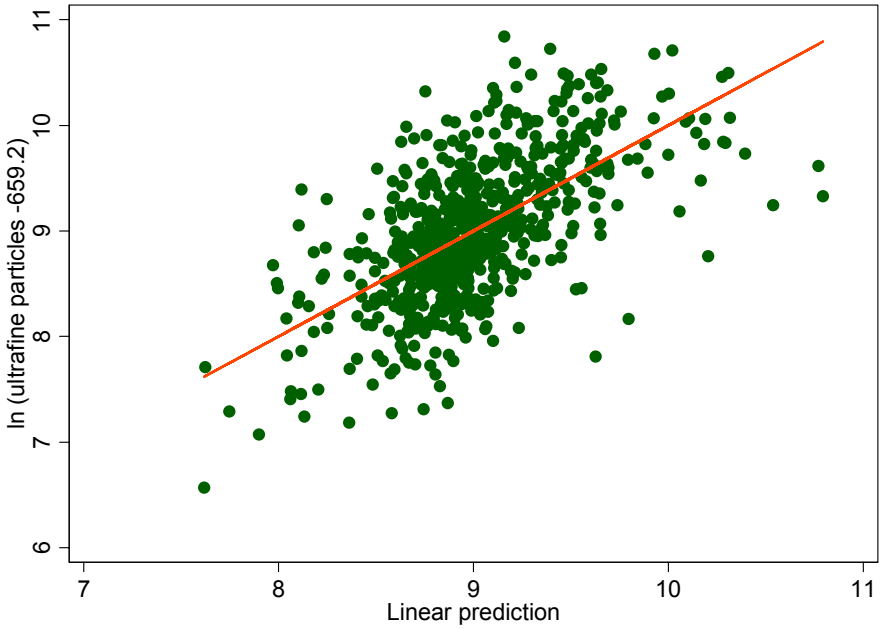
Table S3. Core model for ln (ultrafine particles -659.2) adjusted by monitoring date and hour. UFP are expressed in (part/cm3). Predictor variables are divided by the difference between the10th and 90th percentile indicated in parenthesis. Coefficients are expressed as proportional change in UFP for an increase between the 10th and 90th percentile in each of the predictors. Coefficients for monitoring date are not shown.

	Coef.	P>t	[95%	CI]	R^2_A	R^2(cv)
LUR model adjusted for monitoring date and hour						
					0.51	**0.48**
heavy, light and motorcy. veh in 24 hours (veh/9726)	0.42	<0.001	0.34	0.50		
area of high density residential land within 1000m (m2/1930508)	0.34	<0.001	0.19	0.49		
distance to intersection of two major roads (m/904)	-0.20	<0.001	-0.27	-0.12		
household density within 100m (number/184)	0.17	0.003	0.06	0.28		
monitoring hour (ref. category 9:15-10:30)						
10:30-11:30	-0.30	<0.001	-0.41	-0.20		
11:30-12:45	-0.37	<0.001	-0.47	-0.26		
15:15- 16:00	-0.39	<0.001	-0.50	-0.28		
16:00-16:45	-0.54	<0.001	-0.70	-0.39		
constant	9.38	<0.001	8.68	10.09		

R^2_A: adjusted coefficient of determination

R^2(cv): cross-validation coefficient of determination

Figure S1. Scatter plot of ln (ultrafine particles -659.2) vs. Predicted values using the Core model. Continuous line: y=x



8.3 Effect of the number of measurement sites on Land Use Regression Models in estimating local air pollution (Paper III)

Basagaña, X., Rivera, M., Aguilera, I., Agis, D., Bouso, L., Elosua, R., Foraster, M., de Nazelle, A., Nieuwenhuijsen, M., Vila, J., Künzli, N., In press. Atmospheric Environment. Submitted October 24, 2011, revised version submitted December 23, 2011, accepted January 27, 2012.*

* *This paper is reproduced according to the revised version submitted. References of this paper are included in Section 10.*

EFFECT OF THE NUMBER OF MEASUREMENT SITES ON LAND USE REGRESSION MODELS IN ESTIMATING LOCAL AIR POLLUTION

Authors:

Xavier Basagaña[a,b,c], Marcela Rivera[a,b,c,d], Inmaculada Aguilera[a,b,c], David Agis[a,b,c], Laura Bouso[a,b,c], Roberto Elosua[b,c], Maria Foraster[a,b,c,d], Audrey de Nazelle[a,b,c], Mark Nieuwenhuijsen[a,b,c], Joan Vila[b,c], Nino Künzli[d,e]

Authors' affiliations:

[a] Centre for Research in Environmental Epidemiology (CREAL), Barcelona, Spain;
[b] IMIM (Hospital del Mar Research Institute), Barcelona, Spain;
[c] CIBER Epidemiología y Salud Pública (CIBERESP), Barcelona, Spain;
[d] Universitat Pompeu Fabra (UPF), Barcelona, Spain;
[e] Swiss Tropical and Public Health Institute, Basel, Switzerland;
[f] University of Basel, Basel, Switzerland

Corresponding author:

Xavier Basagaña, Centre for Research in Environmental Epidemiology (CREAL), Doctor Aiguader 88, 08003 Barcelona, Catalonia, Spain (phone +34 93 214 73 06; fax +34 93 214 73 02; e-mail xbasagana@creal.cat).

Funding:

The study and data used in this project were funded by: Fondo de Investigación Sanitaria, FIS PI060258; Fundacio La Marató de TV3, grant number: 081632; the French Agency ADEME (TRI-TABS – 2009/1/121). Ms. Rivera received funding from DIUE GC/FSE (2009FI_EX 00041). Ms. Foraster received funding from the ISCIII (Instituto de Salud Carlos III)

Abstract

Land use regression (LUR) models are often used in epidemiologic studies to predict the air pollution exposure of health study participants. Such models are often based on a small to moderate number of air pollution measurement sites across the study area, and on a set of variables characterizing factors such as traffic patterns and surrounding land uses that are used as potential predictors. We used resampling techniques on a set of 148 measurement sites of NO_2 in the urban area of Girona (Spain) to investigate the effect of the number of measurement sites on the LUR model performance, in particular on predictive ability and on the variables being chosen in the final model. In addition, we investigated the effect of the number of potential predictors and the variable selection algorithm used, and the consequences of the use of LUR predictions in regression models for a health outcome. Our results showed that, especially in small samples, both the adjusted within-sample R^2 and the leave-one-out cross-validation R^2 tended to give highly inflated values when compared to their prediction ability in a validation dataset. When the number of potential predictors was high, LUR models developed with a small number of measurement sites tended to give higher within-sample and cross-validated R^2 than those developed with more sites. Validation dataset R^2 showed a poor performance of models developed with a small number of sites that improved as the number of sites increased. Models developed with a small number of sites tended to select a different set of variables every time, were very sensitive to the number of potential predictors offered and resulted in stronger attenuation of coefficients when air pollution predictions were used in a health model. Our results suggest that LUR models aimed at characterizing local air pollution levels in complex urban settings should be based on a large number of measurement sites (>80 in our setting) and that the set of potential predictors should be restricted.

Keywords: land use regression; measurement error; modeling; NO_2; residential exposure; Spain.

1. Introduction

Measuring individual exposure to outdoor air pollution in a large cohort of study participants is not feasible in terms of cost and logistics. Thus, epidemiologic studies often perform a set of air pollution measurements throughout the study area and use statistical models to predict the exposure at participants' residential addresses. One of these techniques, usually called land use regression (LUR) modeling, incorporates a set of predictor variables that are available for all study participants' residences through Geographic Information Systems (GIS) into a linear regression model to obtain air pollution predictions at unmeasured sites. The set of predictor variables typically includes traffic patterns, surrounding land characteristics and population density. The number of sampling sites often employed are small to moderate (20 to 60), while only a few studies reach 80 or more sampling sites (Hoek et al., 2008a; Ryan et al., 2007a). Most of the LUR models report quite high percentage of variance explained (R^2) with only a few predictors in the final model. Typically, the R^2 of the models is around 70%, with several models reporting values over 80% (Hoek et al., 2008a; Ryan et al., 2007a).

It is known from the statistical literature that, with small datasets and a large number of predictor variables to choose from, it is easy to end up with a final model with a very high within-sample R^2 (or adjusted R^2) that does not reflect the true predictive ability of the model (Derksen et al., 1992; Flack et al., 1987; Hastie et al., 2001; Rencher et al., 1980). A R^2 measure based on leave-one-out cross-validation is often reported to alleviate this problem, although additional difficulties may arise with small datasets (Davison et al., 1997; Hastie et al., 2001; Isaksson et al., 2008; Molinaro et al., 2005). A previous study showed that LUR models developed for New Haven, CT with a small number of measurements had poor out-of-sample predictability even when the in-sample R^2 was high (Johnson et al., 2010). However, their evaluations were conducted on modeled rather than measured air pollution concentrations. In this paper, we use data from the REGICOR-AIR study (www.regicor.org), with 148 NO_2 measurements sites across one geographic urban area, to explore the effects of the number of measurement sites on LUR model performance, including predictive ability and the choice of variables in the final model. In addition, we explore other issues such as the effect of the number of potential predictors used, the

variable selection algorithm, and the consequences of using LUR predictions in regression models for a health outcome.

2 Methods

2.1. Data

We used data from the two adjacent cities of Girona and Salt, two mid-size Mediterranean cities in northeastern Spain covering an area of 45.7 km^2. Measurements of NO_2 were available for 148 sites, mainly study participant residences, evenly distributed across the area (Figure 1). Briefly, NO_2 was measured using Palmes passive sampler tubes in several 1-month campaigns from June 2007 to July 2009, and the different measurements from the same site were averaged to represent annual means. Prior to averaging, each measurement was adjusted for temporal variation, derived from one background in the month with average concentration more approximated to the annual mean. The mean and standard deviation of the annual mean NO_2 concentrations in the 148 sites were 28 $\mu g/m^3$ and 9.1 $\mu g/m^3$, respectively. A set of 106 geographic variables were available to enter a LUR model to predict annual mean NO_2 levels. Geographic variables included surrounding land uses (urban, industrial, green areas), topography (altitude), population and household density, building density, traffic-related variables (average daily traffic, road length, and length of bus lines), and distance to emission sources (roads, parking lots, and bus stops). Some of the variables were computed for circular buffers from 25m to 1000m, depending on the type of variable. As most of the sampling sites were residential addresses, we also included the building floor number where the sampler was placed, as a proxy of the sampler height above ground.

2.2. Statistical Analysis

Training samples of NO_2 measurement sites of size $n = 20$ to 120 were randomly drawn from the original dataset of $N = 148$ sites. In addition, a random sample of $n_v = 28$ sites was held out to validate the model. The process was repeated 300 times for each n. For each dataset, we derived the regression model by applying the following algorithm:

- FORWARD_R2a: a forward selection algorithm. In each step, the variable generating the highest increase in the adjusted R^2 (R^2_a) enters the model, provided that the improvement in R^2_a is greater than 1%, the sign of the regression coefficient agrees with the pre-specified expected sign, and the sign of the remaining coefficients in the model does not change. When no additional variables can improve R^2_a by more than 1%, variables with a p-value >0.10 are sequentially removed from the model.

In order to explore how other variable selection algorithms performed, we also applied the following two algorithms to the same datasets:

- FORWARD_L1OCV: the same algorithm than FORWARD_R2a but based on the leave-one-out cross-validation R^2 (described below) instead of on R^2_a.
- Deletion/Substitution/Addition (DSA) algorithm. This algorithm is described elsewhere (Sinisi et al., 2004) and was developed to provide both good estimation and a good assessment of the model performance. The algorithm uses cross-validation to search for the subset of variables that minimizes the sum of squared residuals. The algorithm allows the incorporation of polynomial transformations of the original variables as well as products of two or more variables (interactions). However, for the purpose of this study, all variables were kept in the original scale and interactions were not allowed. A library implementing the DSA algorithm is available in the R software (R Development Core Team, 2010) (http://www.stat.berkeley.edu/~laan/Software/)

For each final model, we computed four different measures aimed to estimate prediction error:

- The adjusted R^2, R^2_a.
- The leave-one-out cross-validation R^2, R^2_{cv}. In the leave-one-out cross-validation procedure the model is fitted using *n*-1 observations and the resulting model equation is used to predict the response for the observation that has been left out. This process is repeated *n* times so that each site has a prediction based on a model that did not use that same observation in the model fitting step (Hastie et al., 2001). The R^2_{cv} was then obtained from a regression model of the observed against the predicted values.

- Validation set R^2, $R^2_{V(corr)}$. The model derived with the *n* sites is used to predict the observations in the validation dataset of n_v sites. Then, the $R^2_{V(corr)}$ is derived from a regression model of the observed against the predicted values in the validation dataset, or, equivalently, as the squared Pearson correlation coefficient between observed and predicted values.
- Validation set R^2 based on mean squared prediction error, $R^2_{V(MSPE)}$. The previous measure, $R^2_{V(corr)}$, is only based on the correlation of predicted vs. observed values, but it does not take into account their absolute values. When it is important to predict the actual values of air pollution, and not merely a correlate of air pollution, prediction ability is often measured in terms of mean squared prediction error (MSPE), defined as the average of the squared differences between observed and predicted values. One can transform MSPE into a R^2-like formula by computing

$$R^2_{V(MSPE)} = 1 - \frac{MSPE}{\left(\frac{1}{n_v} \sum_{i=1}^{n_v} (y_i - \bar{y}_t)^2 \right)}, \quad (1)$$

where $\bar{y}_t$ is the average of the response in the training sample (Harrell, 2001; Szpiro et al., 2011a). $R^2_{V(MSPE)}$ can yield negative values when, in the validation set, the average of the observed values performs better, in terms of mean squared error, than the predictions of the model.

As a sensitivity analysis, we repeated the same process but sampling the sites in the training sample stratifying by four categories of NO_2 according to quartiles. In this way, all training datasets had the same number of sites in each of the four categories and the range of NO_2 was more homogeneous between datasets. In a separate analysis, we also repeated the $R^2_{V(corr)}$ and $R^2_{V(MSPE)}$ calculation by truncating the values of the predictors in the validation dataset that were outside the range of values observed in the training dataset, in order to prevent the effect of extreme predictions based on extrapolations.

As a result of the resampling and model fitting processes, we obtained 300 final models for each *n*, with their associated R^2 measures and sets of final variables. To summarize the results, we

reported the median and the interquartile range (IQR) of the R^2 measures, the frequency of appearance of each variable in a final model, and listed the ten predictor variables that appeared more frequently for different values of *n*.

To assess how the number of potential predictors affects the R^2 of the final models, we performed two analyses. First, we restricted the number of predictors in the following way. We performed a principal component analysis of the original 106 variables, which resulted in 18 components with an eigenvalue greater than one that explained 90% of the total variation in the dataset. We performed a varimax rotation of the 18 components and picked the variables that had the highest factor loading in each of the 18 components. We repeated all the aforementioned analyses with this restricted set of predictors. Second, we generated an independent set of 106 random predictors which, together with the NO_2 measurements, entered the same resampling and model fitting process described above. The resulting R^2 of these models based on random predictors was then studied for different values of *n* and different numbers of random predictors offered.

2.2.1. Measurement error

In epidemiological studies, the air pollution predictions from a LUR model are used to assign exposure values to study participants, and this exposure is then used as an explanatory variable in a regression model for a health outcome. Because of the imperfect measure of exposure, the health model is then subject to the effects of measurement error in covariates (Carroll et al., 2006). Measurement error from LUR models is often considered to be in the form of Berkson error. For the sake of clarity, in this explanation we do not consider the classical-like measurement error that would be introduced by estimating the regression coefficients of the exposure model and will consider that the regression coefficients are known (Szpiro et al., 2011a; Szpiro et al., 2011b). If we call X_i the true exposure and W_i the imperfect measure of exposure, then Berkson error is defined as

$$X_i = W_i + U_i, \quad (2)$$

where U_i is the error term, which is assumed to have mean zero given W_i. Thus, under (2), the true exposure has more variability than the estimated exposure. A LUR model has the form

$$X_i = \gamma_0 + \gamma_1 Z_1 + \gamma_2 Z_2 + \Lambda + \gamma_k Z_k + U_i \quad (3)$$

where the variables Z_j are the final covariates in the model and the γ_j are their associated regression coefficients. By setting $W_i = \gamma_0 + \gamma_1 Z_1 + \gamma_2 Z_2 + \Lambda + \gamma_k Z_k$, one can clearly see that the LUR equation (3) has the Berkson error form (2).

Suppose that we are interested in a health outcome Y_i, for which we would like to fit the regression model

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i. \quad (4)$$

Pure Berkson error, as the one defined in equation (2), has the important property that, if we fit a regression model of Y_i on W_i, we will obtain unbiased estimates of β_1. This is readily seen by plugging equation (2) into (4) to obtain

$Y_i = \beta_0 + \beta_1 W_i + \beta_1 U_i + \varepsilon_i = \beta_0 + \beta_1 W_i + \varepsilon_i^*$, and then taking expected values.

An often overlooked fact is that performing model selection with a large number of candidate variables may lead to an alteration of the pure Berkson error model because of overfitting, which will in turn bias the coefficients of the health model when W_i is used instead of X_i. When the model is developed in a training sample to predict the exposure in an independent sample, the slope of the regression of observed vs. predicted in the independent sample will be less than one, i.e.

$$X_i = \alpha_0 + \alpha_1 W_i + U_i, \quad (5)$$

where $\alpha_1 < 1$ (Harrell, 2001; Harrell et al., 1996). This is a result of regression to the mean, by which the high model predictions will be too high and the low model predictions will be too low in the new

sample. Equation (5) is equivalent to a regression calibration equation for classical measurement error (Carroll et al., 2006). Now, if we plug (5) into (4) we obtain $Y_i = \beta_0 + \beta_1(\alpha_0 + \alpha_1 W_i + U) + \varepsilon_i = \beta_0^* + \alpha_1\beta_1 W_i + \varepsilon_i^*$, and by taking expected values we can see that the coefficient associated to W_i estimates α_1 times the coefficient we would obtain if we used the true X_i, β_1. Since $\alpha_1 < 1$, using W_i instead of X_i leads to an attenuation of the coefficient. We call α_1 obtained from (5) the attenuation factor (AF). Based on the performance of the model in the validation datasets, we calculated the attenuation factor as a function of *n*.

Measurement error may not only have effects on the regression coefficients (bias) but will also have an effect on the standard errors. In particular, the standard errors of the coefficients obtained in the regression of Y_i on W_i are underestimated and one needs to use special methods to obtain the correct ones (Carroll et al., 2006). In this paper we will not study the effects on standard errors.

3. Results

3.1. Prediction ability

Figure 2a illustrates the median of the four different measures of prediction ability over 300 datasets. R^2_{CV} produced slightly smaller values than R^2_a and both measures decreased when the number of training sites *n* increased. This decrease was faster at smaller values of *n*. R^2_a decreased from around 80% for *n*=20 to around 60% for *n*=120. The two validation measures of R^2 ($R^2_{V(MSPE)}$ and $R^2_{V(corr)}$) showed the opposite pattern, i.e. the R^2 increased when the number of training sites increased. The differences of R^2_a or R^2_{CV} with the validation measures decreased as *n* increased, but the difference was still of around 10% at *n*=120. $R^2_{V(MSPE)}$ had negative values for *n*=20, indicating that, for the validation set, the sample mean of the training set had a better predictive ability, in terms of prediction mean squared error, than the predictions of the model. However, in terms of correlation between predictions and measured values, $R^2_{V(corr)}$ achieved values around 20% at *n*=20, which were

still much smaller than the values around 70%-80% of R^2_a or R^2_{CV}. $R^2_{V(MSPE)}$ and $R^2_{V(corr)}$ tended to give similar values as *n* increased.

Figure 2b shows the variability of the R^2 measures over the 300 datasets.
R^2_a and R^2_{CV} showed the smaller variability, with an IQR of 20% at n=20 that decreased to around 5% at n=120. $R^2_{V(MSPE)}$ showed a very high variability for small values of *n*, which makes its use impractical in real settings. $R^2_{V(corr)}$ had IQR values of around 25%, which did not decrease much when *n* increased.

The results were almost identical when the selection of sites for the training dataset was stratified according to NO_2 levels. When the predictors in the validation dataset were truncated to the range of values observed in the training dataset, the median $R^2_{V(corr)}$ improved around 10% for n=20, and the median $R^2_{V(MSPE)}$ moved from a negative value to 12% for n=20. For both measures, the benefits obtained by truncating predictors vanished at around n=70.

3.2. Attenuation of health-exposure associations

In Figure 3 we show the performance of the LUR models in terms of the attenuation of regression coefficients of a hypothetical health outcome model, when using the predictions of the LUR model as an explanatory variable instead of the true (measured) exposure. With the original set of 106 predictors we found that, for n=20, the regression coefficients from such a model could be halved compared to the true ones. The attenuation was quite pronounced for small values of *n*. It was not until n=70 that the attenuation reached 0.8, and until n=110 that it approached 0.9.

3.3. Variables selected

For every one of the 300 datasets for each *n* a different final model could be obtained. We pooled the results of these models and reported the percent of times each variable was selected into the final models in Figure 4. When the training dataset contained 120 sites, 34 of the original 106 variables made it at least once into the final model, but only seven of them appeared in more than 25% of the models. When *n* decreased, a wider range of variables appeared at least once in the final model, and even the variables appearing more frequently were only included in the final model in a small percentage of datasets. In the extreme case of

n=20, almost all the original variables were in at least one final model, and the variable appearing most often was only present in less than one third of them.

Table 1 includes a list of the ten variables most frequently selected for each of n=120, n=80 and n=20, which after taking repetitions into account, corresponds to a set of 15 unique variables. The three variables appearing in more than 80% of the models for n=120 also appeared in the first three positions for n=80, although in that case they appeared in less than 60% of the models. For n=20, the most frequently selected variable appeared in only 33% of the models, and all the rest had a small probability of being selected.

3.4. Effect of number of potential predictors

When we performed the analysis with the restricted set of 18 predictors, the values of R^2_a and R^2_{CV} were reduced, especially for small n (Figure 5). The decreasing trend of R^2_{CV} with n that was observed with the set of 106 predictors disappeared. This trend was still observed for R^2_a but it had a flatter slope. Both R^2_a and R^2_{CV} still showed higher values than the two validation measures ($R^2_{V(MSPE)}$ and $R^2_{V(corr)}$), especially for small n. All R^2 measures were smaller than in the case of 106 predictors because variables with good predictive properties may not be in the restricted set. The attenuation factors were slightly higher compared to the ones obtained using 106 predictors, but were still low (e.g. 0.57 for n=20) and did not reach 0.9 until n=90 (Figure 3).

Figure 6 shows the R^2_{CV} obtained when applying the variable selection algorithm to the NO_2 measurement sites and a set of randomly generated predictors. The number of predictors showed to be an important factor especially for small n. When n=20, we could achieve cross-validated R^2 as high as 80% by using 106 variables generated at random. This decreased to 20% when only 50 variables were offered to the algorithm. As soon as n increased, the problem became much less severe, and with models developed with n=50 the final R^2_{CV} was less than 10% even with 106 potential predictors.

3.5. Variable selection algorithm

Figure 7a shows the distribution of the number of variables in the final model for the three model selection algorithms. The number of variables in the final model was higher for FORWARD_R2, followed by FORWARD_L1OCV and DSA included the lowest number of predictors. At n=20, the average number of variables was 4.6 for FORWARD_R2, 3.4 for FORWARD_L1OCV, and 1.1 for DSA, which chose an empty model a substantial number of times. In those cases, the DSA algorithm did not provide a LUR model that could be used to predict pollution levels. When n increased, the number of variables in the final model increased as well (average numbers for n = 120 were 5.8, 5.2 and 4.5, respectively).

Figure 7b shows the number of variables, out of the 106 original ones, that appeared in at least one of the final models in the 300 datasets. FORWARD_R2A and FORWARD_L1OCV behaved very similarly. While they included almost all variables at least once for n = 20, DSA was a bit more restrictive. However, for n = 120, the number of included variables decreased to less than 40 for FORWARD_R2A and FORWARD_L1OCV while it was still 60 for DSA.

The predictive ability of the final models obtained with the three algorithms is described in Figure 7c. FORWARD_L1OCV had a similar behavior than FORWARD_R2a but with slightly smaller values for R^2_a. DSA behaved differently. Because the final DSA models included fewer variables, they reached smaller values of R^2_a, starting from 50% for n=20 and slightly increasing as n increased. For training datasets with more than 100 sites, the R^2_a obtained from DSA were very similar to those obtained with the other algorithms. In terms of prediction ability in the validation datasets ($R^2_{V(corr)}$), DSA performed similarly to the other algorithms for n=20, but slightly worse for larger values of n.

In terms of attenuation of the coefficients in a health model, the predictions from a model developed with the DSA algorithm performed slightly better for models developed with small n (Figure 7d). This occurred despite the fact that the validation R^2s were smaller for the DSA models.

4. Discussion

We explored the effect of the number of measurement sites on several performance aspects of LUR models using real data. Our results show that, if a high number of potential predictors is offered to the model, both the adjusted R^2 and the leave-one-out cross-validation R^2 tended to provide higher values when the models were developed with a small number of sites than in cases where more sites were available. When the R^2 were computed from validation datasets, though, models developed with a small number of sites performed rather poorly, and the performance improved as the number of sites increased. The differences between the validation R^2 and the other measures could be as high as 60% for a situation with small number of sites and high number of potential predictors. Restricting the number o potential predictors helped in reducing these differences, but models developed with a small number of sites still showed inflation of the adjusted and cross-validated R^2, and attenuation of the regression coefficients when air pollution predictions were used in a health model

We confirmed with real air pollution measurements that LUR models developed with a small number of sampling sites may have a poor performance in predicting air pollution levels even if the cross-validation R^2 of the model is high. The same result was obtained in a previous study that used the predictions from a hybrid dispersion and regional model as the true ambient levels for NO_2, benzene and PM2.5 to fit and evaluate LUR models with training data sets ranging from n = 25-285 and corresponding validation data sets ranging from n_v = 293-33 in size (Johnson et al., 2010). Of note, our results in Figure 2a were very similar to their results in Figure 6b. That study found that the values of the adjusted R^2 and the validation R^2 began to converge for models developed with 125 sites, which is similar to what we observed. Compared to that study, our measurements were collected in a more compact area, suggesting that, even in small areas, a substantial number of measurement sites are needed to correctly characterize the air pollution levels in a complex urban structure. As in Johnson et al. (2010), we found that the adjusted and cross-validated R^2 were inversely correlated with the number of measurement sites, indicating that even when presenting cross-validated R^2 measures, it is more likely to obtain inflated values in smaller studies. A previous review found no correlation between the published R^2 and the number of

sampling sites (Ryan et al., 2007a), but this result was probably due to the between-study design of that comparison.

The adjusted R^2 of a model is known to inflate its true prediction ability (Derksen et al., 1992; Flack et al., 1987; Hastie et al., 2001; Rencher et al., 1980). Part of this inflation is due to using the same sites to develop the final model and to assess its performance. A solution to that is to test the model in a separate held out dataset, or alternatively to use techniques such as cross-validation (e.g. leave-one-out cross-validation) or the bootstrap that create several training and validation datasets via resampling, making a more efficient use of the data. When variable selection is performed on a large number of potential predictors, there is additional inflation of the adjusted R^2 and also of the cross-validated measures, since they are derived using the same data points that were used to select the variables in the final model (Hawkins, 2004). One solution that has been proposed is to split the data into three parts: a training set to fit the models, a validation set to estimate prediction error for model selection and a test set to evaluate the prediction ability of the final model (Hastie et al., 2001). Since this solution is often logistically unfeasible in practice due to limited number of samples, other methods, such as incorporating cross-validation in the selection process (FORWARD_L1OCV) (Hawkins, 2004) or the DSA algorithm (Sinisi et al., 2004) have been suggested.

In our data, neither FORWARD_L1OCV nor the DSA algorithm solved the problem of inflated R^2. The DSA algorithm was the most restrictive in terms of including variables into the model. This resulted in much smaller values of R^2_a for small n, although they were still too optimistic when compared to the validation R^2. For n=20, none of the 106 variables could enter the model under the DSA algorithm in one third of the datasets. Although in such cases this would leave the researcher without a LUR model to predict air pollution levels, this reflects a true limitation of the reduced set of measurements, which is also confirmed by the fact that the validation R^2 obtained from FORWARD_L1OCV did not reach 15% in a third of the models.

The two validation R^2 measures, $R^2_{V(MSPE)}$ and $R^2_{V(corr)}$, were obtained via held out data. Since the LUR models are intended to predict air pollution levels at locations that were not used to develop the model, these two measures can better reflect the prediction

ability in that separate dataset of locations. $R^2_{V(MSPE)}$ and $R^2_{V(corr)}$ showed an increasing trend when the number of measurements increased. This result is more intuitive than the one obtained for R^2_a and R^2_{CV}, since one expects models developed with more data to be more robust and to have better properties. When the aim of the LUR model is to use the predictions as covariates in a regression model for a health outcome, a variable that correlates with the true values is still useful to detect associations even if the absolute values show strong departures. In some cases, for example when pure Berkson measurement error can be assumed, they are also useful to quantify the size of health effects. $R^2_{V(corr)}$ was based on the squared correlation between observed and predicted values in the validation dataset, and therefore it may be more informative than $R^2_{V(MSPE)}$ in the situation just mentioned. However, we have shown that there are some cases where a model with predictions highly correlated with true values but that are not well-calibrated can still bias the coefficients of the health model. $R^2_{V(MSPE)}$ can be more useful to detect those cases. $R^2_{V(corr)}$ has the limitation that one can get a positive value even if the predictions are negatively correlated with the observed values in the validation dataset. This actually occurred in 5% of the cases for n=20. For large n, $R^2_{V(MSPE)}$ and $R^2_{V(corr)}$ produced similar values, although they differed for small n.

Our results were based on sampling small datasets from a larger one. However, in practice, only one small dataset is often available. Several papers report that holding out some measurements does not make an efficient use of the data and advocate for other procedures such as cross-validation or the bootstrap (Hawkins et al., 2003; Molinaro et al., 2005). In our data, the leave-one-out cross-validation R^2 was upwardly biased, and similar results were obtained when we applied other procedures such as K-fold cross-validation or the 0.632 bootstrap to our data (Figure S1 in the Supplementary material). On the other hand, the cross-validation R^2 is a much less variable measure than the R^2 obtained from a validation dataset, so the investigator is faced with the usual trade-off between bias and variance, with the additional consideration of costs. Small validation datasets have been considered of no value because of bias and very high variance. In particular, it has been suggested that small validation datasets can lead to downwardly biased validation R^2 (Hawkins, 2004; Hawkins et al., 2003). Our results on median $R^2_{V(corr)}$ did not change when the validation set size was increased (Figure S2 in the Supplementary material). If the mean $R^2_{V(corr)}$ was computed instead of the median,

slightly smaller $R^2_{V(corr)}$ were obtained when using larger validation datasets, indicating that computing the mean R^2 may be sensitive to the size of the validation set. In terms of variability, $R^2_{V(MSPE)}$ showed an extreme variability at small training sample sizes (e.g. IQR for $R^2_{V(MSPE)}$ was greater than 80% for n=20), but the variability was smaller for $R^2_{V(corr)}$. The IQR of $R^2_{V(corr)}$ for n=20 was around 25%, very similar to that of R^2_{CV}, although for the latter only n=20 measurement sites were used compared to $n+n_v$ = 20+28 for $R^2_{V(corr)}$. Increasing the size of the validation dataset had the expected consequence of reducing the variability of the estimates (Figure S2 in the Supplementary material).

LUR models are often used in epidemiologic studies to predict air pollution exposure of study participants. This predicted exposure is then used as an independent variable in a regression model for a health outcome. Because the LUR model does not exactly predict the observed exposure, the use of this imperfect measure of exposure introduces the effects of exposure measurement error in the health model, namely an attenuation of the regression coefficient of exposure (compared to the one that would be obtained if the true exposure was used) and a loss of power (Carroll et al., 2006). LUR models are usually evaluated by a R^2 measure, but this does not directly inform about attenuation of regression coefficients. We computed the attenuation factor to directly see how this is affected by the number of measurements. In our data, coefficients could be halved if the LUR model was developed with only 20 measurement sites, although the leave-one-out cross-validation R^2 would wrongly suggest a good fit of over 70% of variance explained. It was not until more than 100 measurement sites were used that the attenuation factor reached values of 0.9. Interestingly, even the validation R^2 does not directly inform about attenuation, as shown by the predictions obtained with the DSA algorithm, which resulted in less attenuation than predictions obtained with FORWARD_R2a, even though they had smaller validation R^2. This phenomenon has been recently described elsewhere (Szpiro et al., 2011a). In addition to attenuation, measurement error also produces inflation of the standard errors of the coefficients in the health model. This was not investigated in detail here because results would be study-dependent, since the inflation of the standard errors depends on the size of the health-exposure association and on the residual variance of the health model (Carroll et al., 2006).

We also investigated the effect of the number of potential predictors in the performance of LUR models. We found that having a small number of measurement sites and a large number of potential predictors, it is very easy to find models with very high cross-validation R^2 even when no relationship exists. This has been known for quite some time in the statistics literature (Flack et al., 1987). In addition, it has been reported that when there is substantial correlation between the potential predictors, which is often the case with the geographic variables used to develop LUR models for air pollutants, the risk of selecting noise variables into the final model increases (Derksen et al., 1992; Flack et al., 1987). In practice, it is therefore recommended to restrict the pool of potential predictors to those that are believed to have the strongest effect, especially for small datasets, and to avoid using highly collinear variables.

The results obtained for the variables selected in the final model were those that were initially expected, i.e. as the number of measurement sites increased, the number of variables appearing in at least one model decreased, which is interpreted as having less chance of including noisy variables in the final model (Flack et al., 1987). However, we acknowledge that our experiment was not perfect in the sense that, given the total sample size (148 sites), the different datasets of size $n = 120$ contained mostly the same set of sites and therefore were more likely to always select the same variables.

Even though we only had data on NO_2, many of the statistical mechanisms involved in producing our results will be shared in problems developing LUR models for other local air pollutants. Our study area was mostly urban space of 45.7 km^2. While this setting is similar to many others where LUR models have been developed, others have developed spatial models for far larger geographic areas, often with various geographic clusters. Whether and how our findings can be generalized to such modeling conditions is uncertain but worth investigating.

In summary, LUR models based on a high number of sampling sites have a better performance and are to be preferred. However, this superiority can be masked by the adjusted R^2 and the leave-one-out cross-validation estimate, which tend to give higher and more inflated values for smaller sample sizes, particularly in cases where the number of potential predictors is high and model selection

is performed. Epidemiological analyses using LUR-based exposure estimates need to address the biases that depend on the number of sites available for the LUR modeling.

References

References of this paper are included in Section 10.

Table 1. Percent of times a variable is selected into the final model for different values of *n*. Only variables appearing in the top ten for *n* =120, *n*=80 or *n*=20 appear in the table.

Variable names	Rank (n=120)	% appearance (n=120)	Rank (n=80)	% appearance (n=80)	Rank (n=20)	% appearance (n=20)
Area of buildings within 500m	1	90	2	58	14	8
24 hour total traffic load of all roads in 25m	2	88	3	58	10	10
Floor	3	83	1	60	1	33
Length of major roads within 500m	4	60	7	27	16	8
Household density within 100m	5	58	5	39	4	13
Distance to highway AP-7	6	50	4	44	2	21
Background concentration of elemental carbon	7	48	6	38	3	15
24 hour total traffic load of major roads in 300m	8	20	8	18	34	6
Area of high density residential land within 100m	9	16	13	10	20	7
Length of major roads within 1000m	10	8	9	14	44	4
Length of bus lines within 25m	16	4	19	6	5	12
Length of all roads within 25m	18	3	12	12	9	10
Area of buildings within 300m	26	1	10	13	7	10
Area of buildings within 100m	-	-	37	2	6	11
Population density within 100m	-	-	46	1.3	8	10

Figure 1. Map of the study area.

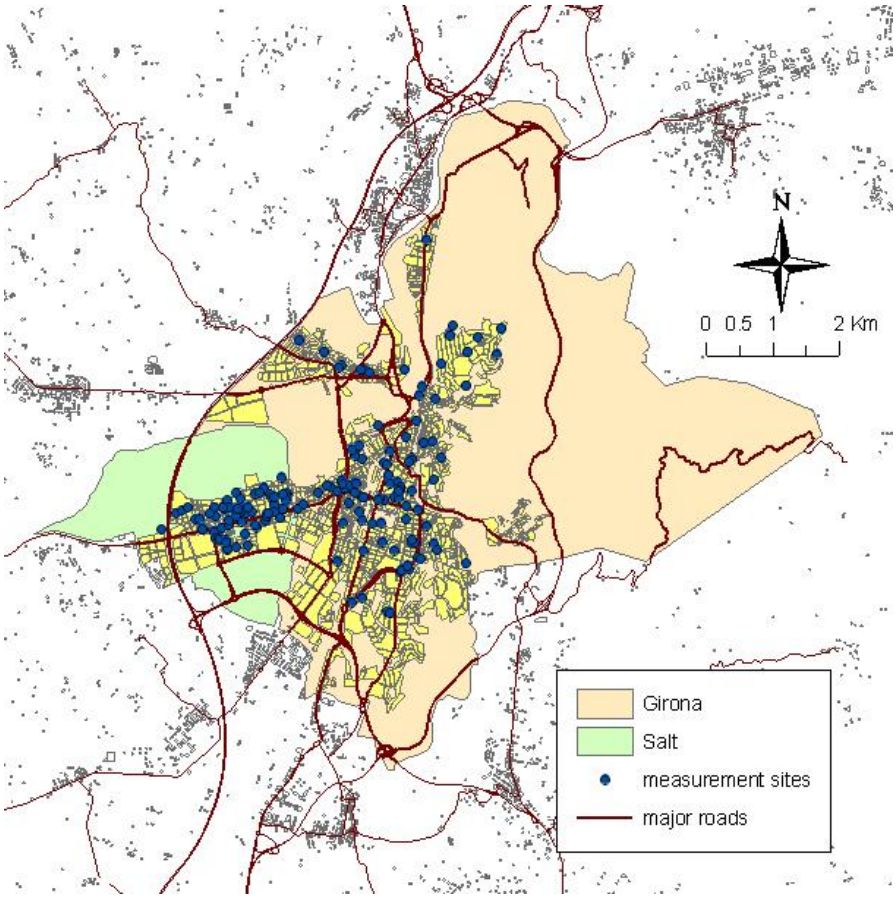


Figure 2. Percent of variance explained, as measured by different R^2 indicators, as a function of the number of measurement sites in the training dataset. Panel (a) shows the median R^2 of 300 final models for each *n*, and panel (b) its variation measured by the interquartile range (IQR).

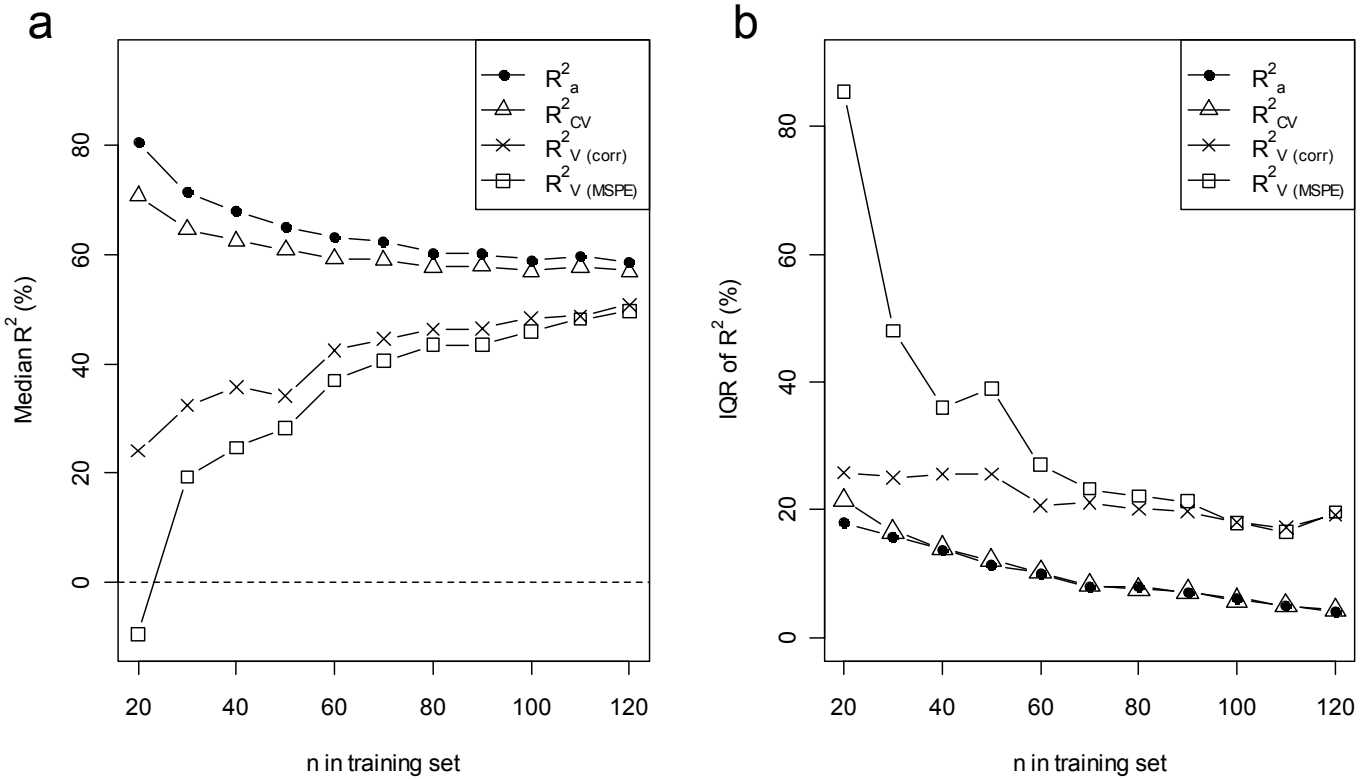


Figure 3. Attenuation of the association between a hypothetical health outcome and air pollution when the association is evaluated using the LUR predictions of exposure instead of the true exposure levels. In other words, the attenuation factor is the ratio between the regression coefficient obtained from the model 'health outcome vs true exposure' and the regression coefficient obtained from the model 'health outcome vs LUR predicted exposure'.

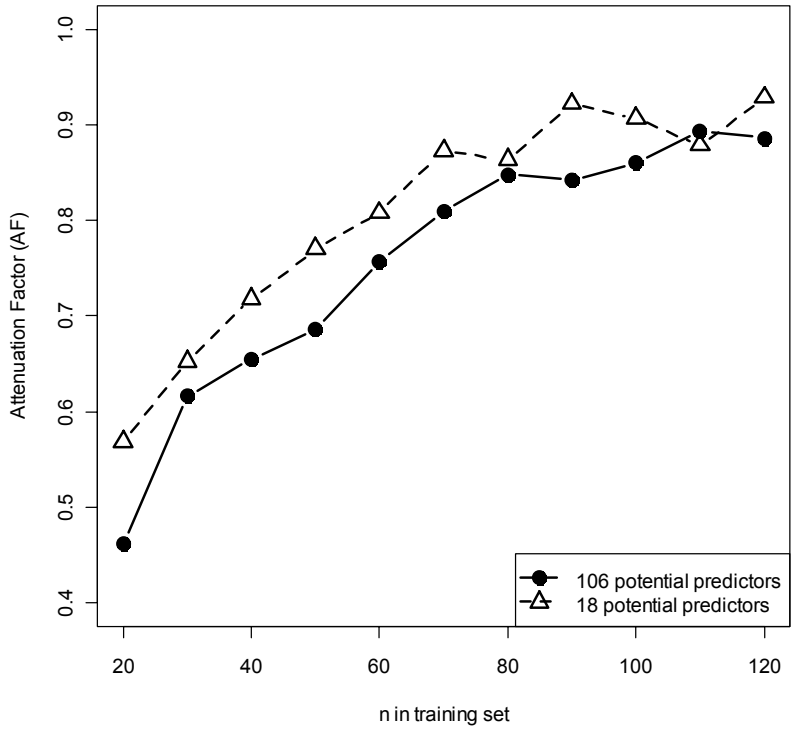


Figure 4. Percent of times a variable is selected into the final model. Variables are sorted in the x-axis according to the percent of times they were selected for the $n = 120$ case.

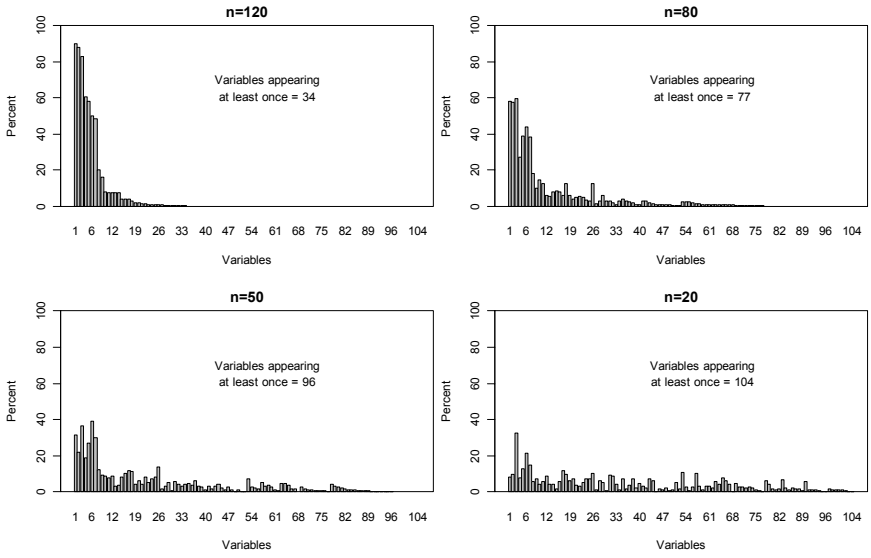


Figure 5. Percent of variance explained, as measured by different R^2 indicators, as a function of the number of measurement sites in the training dataset. Results are based on the restricted set of 18 predictors.

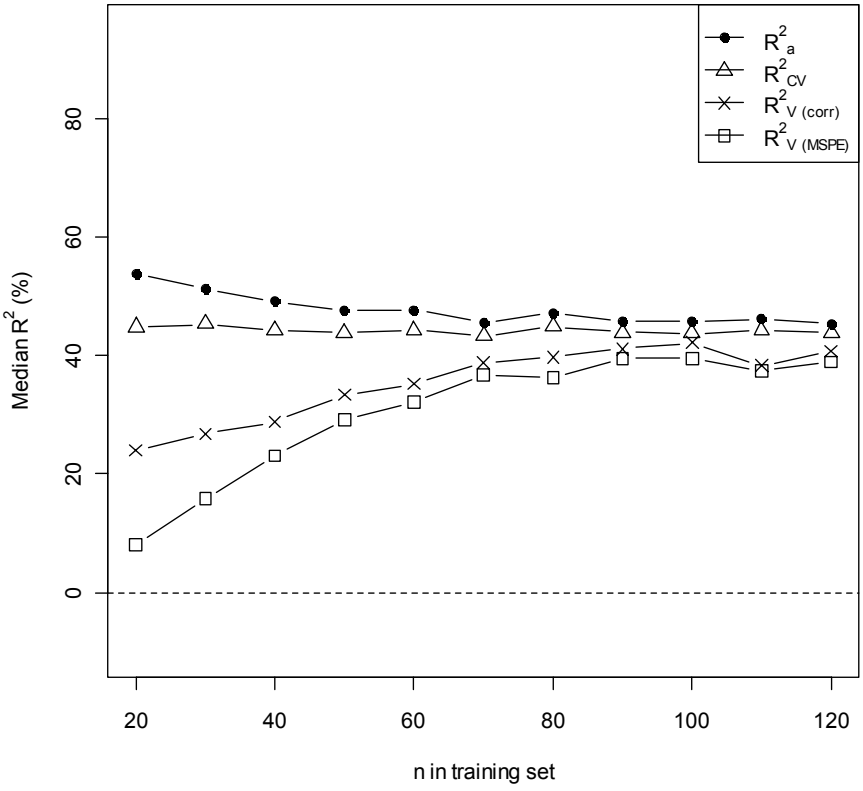


Figure 6. Median R^2_{CV} of 300 final models for each *n* as a function of *n*, when the variable selection algorithm was applied to the NO_2 measurements using as potential predictors a set of 106, 80 or 50 randomly generated variables.

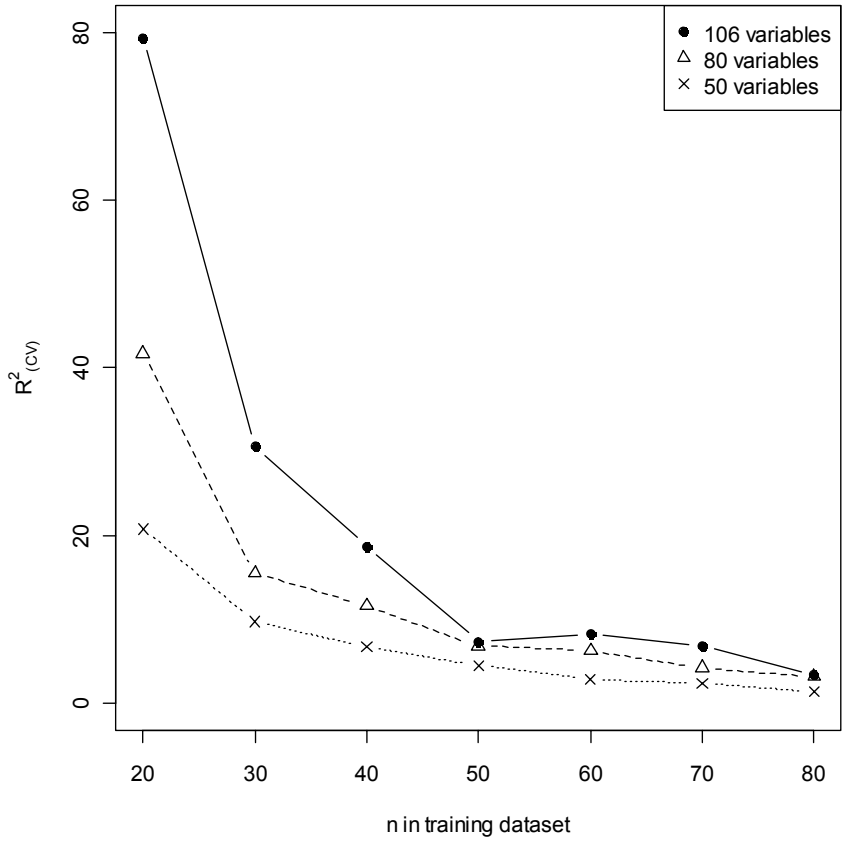
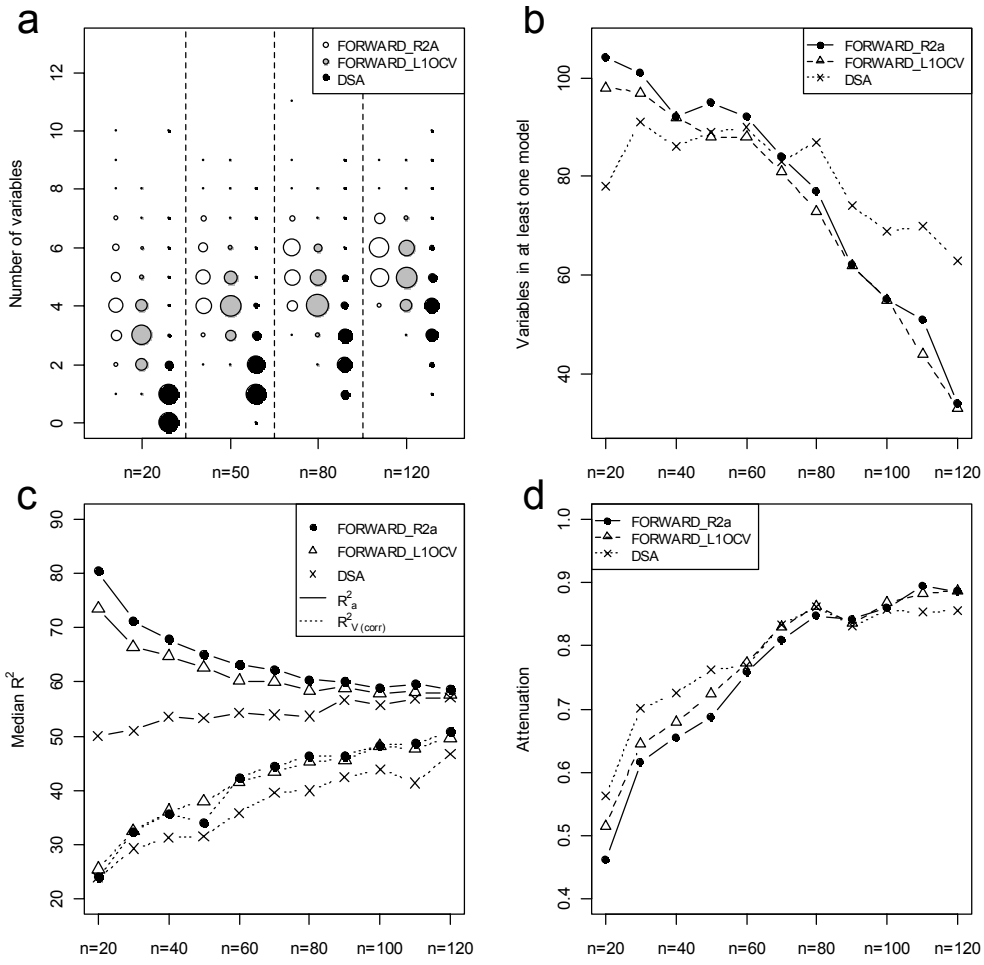


Figure 7. Performance of the three model selection algorithms in terms of: a) Number of variables in the final model. The size of the dots is proportional to the frequency. b) Number of variables appearing in at least one of the 300 final models for each *n*. c) Percent of variance explained, as measured by R^2_a and $R^2_{V(corr)}$, as a function of the number of sites in the training dataset. All points in the graph are medians of 300 R^2 measures obtained from the final models for each *n*. d) attenuation of the regression coefficient when the exposure predicted by the LUR model is used instead of the measured exposure in a model regressing a health outcome against exposure.



Supplementary material

THE EFFECT OF THE NUMBER OF MEASUREMENT SITES ON LAND USE REGRESSION MODELS OF LOCAL AIR POLLUTION

Xavier Basagaña, Marcela Rivera, Inmaculada Aguilera, David Agis, Laura Bouso, Roberto Elosua, Maria Foraster, Audrey de Nazelle, Mark Nieuwenhuijsen, Joan Vila, Nino Künzli

Figure S1. Percent of variance explained, as measured by the 0.632 bootstrap and the *K*-fold cross-validation (CV) R^2, as a function of the number of measurement sites in the training dataset. *K* is defined as the smallest integer not less than min ($n^{1/2}$, 10) as suggested in Davison and Hinkley, (1997). $R^2_{V(corr)}$ is added for comparison. Panel (a) shows the median R^2 of 300 final models for each *n*, and panel (b) its variation as measured by the interquartile range (IQR)

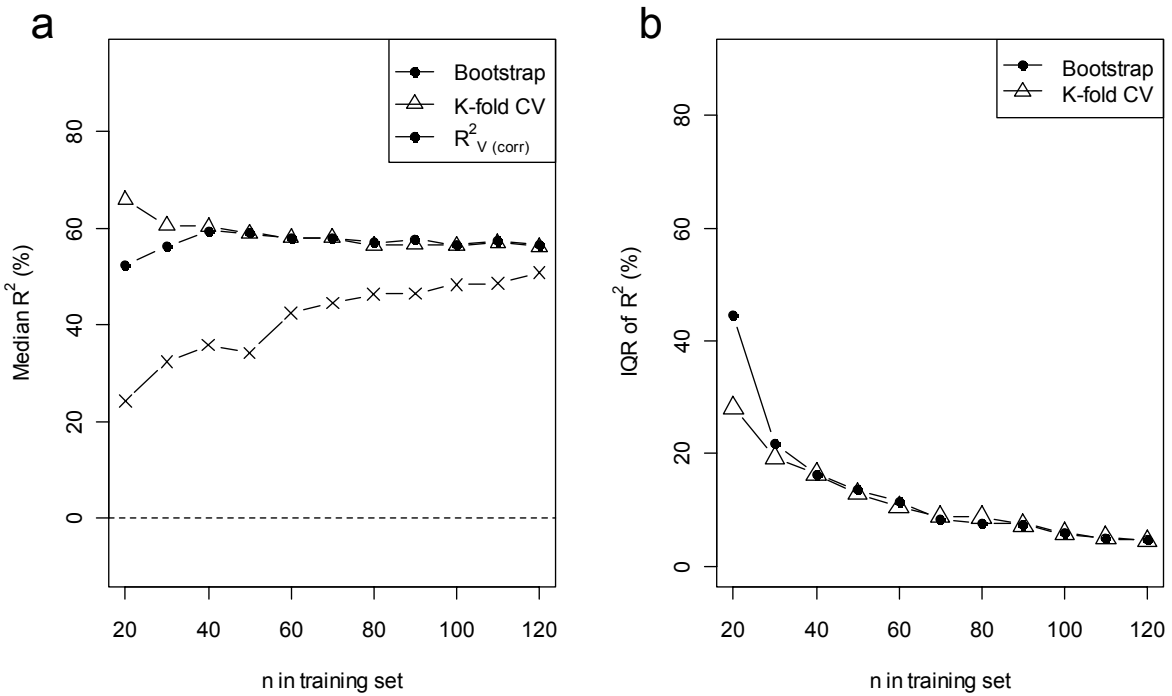
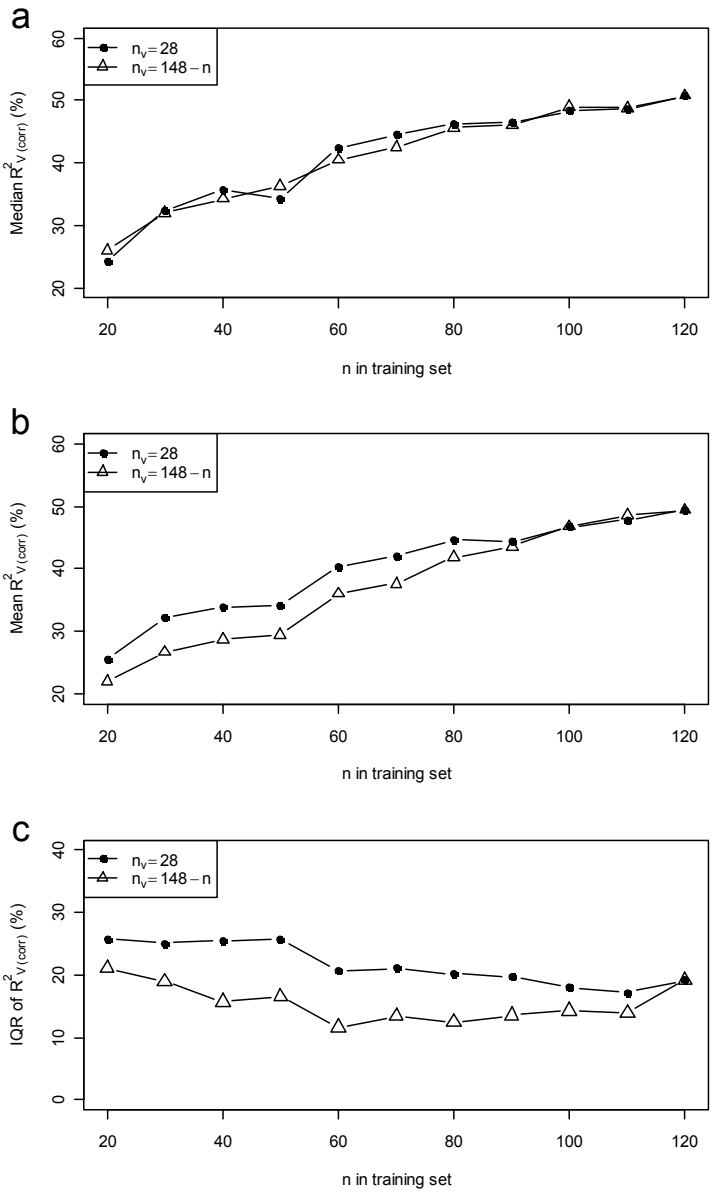


Figure S2. Comparison of the median (panel a), mean (panel b) and interquartile range (IQR) (panel c) of the $R^2_{V\,(corr)}$ obtained using different sizes of the validation datasets. In one case, n_v=28 for all values of n, while in the other, n_v uses all the observations not used in the training dataset (i.e. n_v=148-n).



8.4 Long-term exposure to traffic-related air pollution and subclinical atherosclerosis (Paper IV)

Rivera, M., Basagaña, X., Aguilera, I., Foraster, M., Agis, D., de Groot, E., Perez, L., Mendez, M., Bouso, L., Ramos, R., Sala, J., Marrugat, J., Elosua, R., Künzli, N.

Paper submitted. *

* *This paper is reproduced according to the version submitted. References of this paper are included in Section 10.*

LONG-TERM EXPOSURE TO TRAFFIC RELATED AIR POLLUTION AND SUBCLINICAL ATHEROSCLEROSIS

Authors:

Marcela Rivera[a,b,c,d], Xavier Basagaña[a,b], Inmaculada Aguilera[a,b,d], Maria Foraster[a,b,c,d],David Agis[a,b,d], Eric de Groot[e], Laura Perez[f,g], Michelle Mendez[h], Laura Bouso[a,b,d], Rafael Ramos[i,i], Jaume Sala[k], Jaume Marrugat[l], Roberto Elosua[l], Nino Künzli[f,g]

Authors' affiliations:

[a] Centre for Research in Environmental Epidemiology CREAL, Barcelona, Spain
[b] Hospital del Mar Research Institute (IMIM), Barcelona, Spain
[c] Universitat Pompeu Fabra, Barcelona, Spain
[d] CIBER Epidemiología y Salud Pública (CIBERESP), Spain
[e] Department of Vascular Medicine, Academic Medical Centre, Amsterdam, The Netherlands
[f] Swiss Tropical and Public Health Institute, Basel, Switzerland
[g] University of Basel, Basel, Switzerland
[h] University of North Carolina at Chapel Hill, North Carolima, USA
[i] Research Unit, Family Medicine, Jordi Gol Institute for Primary Care Research (IDIAP Jordi Gol), Catalan Institute of Health, Catalunya, Spain
[i] Department of Medical Sciences, School of Medicine, University of Girona, Spain
[k] Servicio de Cardiología, Hospital Universitari Josep Trueta, Institut Català de la Salut, Girona, Spain
[l] Grupo de Epidemiología y Genética Cardiovascular, Hospital del Mar Research Institute (IMIM), Barcelona, Spain

Corresponding author:

Marcela Rivera, CREAL - Centre for Research in Environmental Epidemiology. Barcelona Biomedical Research Park. Dr. Aiguader, 88. 08003 Barcelona. Spain. Tel. +34 932 147 342. Fax + 34 93 3 160 575.
marcelaepi@gmail.com; marcelariverav@gmail.com

Abstract

Epidemiological evidence of the effects of long-term exposure to air pollution on the chronic processes of atherogenesis is limited. We investigated the association of long-term exposure to traffic-related air-pollution with subclinical atherosclerosis.

We performed a cross-sectional analysis using data of the re-examination (2007-2010) of 2780 participants of the REGICOR study, a population-based prospective cohort in Girona, Spain. Long-term exposure was calculated as the time-weighed average of residential NO_2 estimates, traffic intensity in the nearest street and traffic intensity in a 100m buffer across residences in the last 10 years. The main health outcome was the carotid artery intima media thickness (IMT), a marker of subclinical atherosclerosis. The association between exposure and IMT was assessed using multivariate linear regressions controlling for individual level confounders.

An increase of 25μg/m^3 in NO_2 was weakly associated with IMT (percent change in IMT 0.56%, 95%CI: -1.47 to 2.59%) and reached statistical significance in people with high education level (4.6%, 95%CI: 0.4 to 8.9%) and men above 60 years (4.3%, 95%CI: 0.2 to 8.4%). An increase of 15000veh/day in traffic in the nearest street was associated with a 2.32% increase in IMT (95%CI: 0.48 to 4.17%) and an increase of 7 200 000veh-m/day in traffic load in 100m was associated with a 1.91% increase in IMT (95%CI: -0.24 to 4.06). Stronger associations between all exposure markers and IMT were observed in the subgroups of high education level and men above 60 years.

Long-term exposure to traffic-related air pollution is associated with subclinical carotid atherosclerosis in healthy individuals.

Background

Air pollution from traffic and other sources is an established cause of premature mortality (Brook et al., 2010). A relevant part of this environmental burden of disease relates to cardiovascular diseases (CVD), which were responsible for 10% of total Disability Adjusted Life Years in 2004 and the leading causes of death worldwide in 2008 (World Health Organization, 2008). The common basis of this group of diseases is atherosclerosis, a chronic and degenerative process that mainly occurs in large and medium-sized arteries morphologically characterized by asymmetric focal thickenings of the innermost layer of the artery, the intima . The development of atherosclerosis is the result of the total cumulative exposure to atherogenic risk factors such as hypertension, high cholesterol, diabetes, obesity, smoking, physical inactivity and other lifestyle factors and its interactions with genetic susceptibility (Lusis, 2000). Acute events such as myocardial infarction or stroke can be triggered by short-term exposure to air pollution. However, whether and how ambient air pollution contributes to atherogenesis is subject to debate. While experimental studies on animals provide strong evidence for a causal atherogenic role of air pollution, in particular in obese mice (Sun et al., 2005), the number of epidemiological studies is limited.

The long induction period of atherosclerosis makes the study of subclinical disease feasible. The association between air pollution and intima media thickness (IMT), an established marker of subclinical atherosclerosis, was reported for the first time in volunteers participating in two clinical trials in California (Künzli et al., 2005a). Two population-based cross-sectional analyses, namely the RECALL study in Germany (Bauer et al., 2010) and the MESA Cohort in the United States (Diez Roux et al., 2008) confirmed these results while one study in young adults in the Netherlands found no association (Lenters et al., 2010a). The observed thicker artery walls in highway toll station workers in Turkey may be interpreted as an effect of high exposure to traffic-related pollutants (Erdogmus, 2006). So far only one longitudinal study has been published, based on heterogeneous samples of volunteers participating in five clinical trials (Künzli et al., 2010). The study suggested a possible role of ambient air pollutants, indicated by $PM_{2.5}$ and living close to busy highways, in the progression of IMT.

Most of these studies showed that the impact of air pollution differed across various subgroups; for example, some observed stronger effects in women and the elderly (Hoffmann et al., 2009b; Künzli et al., 2010). However, these patterns were not consistent, thus, a clear understanding of susceptibility factors is still lacking. Although observed in animal studies (Sun et al., 2005), none of the human studies investigated whether diet modifies the effect of air pollution – a plausible hypothesis given its interaction with pathways of oxidative stress and systemic inflammation – both likely involved in the systemic effects of ambient air pollution (Brook et al., 2010).

The Mediterranean region of Girona, and Spain in general, is known to have one of the lowest cardiovascular mortalities in Europe, while the conventional cardiovascular risk factors are among the highest (Masiá et al., 1998). It has been shown though that the paradox may in part be explained by the protective effect of Mediterranean diet (Guallar-Castillón et al., 2011; Martínez-González et al., 2011; Zhu et al., 2011). Our study aimed to investigate the association between long-term exposure to traffic-related air pollution and subclinical atherosclerosis in Spain and the effect modification of such association by demographic characteristics, established risks factors of atherosclerosis and diet. We investigated this in the follow-up of participants of three population-based cohorts recruited in the REGICOR study (Grau et al., 2008). Subclinical atherosclerosis was measured by carotid intima media thickness (IMT), a validated marker of atherosclerosis (Bots et al., 2002; Coll et al., 2008). Additionally we measured ankle brachial index (ABI), a marker of the presence and severity of peripheral artery disease. Both IMT and ABI are associated with cardiovascular events and mortality (Ankle Brachial Index Collaboration, 2008; Lorenz et al., 2007; Polak, 2009; Ramos et al., 2009). In the absence of particulate matter measurements in this region, we used estimates of the 10-year average home outdoor nitrogen dioxide (NO_2) concentrations and residential traffic intensity as markers of exposure to local traffic-related air pollutants.

Methods

Study design

REGICOR-Air is a cross-sectional study nested in the REGICOR (REgistre GIroni del COR, Girona Heart Registry) cohort study (Grau

et al., 2007). We used data from the follow-up (in 2007-2010) of three population-based cohorts of REGICOR originally enrolled in 1995, 2000 and 2005. A two-stage framework was used in the first two cohorts: in the first stage a fixed number of towns of Girona province were randomly selected: 33 towns in 1995 and 17 towns in 2000. In the second stage an equal number of men and women between 25 and 74 years were randomly recruited from each selected town. For the third cohort, residents of Girona city were randomly selected. The response rate was > 71% in the three surveys. For the REGICOR-Air study, 12 of the original towns that provided the maximum contrast in ambient air pollution levels, while keeping the travel distance between towns to a minimum, were selected (see map in Supplement **Figure S1** and Girona.KLZ file in **Appendix 2**). These towns are spread over an area of approximately 65x70 km and represent the geographic diversity of the Girona Province, including settings like the average mid-size Mediterranean city of Girona and adjacent Salt (96 200 and 30 300 inhabitants, respectively), tourism getaways on the beach, industrial centers, maritime ports, historic quarters and small villages in the valleys and mountains. During 2007-2010, the participants residing in these towns, who were alive and not institutionalized, were invited to participate in REGICOR-Air. Response rate for the re-examination was approx. 82%. Information for the complete reassessment of the health status, measurements of IMT and ABI and address history linked to time period during the last 10 years were collected.

This study was approved by the Hospital del Mar Research Institute ethics committee and participants gave written informed consent.

Health and life-style measurements

The main outcomes were: a) the mean IMT of left and right common carotid artery (**IMTcca**) and b) the mean IMT of left and right common carotid artery, internal carotid artery and carotid bulb (**IMT6seg**). Each measurement was the average thickness in a 1cm segment of the artery far wall.

For assessment of carotid intima-media thickness (IMT) three trained and certified sonographers performed ultrasound examinations of the carotid arteries. Standardized scan and image analyses protocols were used. For imaging, an Acuson Aspen ultrasound

instrument (Acuson-Siemens, Erlangen, Germany) equipped with an L7 5-12MHz transducer and dedicated REGICOR application scan protocol (AMC IMAGELAB, University of Amsterdam/Technical University Eindhoven, The Netherlands) were used. Of each arterial segment a still image was saved as a DICOM (Digital Information and Communication in Medicine) file. These source files were locally stored and securely transferred to the IMAGELAB, where trained and certified sonographers analyzed the images using validated software eTrack REGICOR (Department of Physiology and Vascular Medicine Academic Medical Centre, Amsterdam) (de Groot et al., 2008). The IMT was defined as the average distance between the lumen-intima and media-adventitia interfaces of a given segment. To assess repeatability two measurements of the IMT were taken in 42 participants at two occasions with two weeks between the visits. In each visit, up to three sonographers measured IMT in each participant. Between-sonographer and between–visit variability were measured.

ABI measurements were done by operators meticulously trained by a senior vascular surgeon. After a 5-minute rest, systolic blood pressure was measured in the brachial artery in the antecubital fossa of both arms, and subsequently in the posterior tibial and dorsalis pedis arteries in the distal calf of both legs in supine position. A continuous Doppler device (SONICAID 421, Oxford Instruments), 8MHz probe was used. Right and left ABI were calculated as the ratio of the highest systolic pressure in each lower limb to the highest of left and right brachial systolic pressure. The lowest resulting ABI value was categorized (low: <0.9; medium: 0.9-1.3; high: >1.3) for analysis. Categories were selected according to reported mortality risks associated to ABI level (Ankle Brachial Index Collaboration, 2008; McDermott et al., 2005). Repeatability was also assessed by independent measurements.

In addition, we measured fasting lipid profile and glycaemia, blood pressure and anthropometric characteristics by standard protocols. Hypertension was defined as either systolic blood pressure >140mmHg, diastolic blood pressure >90 mmHg or taking antihypertensive medication. Hypercholesterolemia was defined as having total cholesterol levels above 250 mg/dl or taking lipid-lowering medication. Weekly energy expenditure in the leisure-time was measured with the Minnesota questionnaire (Elosua et al., 1994), converted to metabolic equivalents and categorized in tertiles. Adherence to Mediterranean diet was measured by a 10-

point index based on sex-specific intake tertiles of 8 beneficial (cereals, fruits, vegetables, legumes, seafood, nuts, moderate red wine) and 2 detrimental (meats, dairy products) food groups, and categorized in quartiles, as described in previous studies in these cohorts (Schröder et al., 2004). Modifications such as eliminating low-fat dairy products or white meats, or incorporating several unhealthy food groups (e.g. soft drinks, salty snacks, pastries) as detrimental components had no meaningful impact on findings in sensitivity analyses (not shown). The plausibility of reported dietary intakes was assessed based on disparities between reported energy intakes and estimated energy requirements calculated using the Goldberg method modified with updated basal metabolic rate estimates (Mendez et al., 2011). Highest achieved education level (low refers to illiterate or primary school, medium to secondary school and high to college, university or higher degree), occupational status (employed, inactive, retired or unemployed), smoking habits (never, former, current smoker), marital status (single, married or equivalent, divorced, widow, other) were collected in a questionnaire administered during the reassessment visit. Participants also reported any current medication treatment, which was later checked and coded by a physician into main therapeutic groups defined in the REGICOR study protocol. Percentage of people with low education at the census tract level was available from the 2001 Spanish census. We assigned this area-level variable to each participant according to the census tract where they had lived the longest during the 10-year period before the IMT measurement.

Clinical history of CVD was defined as having had myocardial infarction, stroke, angina, catheterization, angioplasty, bypass surgery or amputation due to circulatory problems. Participants with clinical signs of CVD were also excluded (n=227) since medication use or altered health behaviors among these subjects may have influenced IMT/ABI measures obtained for this study.

Address history

Ten-year residential history was collected by questionnaires administered at the time of IMT measurement. Addresses were geocoded at the front door level by a Geographic Information System (GIS) services company. Addresses geocoded with less than maximum precision level (i.e. exact address) were manually

geocoded using the web mapping application of the Cartographic Institute of Catalonia (http://mercuri.icc.cat).

Exposure Assessment

We assigned exposure to each participant estimating the 10-year time-weighted average of the home outdoor concentrations of NO_2 using land use regression models. In the absence of air quality data in most of these towns, we conducted an extensive monitoring campaign based on NO_2 passive samplers (Palmes tubes) used and validated before in Girona and internationally (Targa J. et al., 2008).

A previous study conducted with the same Palmes tubes in the town of Girona indicated surprisingly high levels of NO_2 in some narrow street canyons with rather low traffic density. Thus, we decided to take measurements at a far higher number of locations than those suggested in established protocols developed for non-Mediterranean cities (Brauer et al., 2003), where models are based on repeated measurements taken on some 40-50 locations, but limit the seasonal repeats. The latter decision was based on the observations from previous monitoring data from local authorities and our own pilot study measurements indicating that monthly mean NO_2 levels during spring (mainly April and May) and fall (mainly October and November) approximated the annual mean. Thus, we conducted our largest campaigns during these months.

We measured NO_2 in the balcony of 562 participants' homes for one month, in spring and fall campaigns between June 2007 and July 2009. Homes were selected to cover a broad range of traffic-related pollution (based on mentioned previous study) and urban settings (e.g. low and high building-density areas), to be representative of the residential locations of the cohort participants and to be well distributed across the towns. The added surface area covered by sample monitoring across towns was approx. 81km^2. Since all measurements were not made simultaneously, we adjusted for the temporal variability of the NO_2. To do so, we also measured monthly mean NO_2 concentrations at one fixed location in each town for at least one year, including simultaneous measurements during the campaigns. NO_2 annual means were derived multiplying the monthly means at each location by the ratio of the annual to the same month mean NO_2 at the town's fixed location.

To predict NO_2 at each participant's residence we used land use regression (LUR) models. LUR models were developed using the NO_2 annual means and data on traffic intensity, bus routes and stops, distance to traffic, land cover, building density, and other GIS-derived variables. Given the geographic diversity of the study area and differences in the availability of GIS data among towns, we divided the study area into seven subareas, i.e. clusters of adjacent towns. LUR models were derived for each cluster. The cluster-specific models explained between 33 and 63% of NO_2 (cross-validation R^2 0.32 and 0.61 respectively) (Supplement **Table S1**).

We estimated the outdoor annual mean NO_2 at each residential location by applying the LUR models to the address geocode. Historical data on NO_2 concentrations since 1997 were available from only one urban background fixed station, located 40km south east from Girona city. Long-term temporal trends in NO_2 levels were taken into account by multiplying the NO_2 derived for each address by the ratio of the mean NO_2 for the period living in that address to the mean NO_2 for the monitoring campaigns period at the urban background station. Finally, for each participant we calculated the time-weighed average of NO_2 estimates across all residences in the 10 years prior to the IMT measurement (10yr NO_2 exposure). The time periods when participants lived at addresses geocoded with low precision or outside of the study area were not taken into account to derive the 10yr NO_2 exposure. Participants who lived in the study towns less than 6 years or had maximum precision geocodes for fewer than 6 years of the address history (*n*=365) were excluded from the main analyses.

We also used traffic proximity markers as surrogates of air pollution exposure in independent analyses. Traffic intensity was collected from local registries and urban mobility reports (Urban mobility plan for Girona, 2007) and was assigned to the central road network used within ESCAPE (www.escapeproject.eu). Traffic counts were conducted at approximately 670 streets with missing traffic information (Rivera et al. 2012) to complete the traffic inventory. For each address the traffic intensity at the nearest street and traffic load (sum of traffic intensity multiplied by length of road segment in all segments) in a 100m buffer were calculated. The 10-year average traffic intensity and 10-year average traffic load in 100m were then derived for each participant.

All GIS calculations were done using ArcGIS 9.2 (ESRI, Redlands, CA).

Statistical Analysis

The crude and adjusted association of IMT with individually-assigned air pollution exposure was assessed with linear regression models. The residuals of the regression analyses showed heteroscedasticity, thus we used the natural logarithm of IMT as the outcome variable. The linearity of the associations was explored using generalized additive models (GAM). We analyzed ABI as a categorical variable using multinomial logistic regression. We initially adjusted our models by age and sex. The final models included two sets of adjustment variables. Using directed acyclic diagrams (**Figure 1**) that represent the hypothetical relationship between long-term exposure to air pollution and IMT, according to epidemiological evidence (Brook et al., 2010; Künzli et al., 2005a; Künzli et al., 2011), we defined a minimal adjustment set (Model 1), which included: sex, age, sex-age interaction, smoking status, education level (as a proxy of socioeconomic status) and marital status.

We further adjusted for potential intermediates and other covariates (Model 2): Model 1 plus occupational status, body mass index (BMI), high-density lipoprotein (HDL), waist circumference, systolic and diastolic blood pressure, weekly energy expenditure in physical activity during leisure-time, adherence to Mediterranean diet score, plausibility of reported dietary intakes, any cardiovascular or antihypertensive medication treatment (defined as any of beta blockers, calcium antagonists, aspirin to prevent heart disease, cumarinics, statins, antiaggregants or any other treatment for hypercholesterolemia or hypertension) and percentage of people with low education at the census tract level. In models for traffic load, outcomes were also adjusted for occupational status.

The results are expressed as the cross-sectional percent change in IMT associated with a 10-year exposure contrast corresponding to the difference between the 95th and 5th percentiles in the study population.

We tested heterogeneity of the effects of air pollution by including interaction terms in the covariate adjusted models (Model 2). Potential modifiers tested were sex, age, education level, smoking

status, menopausal status in women, obesity (bmi≥30), hypertension, diabetes, any cardiovascular or antihypertensive medication treatment and adherence to Mediterranean diet.

Analyses were done using Stata 10.1 (StataCorp, College Station, TX) and R 2.12 (http://www.R-project.org).

Sensitivity analyses

We stratified the analysis by movers and non-movers (participants who lived in the same address for the 10 years before the IMT measurement). Errors in exposure assignment and geographic co-variates are likely reduced in the latter, thus, effects may be expected to be less biased. We also tested whether the results using exposures at the address of longest residence differed much from those using the average for the complete 10-year address history. To assess the sensitivity of the results to the NO_2 LUR models, instead of using the modeled NO_2 estimates, we restricted the analyses to participants living within 100m and within 200m of a monitoring location and assigned them the annual mean NO_2 at the monitor closest to the address of longest residence. We did not adjust the exposure-outcome associations for area of residence in the main analyses since the study area was relatively small (65x70Km), data collection in every town was done by the same team and using the exact same procedures and it would partially remove the exposure contrast corresponding to between-town variability. We then explored the sensitivity of the results to the inclusion of area of residence (corresponding to the address of longest residence) as a random effect variable using the *xtmixed* function of Stata. Some recent studies suggest that the association of air pollution and cardiovascular health could be confounded by traffic noise. Traffic noise data was only available for a subsample of participants living in Girona city (*n*=1084). Thus, in this subsample we evaluated the association of IMT with traffic noise exposure. Night-time (11pm – 7 am) road traffic noise levels (dB(A)) at the building's façade were estimated for the address of longest residence with a validated city-specific noise model (Environmental Noise Directive 2002/49/EC) (Foraster et al., 2011).

Results

Information on IMTcca, IMT6seg and ABI was available for 2780, 2188 and 2738 participants respectively. The characteristics of the study populations are summarized in **Table 1**. Participants were 32 to 86 years old. Percentages of the study population with low, medium and high education reflect those of Spain (census 2001). Participants included in the analyses did not differ from those excluded in terms of personal characteristics and exposure levels. The median IMTcca was 0.68mm and it ranged between 0.40 and 2.05mm. The repeatability study showed intraclass correlation coefficients for sonographers and visits of 0.83 for the IMTcca and of 0.77 for the IMT6seg. ABI was on average 1.10 and ranged between 0.5 and 1.75, with 2.0% of the study population with low ABI (<0.9) and 4.2% with high ABI (>1.3). Participants with low ABI were on average nine years older than participants with normal ABI. Participants with high ABI did not differ on age with participants with normal ABI. Inter- and intra-operator variability of ABI measurements were also low, with an intraclass correlation coefficient of 0.92 and 0.94, respectively.

The 10-year average home outdoor nitrogen dioxide concentrations varied from 5 to 48 ug/m^3 and its correlation with NO_2 at the address of longest residence was >0.99 over all subjects, and 0.96 among those who moved at least once during the 10 years. A broad variability was also observed in the distribution of the time-weighted average traffic in the nearest street and in a 100m buffer around the addresses reported for the 10-year period (**Table 2**). Nitrogen dioxide and traffic exposure variables were moderately correlated (correlation coefficients between 0.52 and 0.72), and the two traffic variables (intensity and load) had a correlation of 0.58 (**Supplement Table S3**).

In univariate analyses, both 10-year averaged NO_2 and NO_2 at the address of longest residence were directly and significantly associated with age, HDL, higher education level, being divorced, being a former smoker, having high adherence to Mediterranean diet, and inversely associated with BMI, blood pressure and percentage of low education at the census tract level. Traffic load in 100m buffer was associated with the same variables as NO_2 (same direction of association) except for smoking status (no association). Additionally, exposure to higher traffic load was associated with

being unemployed. Finally, traffic in the nearest street was associated with the same variables as NO_2 (same direction of association) except for smoking status, marital status and diet (no association).

The associations of log-transformed IMT with exposures were linear. In unadjusted models, nitrogen dioxide, traffic in the nearest street and traffic in the 100m buffer were directly and strongly associated with the intima media thickness (both IMTcca and IMT6seg) (**Table 3**). Associations decreased in the adjusted models, particularly after adjusting for age. The two sets of adjustment variables provided similar results. The positive association of 10-year averaged NO_2 with carotid intima media thickness (both IMTcca and IMT6seg) was smaller and no longer significant. A difference in exposure between the 5th and 95th percentiles of 7 200 000 veh m/day in traffic load in 100m was directly associated with a nonsignificant 1.91% increase in IMTcca (95%CI: -0.24 to 4.06) and with a 2.06% increase in IMT6seg (95%CI: -0.09 to 4.21). A difference of 15000veh/day in traffic in the nearest street was associated with a 2.32% increase in IMTcca (95%CI: 0.48 to 4.17%) and a 1.80% increase in IMT6seg (95%CI: 0.01 to 3.59%).

Both 10-year NO_2 and residential traffic were associated with increased risk of high ABI (see **Table 4**). The adjusted relative risk ratio (RRR) of high ABI for a difference of 25µg/m^3 in NO_2 was 1.98, 95%CI: 1.09 to 3.60; for a difference of 7 200 000veh/day in traffic load in 100m was 1.89, 95%CI: 1.07 to 3.34; and for a difference in 15 000veh/day in the nearest street 1.70, 95%CI: 1.13 to 2.57, as compared to persons with medium ABI. The RRR of low ABI compared to medium ABI were non-significant for all exposures.

Effect modification

The association of all exposure markers with IMT differed across education level (**Figure 2**). In people with higher education, the association of 10-year exposure to air pollution with IMT was stronger, with a p-value for interaction of 0.07 (Model 2 with NO_2 and IMTcca). Effects estimates of the associations were 4.6% (95% CI: 0.4 to 8.9%) for NO_2; 4.8% (95% CI: 0.7 to 8.9%) for traffic load; 3.3% (95% CI: -0.02 to 6.7%) for traffic in the nearest street). Similar patterns were observed for the associations of

exposure to air pollutants with IMT6seg. Effects estimates were also stronger in men older than 60 years, showing significant associations of NO_2 and traffic load in 100m with IMT (percent change in IMTcca 4.3% (95% CI: 0.2 to 8.4%) for NO_2; 5.9% (95% CI: 1.6 to 10.3%) for traffic load; 3.4% (95% CI: -0.07 to 7.0%) for traffic in the nearest street). No significant differences were observed in the associations of exposures with IMT by groups of age, sex, smoking status, menopause, obesity (bmi≥30), hypertension, diabetes, any cardiovascular or antihypertensive medication treatment or adherence to Mediterranean diet (results for menopause, obesity, hypertension diabetes and diet not shown).

Sensitivity analyses

Restricting the analyses to non-movers did not influence the results except in the precision of the effect estimates. Using the exposure estimated at the address of longest residence instead of the 10-year average exposures or the NO_2 unadjusted for the10-year temporal trends did not affect the results. Using the exposure estimated at the current address at the time of examination the effect estimate and its precision decreased. E.g. Excluding 57 participants who did not live in the study area at the time of examination, the effect of NO_2 on IMT for the people with high education was 4.56, 95% CI: 0.32 to 8.81 using 10-year average NO_2, 4.33, 95% CI: 0.41 to 8.25 using NO_2 at the address of longest residence and 3.54, 95% CI: -0.44 to 7.51 using NO_2 at current residence. Similar results were observed for residential traffic.

Using annual mean NO_2 at the closest monitor instead of 10-year modeled NO_2 restricted the sample size because only those living close to a monitor were kept in this analysis. There were 2265 and 1778 participants living within 200m of a monitoring location and with data available on IMTcca and IMT6seg, respectively. Median distance to the closest monitor was 90m. As expected, modelled concentrations included fewer observations at the extremes of the distribution (see Supplement **Figure S2**). All associations were in the same direction with effect estimates for IMTcca being approximately 0.6% lower and less precise as compared to the estimates based on modeled NO_2 done in the same subsample.

Including area of residence as a random effect variable, the trend of the associations remained the same but the estimates were closer to the null and less precise, e.g. 1.93% change in IMTcca 95%CI: 0.05 to 3.8 (Model 2 with traffic on the nearest street). Entering night-time traffic noise to the models of IMT with NO_2 did not influence the results.

Discussion

Our multivariate models revealed positive correlations between three markers of long-term exposure to traffic-related air pollution, namely traffic load within 100m of the residence, traffic intensity in the nearest street and modeled concentration of home outdoor NO_2 concentrations with carotid subclinical atherosclerosis in a random healthy population sample of the Spanish Mediterranean region. However, associations with NO_2 were weak and reached statistical significance only in the subgroups of people with high education level and men above 60 years. Residential traffic (load and intensity) was associated with 2% thicker carotid intima-media. High exposures to NO_2 and traffic were also associated with increased risk of high ABI (RRR between 1.70 for traffic intensity in nearest street and 1.98 for NO_2). To put these results in context with common risks factors for atherosclerosis, in the same study population included in main analyses, a 10 year difference in age was associated with 8% (95%CI: 7.3 to 9.0) thicker carotid walls (IMTcca) and a RRR of 10.3 (95%CI: 10.0 to 10.5) for high ABI. Interestingly, age strongly confounded the association of pollution with IMT. Atherosclerosis is a life-time process, thus, age is a strong correlate of the cumulated life-time exposure to air pollution as well as the main determinant of atherosclerosis.

Our results indicate that long-term exposure to traffic-emitted pollution might induce atherosclerosis at the carotid artery level. This is in line with prior results indicating the association of medium- to long-term exposure with subclinical atherosclerosis (Brauer et al., 2003; Diez Roux et al., 2008; Hoffmann et al., 2007; Künzli et al., 2005a) and systemic inflammatory markers (Hoffmann et al., 2009a). Direct comparison of our results with published data is limited given our use of NO_2 as a marker of pollution whereas others used measures of PM mass. However, our effect estimates are in the same order of magnitude like in previous studies.

In contrast, the association between pollution and ABI was less clear. While pollution was associated with having high ABI (>1.3), the risk of a low ABI was not associated with pollution. The segment of ABI that is more clearly associated to higher mortality and cardiovascular risks is ABI<0.9. ABI higher than 1.3 have been less studied. It has been associated to calcification of the arterial wall, higher levels of many CVD risk factors (McDermott et al., 2005), higher risk of all-cause mortality (Ankle Brachial Index Collaboration, 2008), and foot ulcers and weakly associated with heart failure and stroke (Allison et al., 2008). Thus, there is higher cardiovascular risk at low and high levels of ABI. This u-shape association with cardiovascular risk makes ABI a complex marker and makes comparison between studies difficult.

Regarding the association of low ABI with air pollution we can refer to only two studies, providing conflicting results. A study in Germany found that living at 50m from a major road compared to living more than 200m away was associated with a decrease in ABI (modeled as continuous variable) of -0.024 (95%CI: -0.047 to -0.001) and with an OR of 1.77 for peripheral artery disease (95%CI: 1.01 to 2.1), while no associations were found with annual residential $PM_{2.5}$; in this study participants with ABI higher than 1.3 were excluded (Hoffmann et al., 2009b). On the other hand, in the MESA study, 20-year exposures to $PM_{2.5}$ and PM_{10} were not associated with ABI, while 1-year PM_{10} exposures were associated with higher ABI (modeled as continuous and dichotomous variable ABI≤0.9); participants with ABI higher than 1.3 were included in this analysis (Diez Roux et al., 2008). Smoking could be considered a proxy for high exposure to combustion related pollution. Strong direct associations of current smoking with ABI<0.9 have been reported in several studies (Agarwal, 2009; Hobbs et al., 2005; Ramos et al., 2009). No association between passive smoking (self-reported nonsmokers with serum cotinine levels above 0.05ng/ml) and low ABI was found in a population-based study in the US (Agarwal, 2009). Low ABI was associated with cotinine levels above 155mg/ml (threshold). In another population-based study in women older than 60 years in China, passive smoking was associated with ABI<0.9 (OR 1.47, 95%CI: 1.07 to 2.03) and a dose response relationship was found (He et al., 2008).. Exposure levels in China were higher than in the US. This might indicate that effects on low ABI could be detectable only after high exposures to air pollution (a threshold effect for pollution is unlikely). The lack of association between air pollution and low ABI in our study might be due to that in addition to

the small number of participants with low ABI and the resultant low power to detect an association.

To our knowledge, the association of air pollution with high ABI has not been investigated before. Current and former smoking have been negatively associated to high ABI (Allison et al., 2008; McDermott et al., 2005; Wattanakit et al., 2007) compared to normal ABI. In our study both were associated to lower risks of high ABI although for current smokers this risk was not significant. ABI is a ratio of systolic blood pressures, thus higher ABI could result from low brachial or high ankle pressures or both (Allison et al., 2008). While our results need confirmation from future studies, they lead to speculate that the atherosclerotic response to varying levels of air pollution differs between the carotid arteries and the peripheral arteries. This hypothesis would also be supported by our results that high ABI was not associated to IMT compared to normal ABI levels, while low ABI was associated with higher RRR of thicker IMT.

In contrast to our findings, the effects of 20-year exposure to $PM_{2.5}$ and PM_{10}, in the MESA study, were weakly associated with carotid IMT (1 to 3% increases in IMT for increases of 12.5 ug/m3 in 20-year average $PM_{2.5}$ or 21ug/m3 in 20-year average PM_{10}) but not with ABI, or coronary artery calcification (Diez-Roux 2008). In Los Angeles, a 10µg/m^3 increase in modeled $PM_{2.5}$ assigned to the current address only, was associated with a 4.2% (95%CI: -0.2 to 8.9) increase in IMT (Künzli et al., 2005a). Positive, though not significant, associations between traffic intensity and IMT, but no association between NO_2 and IMT, were found in a population-based study of young adults in Utrecht. A 17 000veh/day increase in traffic in the nearest road was associated with a 0.47% increase in IMT (95%CI: -1.79 to 2.73) and an increase in 472 000veh/day in the sum of traffic intensity in a 100m buffer with a 1.04% increase in IMT (95%CI: -0.95 to 3.03) (Lenters, 2010).

Given the worldwide rising burden of atherosclerotic vascular diseases (World Health Organization, 2008), the association of long-term pollution with atherosclerosis has important public health implications. Ninety five percent of our study population was exposed to traffic intensity levels in the nearest street higher than 15 000veh/day. However, these levels are low when compared to other European cities. In larger cities, such as Barcelona, many people live within very short distances to streets with traffic intensity

of about 50 000 to 100 000veh/day. Likewise, levels of NO_2 can exceed 100µg/m^3 near streets with high traffic intensity.

Our population-based study used exposure measurements (both NO_2 and traffic) collected outdoors of participant's residences. Similar studies rely on models based on measurements at routine air quality monitoring stations as far away as 10km or more from participants' residences (Diez Roux et al., 2008). Besides, our traffic markers were based not only on proximity and road type classification, which probably lead to larger exposure misclassification (Allen et al., 2009; Hoffmann et al., 2009a), but rather on actual traffic intensity derived from a dense traffic-count network.

Reasons that may have led to not detecting a significant association of IMT with NO_2 for the overall population, if the association exists, are exposure misclassification and NO_2 not being a good proxy of the agents in the pollutant mixture that actually promote atherosclerosis. Moderate R^2s of the NO_2 models introduce exposure measurement error. Reasons that led to the variability of LUR's R^2 across towns are not known. Possible explanations include that available geo-spatial variables did not fully capture the complex air quality conditions in street canyons or the local wind fields that determine dispersion. Nevertheless, using the measured (at the closest monitor) instead of modeled NO_2 did not influence the results, thus the measurement error introduced by the models, if any, should be low.

Although findings using traffic as a surrogate of exposure appear to be more consistent across studies than those based on pollutants, more research is needed to clarify the role of different constituents of the air pollution's complex mixture on atherogenesis. In the referred study in Utrecht, the associations of NO_2, black smoke, $PM_{2.5}$ and SO_2 with the IMT, pulse wave velocity and augmentation index (the last two are markers of arterial stiffness) were investigated. Significant associations were found only for NO_2 with pulse wave velocity and augmentation index and for SO_2 with augmentation index. The evidence on the effects of NO_2 is inconsistent. It has been associated with stronger effects on cardiovascular mortality compared to $PM_{2.5,}$ black smoke and SO_2 (Beelen et al., 2007b) and on cardiopulmonary mortality compared to PM_{10}, total suspended particles, black smoke and SO_2 (Gehring et al., 2006). While Pope et.al (2002b) found no association

between NO_2 and cardiopulmonary mortality. Thus, as it happens for the correlation of NO_2 with traffic intensity (Raaschou-Nielsen et al., 2000), the correlation of NO_2 with the components of the traffic emissions cocktail that are responsible for promoting atherosclerosis, might vary across locations.

Effect modification

We tested effect modification by a pre-determined set of variables, based on differences in the atherosclerosis progression given by sex, age, menopause, etc, in addition to effect modifiers found in previous studies (see **Supplement Table S2**), Stronger positive effects in people with high education level and in men older than 60 years were consistent across IMT measurements, IMTcca and IMT6seg, and across all markers of pollution. No evidence of effect modification by age, sex, menopause, Mediterranean diet, common risks for CVD or subclinical disease (indicated by medication treatment) was observed. Other studies have found heterogeneous effects across subgroups of age, sex, BMI, smoking status, socioeconomic status, town of residence and other cardiovascular risks factors (Bauer et al., 2010; Künzli et al., 2005a; Lenters et al., 2010a). The detection of interactions in epidemiologic studies is often underpowered (Greenland, 1993) and testing many interactions can also lead to a multiple comparison problem. However, our subgroup analyses were based on the results of previous studies, and although our results may be subject to the aforementioned problems, the accumulated evidence over different studies will help to shed light on the very important question of susceptible subgroups.
In Girona, exposure to traffic-related pollutants (all markers) was higher for people with high education level. The same was observed at the census tract level: higher mean concentrations of NO_2 were found at the most privileged census tracts (see **Supplement Figure S4**). This has been reported before for other South European cities (Cesaroni et al., 2010) where wealthy people live in downtown areas that are more polluted. In Strasbourg (central Europe), Havard et al. (2009) observed an inverted-U shaped association between annual NO_2 at the census tract and deprivation index at the census tract level. These two patterns are in sharp contrast to what has been observed in North American (Gunier et al., 2003; Jerrett et al., 2001; Morello-Frosch et al., 2005) and other European studies (Chaix et al., 2006; Wheeler et al., 2005) in which

the most deprived bear the highest air pollution concentrations. Repeating our analysis using socioeconomic variables at the census tract level (percentage of people with low education, deprivation index) yielded the same results.

The people with higher levels of education were also younger, have lower blood pressure, BMI, glucose levels, lower LDL, a higher percent have quitted smoking and a lower percent have never smoked compared to people with low and medium education. The stronger effects of air pollution on atherosclerosis for this subgroup were, nevertheless, not explained by age interactions. This may indicate higher susceptibility but, more likely, a better detectability among people with fewer competing risks for atherosclerosis (due to less confounding). It is difficult to evaluate if this is supported by previous studies given that results are very heterogeneous. No effect modification by education level has been observed before for the association of $PM_{2.5}$ with IMT, ABI or coronary artery calcification (Diez Roux et al., 2008), neither for the association of NO_2 or $PM_{2.5}$ with IMT, pulse wave velocity or augmentation index (Lenters et al., 2010a). Whereas increasing effects of roadway proximity on aortic artery calcification have been reported for increasing income (trend $P < 0.01$) (Allen et al., 2009). Effects of $PM_{2.5}$ on systemic inflammation markers have been reported to be stronger in men, and more specifically, in highly educated men (Hoffman 2009), in a population-based study in Germany. On the contrary, stronger effects of $PM_{2.5}$ on IMT have been found for women in a study in participants of clinical trials in California (Künzli et al., 2005a). The change in augmentation index associated with increased levels of NO_2, black smoke and $PM_{2.5}$ was also higher for women in the study of young adults in Utrecht. These interactions might also be partially explained by differential misclassification bias across subgroups. NO_2 and traffic measurements were done outdoors of residences. These might not accurately represent the personal exposure and, thus, potentially introduce differential measurement error given the different time activity patterns for men and women, younger and elderly populations. Further research to identify susceptibility factors is still needed.

Sensitivity analyses

Using the exposure at the address of longest residence would minimize errors from reported addresses, geocoding and averaging

NO_2 estimates for the different addresses. Our results were indistinct when using exposure at the address of longest residence compared to 10-year averaged exposure. Thus, for settings similar to the Spanish Mediterranean region in terms of mobility (i.e. 20% of people changed address at least once in the 10 years) and patterns of spatial distribution of NO_2, the exposure at the address of longest residence is a good proxy for long-term exposure. Collecting, geocoding and assigning exposure levels to only one address per participant would be simpler and more efficient both for participants and researchers. On the other hand, using the exposure at the current address (at the time of examination), as is common in other studies, the effect estimates and their precision decrease, indicating exposure misclassification.

Adjusting for random-effects of area of residence yields estimates closer to the null and less precise. This is consistent with a decrease in the exposure contrast, which makes the detection of an effect more difficult, but on the other hand it may indicate correction for some confounding by area. Although some residual confounding is possible, we had already adjusted our estimates by a large array of confounders, including education at the census tract level, which makes the reduction in exposure contrast a more likely explanation. Living near busy roads is also associated with traffic-related noise. As shown for Girona, this correlation is moderate and differs across neighborhoods (Foraster et al., 2011). Our results were not sensitive to adjustment for noise.

Other limitations of our study are the cross-sectional design, possibility of unmeasured confounding, including confounding related to environmental tobacco smoke and lack of daily activity patterns to assess time spent at home. Major strengths include being the first study of traffic-related pollution and atherosclerosis in the Mediterranean region; the large population-based sample size; a thorough assessment of health and potential confounders; availability of 10-year address history that allowed estimating long-term NO_2 and traffic exposure markers; and a dense NO_2 and traffic monitoring network at participant's home outdoor.

In conclusion, long-term exposure to traffic-related air pollution is associated with subclinical carotid atherosclerosis, with stronger associations in people with high education level and men above 60 years, and with high ABI levels. Longitudinal studies are needed to

confirm whether air pollution contributes to the chronic processes of atherogenesis.

Acknowledgements

We are grateful to Mercè Medina-Ramón, Helmut Schröder, Jaume Targa, Maria Grau, Joan Vila, Pascual Solanas, Núria Pagès, Martina Sidera, Montserrat Pique, Susanna Tello, Ruth Martí, Marta Cabañero and Leny Franco and the REGICOR team for their contributions and to the Institut Cartogràfic de Catalunya and the City Councils of the participating towns for GIS and complementing traffic data.

Research supported by the Government of Spain through the Ministry of Science and Innovation, Carlos III Health Institute, European Regional Development Found (*Ministerio de Ciencia e Innovación, Instituto de Salud Carlos III*) [Red HERACLES RD06/0009, Red RedIAPP RD06/0018, FIS PI081327, FIS PI060258, CIBERESP, and Rio-Hortega contract CM08-00141], Fundació La Marató de TV3 (Grant number: 081632), and Agència de Gestió Ajuts Universitaris de Recerca (2009 SGR 1195).

References

References of this paper are included in Section 10.

Table 1. Descriptive statistics of the study population included in main analyses *n*=2780. N (%) unless otherwise indicated.

IMTcca (mm), median ± IQR	0.68 ± 0.19
IMT6seg (mm), median ± IQR	0.67 ± 0.18
Ankle Brachial Index (minimum of left and right), median ± IQR	1.10 ± 0.12
Age (yrs), median ± IQR	58 ± 18
Sex (women)	1491 (53.6)
Education level	
Primary school or illiterate	1476 (53.1)
Secondary school	758 (27.3)
Technician or higher education degree	526 (18.9)
Occupational status	
Employed	1447 (53.1)
Inactive or house keeper	358 (12.9)
Retired	852 (30.7)
Unemployed	68 (2.5)
Smoking status	
Never smoker	1202 (54,5)
Former smoker	628 (28,5)
Current smoker	377 (17,1)
Marital status	
Single	165 (5.9)
Married / living together	2178 (78.4)
Divorced	171 (6.2)
Widow	247 (8.9)
Other	9 (0.3)
Menopause (% in women)	1075 (38.7)
Body mass index, median ± IQR	26.6 ± 5.5
Waist circumference (cm), median ± IQR	93 ± 17
Hypertension†	1265 (45.5)
Diabetes mellitus	349 (12.6)
High density lipoprotein (mg/dl), median ± IQR	52.9 ± 15.6
Hypercholesterolemia	690 (24.8)
Any cardiovascular or antihypertensive medication treatment	1137 (40.9)
Mediterranean diet index ‡, median ± IQR	25 ± 4
Energy expenditure in leasure time (MET-min/week)*, median ± IQR	1515 ±1937
People with low education in the census tract (%), median ± IQR	11 ±12.2
Living at the same address for 10 years before IMT measurement	2252 (81)

* MET: Metabolic equivalent

† Systolic blood pressure >=140 mm Hg or diastolic blood pressure >= 90 mm Hg. or treatment

‡ Adherence to Mediterranean diet was defined as quartiles of the index

Table 2. Descriptive statistics of exposure to air pollution and traffic. *n*=2780. Minimum, 5th percentile, mean, standard deviation, 95th percentile and maximum.

	min	5th perc.	mean	sd	95th perc.	max
10yr. average NO_2 ($\mu g/m^3$)	4.9	9.7	20.7	8.4	35.5	47.7
NO_2 at address of longest residence ($\mu g/m^3$)	3.5	10.3	22.0	9.0	37.9	50.4
NO_2 at current address ($\mu g/m^3$)*	3.5	10.2	21.9	9.0	37.7	50.4
NO_2 at the closest monitor within 200m ($\mu g/m^3$)†	3.7	9.2	22.3	9.3	40.5	52.9
10yr. average traffic load 100m buffer (1000 veh m/day)	0	231	2551	2246	7436	11149
10yr. average traffic intensity nearest street (1000 veh/day)	0	0.1	3.2	5.4	15.2	34.2

* Sample size restricted to participants living in study area at moment of examination N=2723

† Sample size restricted to participant living within 200m of monitoring site N=2265. NO_2 was monitored with passive samplers

Table 3. Effect estimates of percent change in intima media thickness (IMT) associated with a 10-yr average exposure contrast between the 5th and 95th percentiles. Exposure contrast indicated in parenthesis.

	IMTcca			IMT6seg		
	N	**%change***	**(95% CI)**	**N**	**%change***	**(95% CI)**
NO_2 (25 μg/m^3)						
crude	2780	3.67	(1.37 to 5.98)	2188	4.98	(2.65 to 7.31)
adjusted for sex	2780	3.67	(1.38 to 5.96)	2188	4.88	(2.58 to 7.18)
adjusted for age & sex	2780	0.04	(-1.83 to 1.92)	2188	0.84	(-1.02 to 2.71)
Model 1	2738	0.35	(-1.63 to 2.32)	2155	0.71	(-1.25 to 2.67)
Model 2 (possible intermediates)	2632	0.56	(-1.47 to 2.59)	2074	0.52	(-1.52 to 2.57)
Traffic load in a 100m buffer (7.200.000 veh m/day)						
crude	2780	5.25	(2.76 to 7.74)	2188	6.38	(3.89 to 8.88)
adjusted for sex	2780	5.21	(2.73 to 7.68)	2188	6.31	(3.85 to 8.78)
adjusted for age & sex	2780	1.39	(-0.64 to 3.42)	2188	1.99	(-0.02 to 4.00)
Model 1	2738	1.78	(-0.33 to 3.89)	2155	2.08	(0.00 to 4.17)
Model 2 (possible intermediates)	2609	1.91	(-0.24 to 4.06)	2053	2.06	(-0.09 to 4.21)
Traffic intensity in nearest street (15.000 veh/day)						
crude	2780	4.18	(2.01 to 6.35)	2188	4.55	(2.43 to 6.68)
adjusted for sex	2780	4.13	(1.98 to 6.29)	2188	4.42	(2.32 to 6.51)
adjusted for age & sex	2780	1.74	(-0.02 to 3.50)	2188	1.75	(0.05 to 3.44)
Model 1	2738	1.96	(0.14 to 3.77)	2155	1.70	(-0.04 to 3.44)
Model 2 (possible intermediates)	2632	2.32	(0.48 to 4.17)	2074	1.80	(0.01 to 3.59)

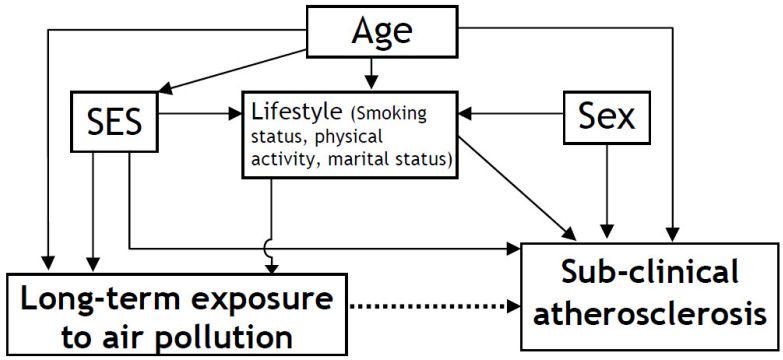
Model 1: sex, age, sex-age interaction, smoking status, education and marital status
Model 2: model 1 plus BMI, HDL, waist circumference, systolic and diastolic blood pressure, weekly energy expenditure in physical activity during leisure-time (tertiles), adherence to Mediterranean diet, plausibility of reported diet, medication treatment and percentage of people with low education at the census tract level, In models for traffic load outcomes were also adjusted for occupational status.

Table 4. Effect estimates of relative risk ratio of categories of ankle brachial index (ABI) (reference category ABI= 0.9 to 1.3) associated with a 10-yr average exposure contrast between the 5th and 95th percentiles. Exposure contrast indicated in parenthesis.

	ABI<0.90		ABI>1.3	
	RRR	(95% CI)	RRR	(95% CI)
NO_2 (25 μg/m^3)				
Model 2 (possible intermediates)	0.72	(0.29 to 1.75)	1.98	(1.09 to 3.60)
Traffic load in a 100m buffer (7.200.000 veh m/day)				
Model 2 (possible intermediates)	1.02	(0.40 to 2.61)	1.89	(1.07 to 3.34)
Traffic intensity in nearest street (15.000 veh/day)				
Model 2 (possible intermediates)	0.48	(0.16 to 1.46)	1.70	(1.13 to 2.57)

Estimates adjusted by Model 2: sex, age, sex-age interaction, smoking status, education and marital status, BMI, HDL, waist circumference, systolic and diastolic blood pressure, weekly energy expenditure in physical activity during leisure-time (tertiles), adherence to Mediterranean diet, plausibility of reported diet, medication treatment and percentage of people with low education at the census tract level, In models for traffic load outcomes were also adjusted for occupational status.

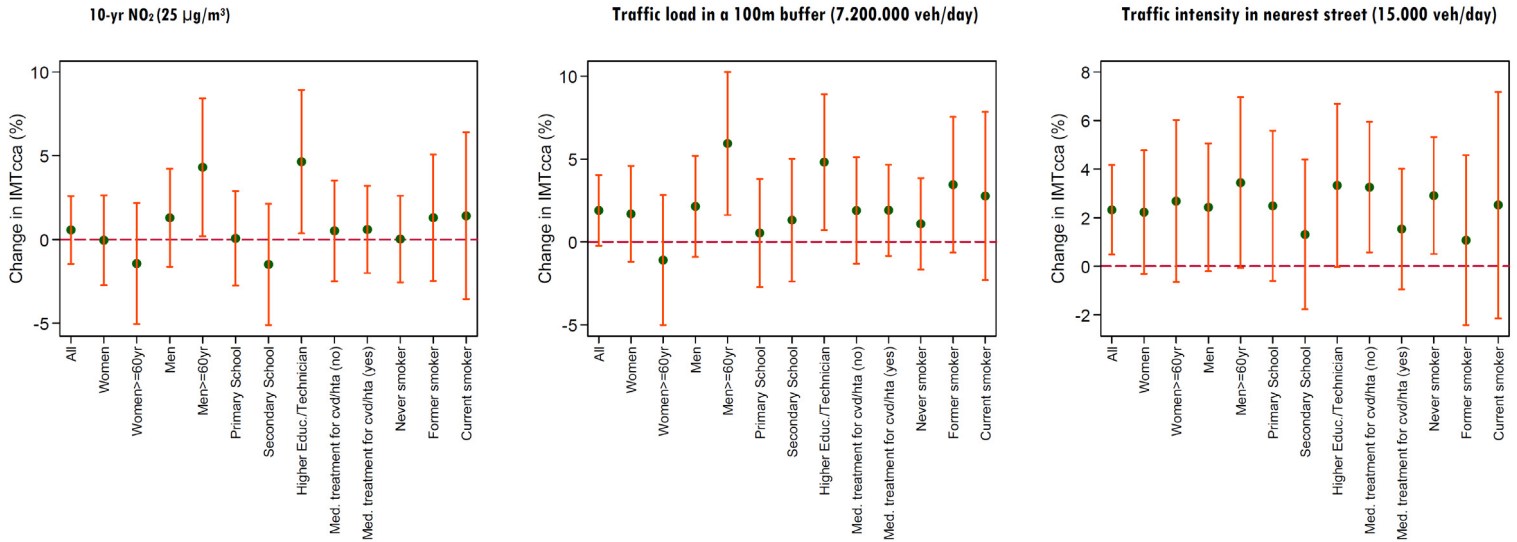
Figure 1. Directed Acyclic Graph (DAG) for the association of exposure to traffic-related pollutants with atherosclerosis.



SES: Socioeconomic status surrogate used educational level.

Figure 2. Subgroup analysis: Effect estimates of percent change in intima media thickness (IMT) associated with an exposure contrast between the 5th and 95th percentiles by groups of sex, sex and age, education level, medical treatment and smoking status.

Estimates adjusted by sex, age, sex-age interaction, smoking status, education and marital status plus BMI, HDL, waist circumference, systolic and diastolic blood pressure, weekly energy expenditure in physical activity during leisure-time (tertiles), adherence to Mediterranean diet, plausibility of reported diet, medication treatment and percentage of people with low education at the census tract level. In models for traffic load outcomes were also adjusted for occupational status. Effect estimates for traffic load in 100m and for IMT6seg are shown in Supplement Figure S3.



Supplemental Material

Table S1. Variables and performance of land use regression models for each group of towns.

Town(s) *(number of monitoring sites)*	Variables of land use regresion model	R^2_A	R^2(cv)
Girona Salt (N=151)	Building's floor number 24 hour total traffic load of all roads in 25m 24 hour total traffic load of all roads in 500m (excluding traffic load in 25m) length of all roads within 1000m	0.63	0.61
Banyoles Porqueres (N=56)	24 hour total traffic load of all roads in 300m product of traffic intensity on nearest road and inverse distance to nearest road nitrogen oxide concentration derived from a regional dispersion model at 1x1 km resolution	0.33	0.32
Blanes (N=91)	24 hour total traffic load of major roads in 500m length of all roads within 100m 24 hour total traffic intensity on nearest road area of high density residential land within 100m	0.38	0.36
La Bisbal Palafrugell Palamós (N=105)	household density within 1000m 24 hour total traffic load of all roads in 50m area of industrial land within 1000m	0.44	0.42
Olot Sant Joan (N=88)	length of all roads within 1000m distance to major road 24 hour total traffic load of all roads in 25m area of buildings within 100m area of industrial land within 300m	0.53	0.50
Llagostera Sta.Cristina (N=43)	24 hour total traffic load of all roads in 500m	0.51	0.51

R^2_A: adjusted coefficient of determination

R^2(cv): cross-validation coefficient of determination

Table S2. Population characteristics, exposure assessment, main results and effect modifications of the published epidemiological studies on subclinical markers of CVD. Adapted with permission from Künzli et al. (2011).

*This table is the **Table 2** of the thesis Introduction Section (page 34)*

Table S3. Correlations between markers of exposure to traffic-related pollutants N=2780. All correlation coefficients are significant at the 0.001 p-value level.

	10yr averaged NO2 (ug/m3)	NO2 at address of longest residence (ug/m3)	NO2 at current address (µg/m3)*	NO2 at the closest monitor within 200m (µg/m3)†	10yr. average traffic load 100m buffer (1000 veh m/day)	10yr. average traffic int. nearest st.(1000 veh/day)
10yr averaged NO2 (µg/m3)	1					
NO2 at address of longest residence (µg/m3)	0.990	1				
NO2 at current address (µg/m3)*	0.986	0.979	1			
NO2 at the closest monitor within 200m (µg/m3)†	0.770	0.780	0.761	1		
10yr. average traffic load 100m buffer (1000 veh m/day)	0.720	0.720	0.713	0.570	1	
10yr. average traffic int. nearest st.(1000 veh/day)	0.520	0.520	0.514	0.360	0.580	1

* Sample size restricted to participants living in study area at moment of examination N=2723

† Sample size restricted to participant living within 200m of monitoring site N=2265. NO_2 was monitored with passive samplers

Figure S1. Map of study region and locations of study participants. Points represent address of longest residence.

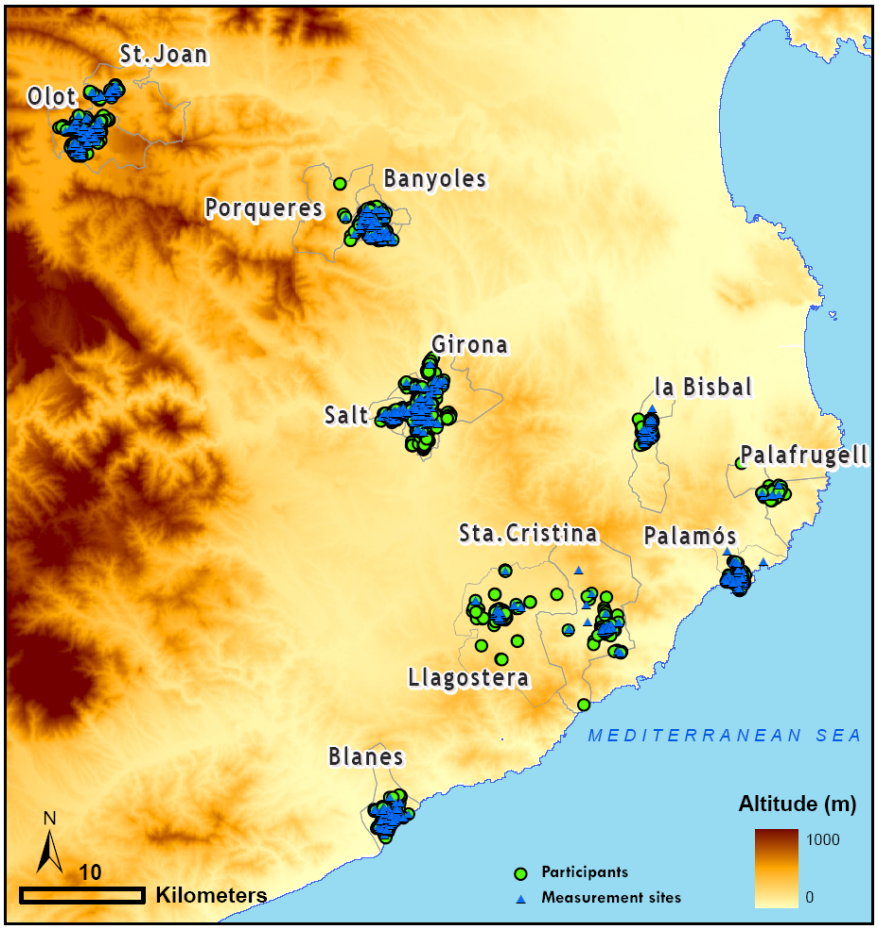


Figure S2. Scatter plot of concentrations derived with the models vs. concentrations at the measurement locations for participants living within 200m of a measurement location.

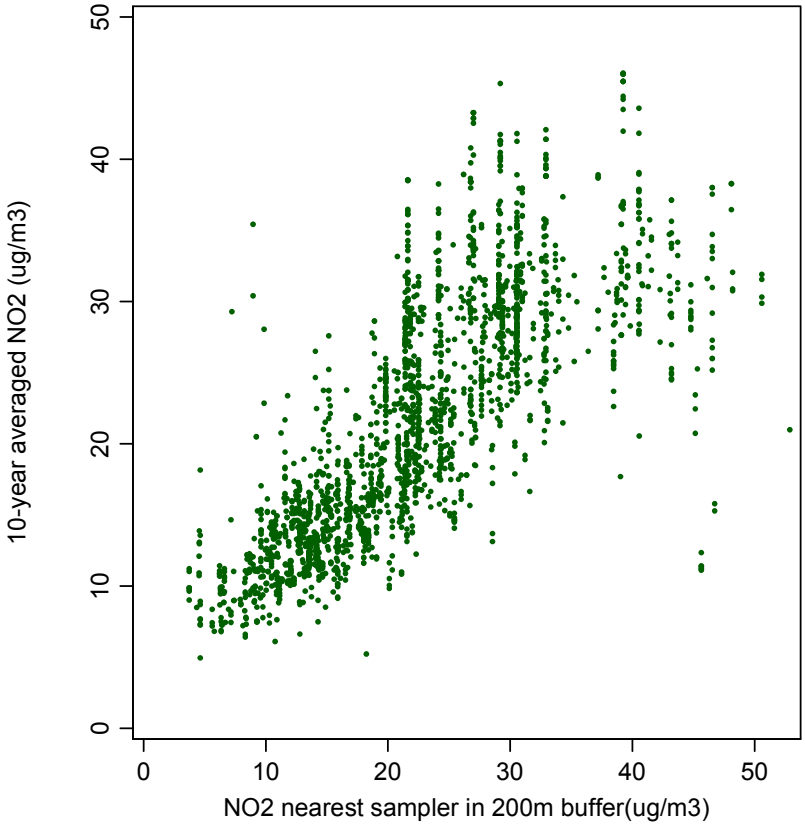


Figure S3. Subgroup analysis: Effect estimates of percent change in intima media thickness (IMTcca and IMT6seg) associated with an exposure contrast between the 5th and 95th percentiles by groups of sex, sex and age, education level, medical treatment and smoking status.

Estimates adjusted by sex, age, sex-age interaction, smoking status, education and marital status plus BMI, HDL, waist circumference, systolic and diastolic blood pressure, weekly energy expenditure in physical activity during leisure-time (tertiles), adherence to Mediterranean diet, plausibility of reported diet, medication treatment and percentage of people with low education at the census tract level. In models for traffic load outcomes were also adjusted for occupational status.

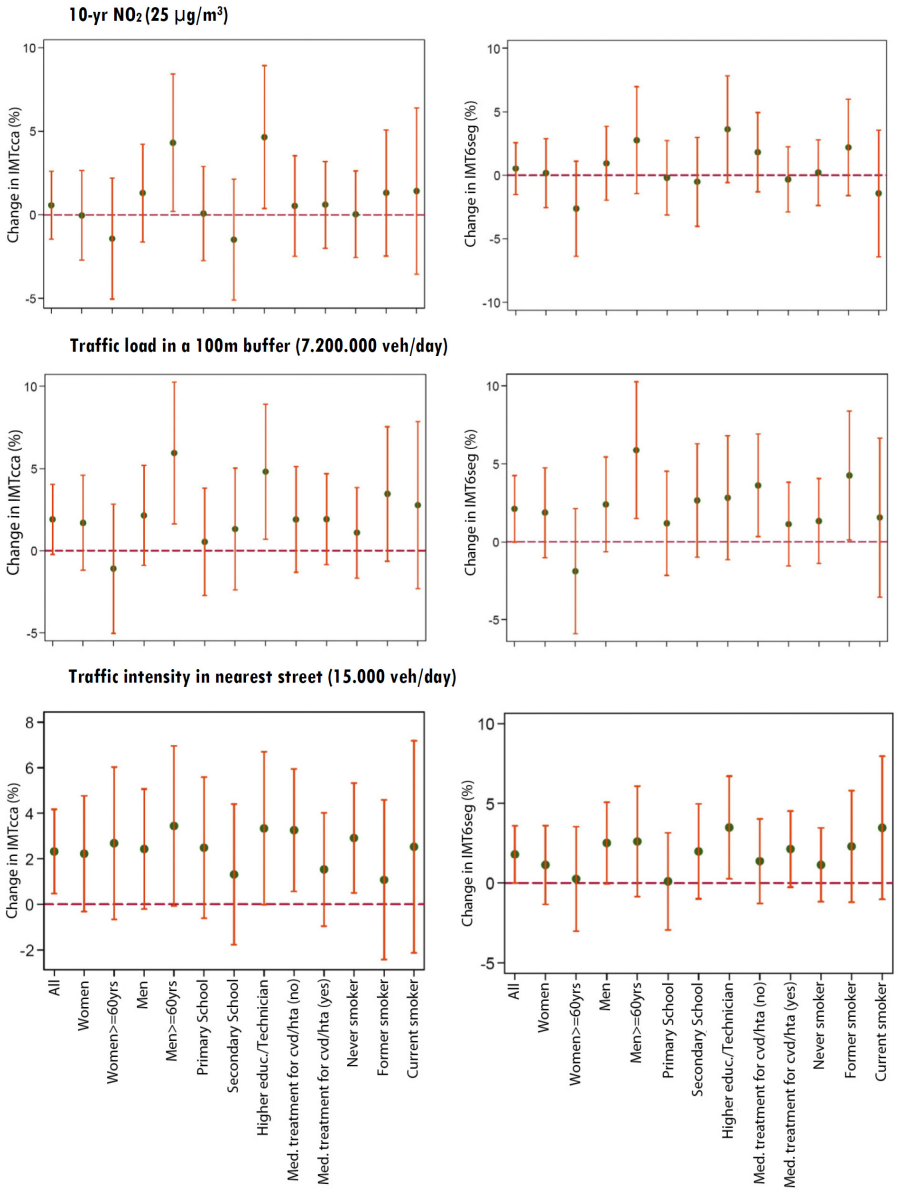
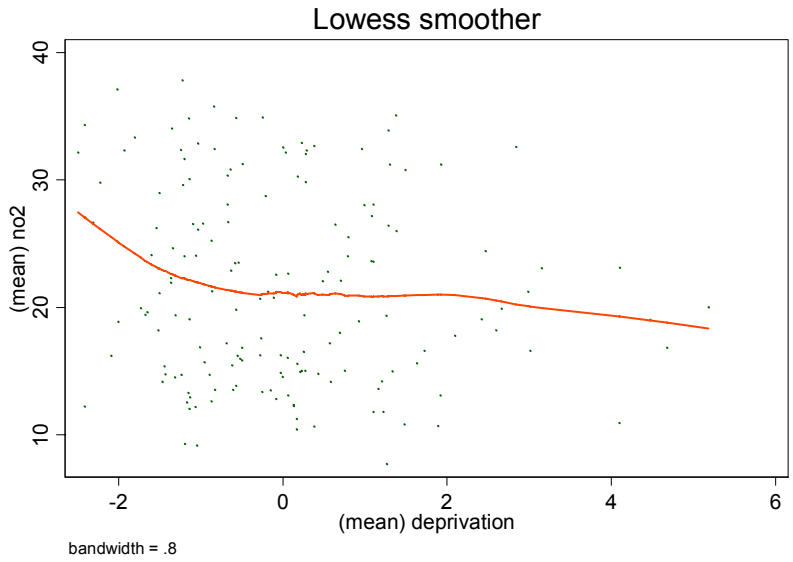


Figure S4. Scatter plot of residential NO_2 and deprivation index at the census tract level. Both measured at the address of longest residence



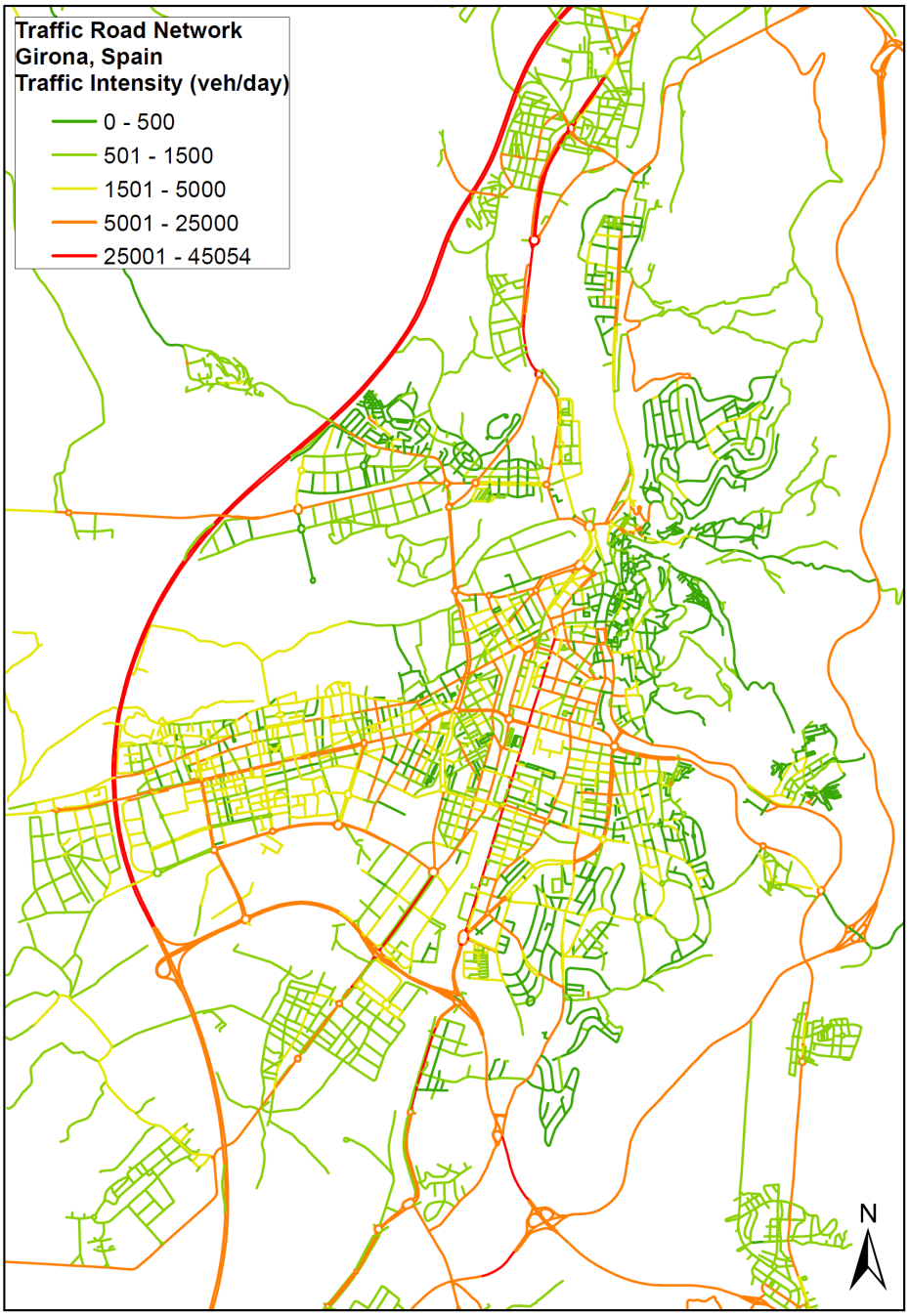


Figure 12. Traffic intensity road network database. Area shown: Girona city and Salt.

Palamós Photo: Talavan

9 DISCUSSION

This section is complementary to the discussion paragraphs included in each of the four papers. It aims to give a broader and more integrated interpretation of the entire thesis project.

9.1 Choice of air pollution marker and monitoring technique

The project used a range of markers and methods to characterize peoples' long-term exposure to traffic-related air pollution, including the novel approach using mosses deployed for 2 months, 4-week NO_2 monitoring with passive sampler, and short measurements of UFP and traffic intensity. Each method has inherent advantages and limitations to be discussed.

Mosses are useful as passive samplers for long-term monitoring of heavy metals, particularly for residential monitoring. The moss samplers need neither maintenance nor electricity and do not emit noise, which is a challenge if monitors are run with pumps. Moreover, the moss samplers are small, thus can be employed almost everywhere, including places where conventional PM monitors are rather intrusive such as indoor sampling at homes of study participants. On the other hand, the resulting concentrations of elements deposited in mosses include all particle size ranges, thus it is not possible to distinguish if an element is enriched in e.g. the fine or coarse fraction. However, this might not be necessary for epidemiological studies given that all PM fractions –ultrafine, fine and coarse- have been associated to health effects (Hoek et al., 2009) (Bell et al., 2009) (Perez et al., 2008).

Passive monitoring of NO_2 is a widely used method given its relatively low costs, ease to install and possibility to cover a wide range of location types (indoor and outdoor homes, traffic lights, street poles and personal monitoring). As shown before and confirmed in our study, passive samplers can be used for periods of up to four weeks (Targa, 2008), being thus appropriate for long-term measurements. Compared to the two-week average NO_2 measured in several studies (Beelen et al., 2011; Lewné et al., 2004), four-week average might be less affected by short-term

peaks and, depending on the study area, might be sufficient to characterize the annual mean if deployed in the month with average concentration closer to the annual mean. The use of residential outdoor NO_2 as an exposure indicator is discussed in the next sections.

For the UFP, we tested a completely different and novel approach of fast mobile, monitoring of ultrafine particles for short-term periods to estimate long-term averages. The technique is very appealing considering its cost-effectiveness. The lack of studies on the long-term exposure to UFP was rated, in a recent expert panel elicitation study (Hoek et al., 2009), as the main factor of uncertainty to estimate effects of UFP on all-cause mortality. Implementing relatively inexpensive techniques, such as the one we used, would allow more studies on UFP as well in resource constraint settings. The mobile monitoring technique allows profiting from the trade-off between sampling period and number of monitoring locations: instead of using conventional equipment to conduct measurements during one or more weeks in one or a few locations, hand-held particle counters (such as p-Traks or condensation particle counters) can be used to do short-term measurements at several locations in one day. This is particularly relevant when considering the increase in model performance resulting from a higher number of monitoring sites (see Section 9.2 for discussion of this point). Nevertheless, the question of what is the shortest time period needed to adequately capture an UFP concentration that approximates the long-term mean remains unexplored.

Annual averages of traffic intensity are surprisingly well characterized with one single 15-minute measurement taken during non-rush hour periods (Van Roosbroeck et al., 2007). With traffic being the main source of UFP in a city like Girona, we piggy-backed for the first time a 15-min UFP measurement protocol to the traffic count protocol to pilot its use to establish a long-term average of the UFP number concentration. The UFP variability, however, depends not only on the number of vehicles but on physicochemical processes determined by distance to the source, meteorological conditions, dispersion patterns, etc., thus being much more complex than the variability in traffic count. We have tempered this problem through adjustment of temporal variability and repeated measurements in a subset of locations. Temporal variability was adjusted based on measurements at a fixed monitoring station in Girona, city. However, this adjustment was not sufficient as shown by the improvement of

the models when including monitoring day and hour. To properly adjust for temporal variability, having at least one central monitoring site or, if possible, one per urban area is recommended. They should be located in an urban background location (i.e. >75 m away from road traffic or other local sources). Repeated measurements in a subset (N=25), showed that performance of models to estimate the long-term averages improved from a moderate 0.40 using one measurement to a rather high cross-validation R^2 of 0.69 when using the average of two measurements. While our pilot study was restricted to only one repeat, further improvements could be expect for protocols based on 2-3 repeated measurements to estimate long-term mean UFP. This would still be far less resource demanding than monitoring during some 2 week periods in each season as done in traditional study protocols.

Based on these novel findings, I expect substantial benefits from collecting repeated measurements on randomly selected days and hours for every location and from having continuous measurement of UFP at multiple reference sites. Further studies may determine the optimal number of short-term repeats needed to efficiently estimate long-term exposures.

Traffic-intensity measurements can be used directly or as predictor variables to model any pollutant component of the traffic-related mixture. This data is sometimes not available, not accessible or not precise. A common example of the latter is when the traffic in streets considered “low traffic” streets is not measured, but assigned using heuristics. Using such data to model pollution might result in error, particularly in situations when low traffic is associated to high pollution e.g. in street canyons; automotive park with poor quality/efficiency engines, etc. Collecting direct measures of traffic intensity is a simple and relatively inexpensive alternative; traffic counting can be easily integrated within mobile monitoring or any other field measurements; traffic measures are stable over years (Beelen et al., 2007a) provided the street's traffic-direction are not changed and traffic restrictions are not implemented; using these measures to derive exposure is straightforward, particularly now with the development of recent GIS applications.

Collecting traffic counts also have the advantage that the vehicle type (light or heavy duty, autobuses) can be distinguished. For example, using tally counters we counted heavy vehicles with the left hand, and light vehicles with the right hand whereas motorcycles

counts were just kept in memory. With two people doing the count, buses can be differentiated also, particularly in absence of bus routes and frequencies. For total traffic (sum of heavy, light, buses and motorcycles) 15-minute counting has been reported to properly capture the annual average traffic intensity and this was later confirmed with official data for our study area. Nevertheless the shortest sufficient counting time to reliably estimate annual average traffic intensity of the different types of vehicles (heavy, light, bus and motorcycles) need to be investigated. The different types of traffic might be independently associated to different components of pollution. Thus, differentiating them in the traffic count would allow investigating associations of specific vehicle-type traffic intensity with health outcomes. This could be of added value in policy making.

9.1.1 Comparison between markers of long-term exposure to air pollution

Another key issue is whether each of these markers, independently, is a reasonable proxy for different aspects of the complex mixture of local traffic-related pollution. The validity of using individual components as indicators of the mixture remains unexplored. However, from the exposure assessment perspective, the suitability of the markers we used can be evaluated. The following evaluation is based on the specific monitoring protocol that we used for each marker and, thus can not be extrapolated to all monitoring methodologies.

Heavy metals showed a higher spatial variability than NO_2. This indicates that metals capture spatial components of local pollution that an NO_2 surface may not fully characterize. Metals also showed a stronger association with traffic intensity in the nearest street than NO_2. Compared to NO_2 and other markers of combustion processes, some metals such as antimony and molybdenum are markers almost exclusive of road traffic, reflecting though not tail pipe emissions–such as in case of NO_2–but emissions from break pads and resuspensions. Finally, the use of heavy metals is appealing given its established toxicity. Thus, monitoring heavy metals as markers of traffic-related local pollution should be considered complementary to the monitoring/modeling of markers of local tail-pipe primary and secondary pollutants such as NO_2 and of markers of urban background and secondary pollutants such as PM_{10} or $PM_{2.5}$ mass,

which do not capture local contrasts but combine the impact of a broad range of sources.

In our case, traffic was the main determinant of all studied markers of pollution: for UFP and NO_2 it was the total traffic (sum of heavy and light traffic) and for heavy metals it was the buses circulating in the nearest street. The correlations between two-month averages of NO_2 and metals simultaneously measured were low to moderate (Spearman rank coefficient between NO_2 and cadmium was 0.15 and between NO_2 and antimony was 0.61). The correlations between annual average NO_2 and short-term UFP measurements were higher (Pearson correlation coefficient by town ranged between 0.66 and 0.87). Nevertheless, in other studies, only moderate spatial correlation of long-term mean UFP with NO_2 have been observed, with Pearson correlations of 0.34 in Raleigh, North Carolina (Hagler et al., 2009); 0.62 in Vancouver Canada (Abernethy et al., 2011); 0.42 to 0.48 in Basel, Switzerland (Phuleria et al., 2011). This raises the question to what extent spatial models of NO_2 -available since many years in many cities- may be considered markers of UFP exposure. Developing spatial models of UFP for the assessment of long-term exposure should thus be considered an alternative.

Direct measures of traffic, in spite of some limitations, are appealing markers of local exposure to traffic-related pollution, particularly for communities where information in GIS format such as maps of urban structure, population density by given area, etc. may not be available. Although there are many towns that lack such data, the road networks are more easily available. Thus, if traffic counts are conducted, they can be integrated to the roadnetwork and exposure measures can be estimated. The limitations of using traffic intensity as a measure of exposure include: "*they are error prone because ignore dispersion parameters and physicochemical activity of the pollutants*"; "*estimates based on proximity can be confounded by factors such as socioeconomic status and noise*" (HEI Panel on the Health Effects of Traffic-Related Air Pollution, 2010). Nevertheless, it is recognized that these limitations especially apply to simple measures of proximity to roads or road length and of pollutant surrogates without specific traffic data; while more elaborated measures, such as traffic intensity in a 100m buffer, might be more valid. Although the traffic intensity is stable over years, the emissions from traffic would be changed if the traffic fleet (trucks, diesel and gasoline powered vehicles) changes or if the maximum allowed

speed changes. On the other hand, the use of direct traffic measures have some advantages: since they measure the source of pollution, results of epidemiology or risk assessment studies may be more directly applicable to policy related questions and interventions; besides, intermediate evaluations of such policies are easily measurable (traffic counts); health effects associated to measures of traffic intensity, as well as, prevention or abatement policies based on them should be easier to communicate to the public, compared to results associated to pollution levels.

Our direct-traffic measurements showed higher spatial variability than NO_2 (IQR/p50 was 2.51 for 10-yr traffic in the nearest street and 0.73 for 10-yr NO_2). The traffic measurements were input as potential predictors to the NO_2 LUR model selection process. They were selected in all the town-specific LUR models along with other measures of traffic such as length of roads within buffers; therefore, some of the spatial variability of traffic intensity is included in the modeled NO_2. Thus, the main differences between what these two traffic markers capture are that annual average NO_2 outdoor homes includes NO_2 emitted by other sources different than traffic such as residential heating, gas cooking, cigarette smoking, etc; while traffic measurements do not take into account dispersion patterns except for distance (as for traffic load in a 100m buffer).

In our study, the traffic intensity might have captured the variability of personal exposure to traffic-related pollution (adjusted for indoor sources) better than the outdoor NO_2. The correlation between outdoor and personal NO_2 varies from as low as 0.06 for children living near busy roads in urban areas of Dusseldorf (Krämer et al., 2000), to correlations as high as 0.72 for elderly population in Hamilton, Canada (Sahsuvaroglu et al., 2009). Validation studies between direct measures of traffic and personal exposure are scarce, particularly for long-term exposure measurements. The existent studies use measurements of traffic intensity on the nearest street or proximity to high versus low traffic, however, traffic intensity within buffers has not been validated. Differences in the validation of these measurements are expected given the moderate correlation between them, e.g. in our study the correlation between traffic intensity in the nearest street and traffic intensity in a 100m buffer was 0.58. In three studies in The Netherlands, traffic intensity in the nearest street and proximity to high versus low traffic were consistently associated with children's personal exposure to soot (also called elemental carbon and measured as $PM_{2.5}$ absorbance).

However, outdoor NO_2 was not associated with personal exposure to NO_2 (Van Roosbroeck et al., 2008; Van Roosbroeck et al., 2006; Van Roosbroeck et al., 2007).

In summary, there is large variability in the percentage of personal exposure explained by both traffic measurements and residential NO_2. In some instances, traffic has been a better indicator of personal exposure to traffic related pollution. Nevertheless, it is very difficult to generalize results from other studies given that the association between personal exposure and outdoor pollution or traffic is modified by the segment of the population under study (children, adults, elderly; men, women, etc.), indoor sources, time-activity patterns, life-style, weather, ventilation, housing stock, traffic composition, etc., most of them specific to the study area. Thus, ideally, measurements of these factors should be included in the exposure assessment. Or as an alternative, the exposure measurements used should be validated by personal exposure measurements for a subset of participants to adjust for measurement error (Baxter et al., 2009).

Another factor that may explain the stronger associations found between IMT and traffic intensity than between IMT and NO_2 is that traffic is directly used in the health model, while NO_2 was modeled in terms of several additional covariates, in order to extrapolate the measurements to all participants. Modeling exposure can also introduce measurement error. This is further discussed in the next section.

9.2 Modeling of air pollution markers by land use regression

The measurements of the different markers of exposure to traffic-related pollution were used to build (LUR) models so that predictions could be drawn for places where measurements were not taken. NO_2 measurements were collected in a highly dense monitoring network and annual averages were successfully derived, thus we used these measurements to test the effect of the sample size in the LUR models. Our study was one of the first to formally address and quantify the impact of having large number of sites to model the local distribution of pollution.

Performance of land use regression models is usually measured by the adjusted R^2 and leave-one-out cross-validation R^2, nevertheless, these measurements largely overestimate the true R^2 (held-out validation R^2) when using small sample sizes and the overestimation is less pronounced when *n* increases. This this is true especially when the final variables included in the model are selected from a large set of potential predictors. To correct for this overestimation and have models with better performance, a larger number of monitoring sites is required. This is reflected in the increase in the attenuation coefficient (decrease in exposure measurement error) with increasing sample size. For sample sizes above 80 measurement sites the attenuation coefficient is higher than 0.85 (on average). This means that the estimates of the health model would be biased to the null by 15% as compared to the association estimated from a model with direct measurements. For most of our town-specific LUR models we had more than 80 monitoring sites, however, for two of the towns the sample size was smaller (N=43 for Llagostera, Sta. Cristina and N=56 for Banyoles, Porqueres). In those cases there could be an attenuation of the effects of NO_2 on IMT of approximately 0.65 to 0.7. To formally test the influence of modelled versus measured NO_2 values, I compared the estimates effects yield by both measurements for the subset of participants living within 100m and within 200m of a monitoring location. Interestingly, there was no difference in the estimates, which indicates that the measurement error introduced by the variable selection in the LUR modeling did not lead to attenuation of the effects. This might be due to few participants living in the towns with fewer monitoring sites and to random variation.

If regression based modeling (LUR) is intended it is advised to use dense monitoring networks and consider doing the monitoring in batches to use the information collected on the first batch to inform the site selection process for the second, using methods as the one proposed by Kanaroglou et al. (2005).

There are 2 main types of measurement error, "classical" and "Berkson". Classical error is introduced because the measured exposure is likely to vary around the true exposure. Thus, the measurements are expected to have more variation than the true levels (e.g. blood pressure). As a result of Berkson type error, the measurements are expected to have less variation than the true levels (e.g. assigning the value of a central monitor to all inhabitants of a town). Both types of measurement error increase the standard

errors in the health model, but only classical measurement error biases the coefficients in the health model (towards the null). It is usually reported that exposure estimates derived through regression models introduce Berkson-like measurement error, being this the largest component of the error (Suglia et al., 2008) (Alexeeff et al.). Nevertheless, such exposure estimates might introduce classical error as well. We showed that in the presence of variable selection, there is bias (attenuation) in the exposure-effect relationship, and this bias depends on the number of monitoring sites (and variables) used to build the exposure model. Thus considering the error introduced by LUR models to be Berkson type only, with the subsequent assumption that the coefficients of the health model will not be biased, is erroneous. Classical measurement error resulting from predictors' selection in regression modeling (of exposure) should be taken into account when quantifying the effects of measurement error. Also, a correction or evaluation of the impact on the associations between exposure and health should be performed.

Because of measurement error, model assumptions and low quality/precision of the data input to the model, the measurement error introduced with the model might offset the advantages of using modeled estimates of exposure over using direct measures of traffic intensity (HEI Panel on the Health Effects of Traffic-Related Air Pollution, 2010). Thus, the marker of pollution used to build the model should be a better marker of personal exposure to the hypothesized causal agents than traffic intensity itself.

A limitation of assessing exposures based on measurements or models of residential pollution/traffic is that exposures at places other than the residence are assumed negligible. In our case, time activity and mobility patterns are not taken into account as only residential location was used in the LUR models. This could have introduced bias towards the null hypothesis (Setton et al., 2011). It is difficult to have an approximated estimate of this bias given that it depends on time spent away from home and the levels of pollution at the alternative locations.

9.3 Effects of long-term exposure to air-pollution on atherosclerosis

We have found significant associations between residential outdoor traffic intensity and subclinical atherosclerosis in healthy population of the Spanish Mediterranean region; instead modeled NO_2 concentrations were associated with these outcomes only in the subgroups of high education levels and men above 60 years. In general, results using direct traffic measurements as a surrogate seem to be more consistent across studies (see **Table 2**).

This is the first study on the association of long-term exposure to air pollution and atherosclerosis in the Mediterranean region. This region is particular because it has one of the lowest cardiovascular mortalities in Europe, while the conventional cardiovascular risks factors are among the highest (Masiá et al., 1998). In this region, vehicle densities within cities are among the highest in Europe. For example, the number of vehicles per total surface in Barcelona and Milan were 6100 and 4100 veh/km^2 respectively (in 2007). Our study was conducted in small- to medium size- towns, where vehicle density is moderate compared to large Mediterranean cities–1500 veh/km^2 in Girona city- but still high compared to northern European cities–e.g. 1600 veh/km^2 in London. In our study, a contrast between the fifth and ninety-fifth percentiles of 10-yr exposure to residential traffic (load and intensity) was associated with approximately 2% thicker IMT.

Reasons that might have led to find an association of IMT with traffic intensity but not with NO_2 include the lower spatial variability of NO_2 compared to traffic intensity. Other hypothetical reasons (not addressed in our study) include, lower spatial correlation (and agreement) between outdoor NO_2 and personal exposure to traffic-related pollution than between traffic intensity and personal exposure; low spatial correlations between NO_2 and other components of the traffic-related air pollution mixture that are responsible for inducing atherosclerosis, e.g. heavy metals, PM, UFP, volatile compounds, resuspended dust and noise; and traffic markers not taking into account dispersion patterns of pollution, except proximity in the case of traffic intensity within buffers (discussed above).

Regarding ABI, we found that higher exposures to traffic-related pollution were associated with increased risk of high ABI. Also that high ABI was not associated to thicker IMT while low ABI was. The segment of ABI that is more clearly associated to higher cardiovascular risks is ABI<0.9, it has thus received more interest in the literature. We have very few participants with ABI<0.9, which might explain not finding an association between air pollution and low ABI. The segment of ABI>1.3 have been also associated to higher cardiovascular risk. This u-shape of the association between ABI and cardiovascular risk makes ABI a complex marker and makes comparison between studies difficult. The more common approach in studies on the association of ABI with air pollution is to categorize ABI in lower or higher than 0.9. However, some have opted for excluding the ABI higher than 1.3 (Hoffmann et al., 2009b) while others include them (Diez Roux et al., 2008). Dichotomizing the ABI using the 0.90 cut point may lead to underestimation of effects (Allison et al., 2008). And excluding the segment of high ABI would not allow a better understanding of the risks and possibly the mechanisms of peripheral atherosclerosis. We opted thus for a multinomial analysis including low, normal and high ABI levels. Our results suggest that higher air pollution is associated to high ABI and thicker IMT. While further confirmation from other studies is needed, our results lead to speculate that the atherosclerotic response to varying levels of air pollution in the carotid arteries is different to that in the peripheral arteries.

Long-term exposure to NO_2 and traffic intensity was associated to many of the established risk factors of atherosclerosis even after adjusting for age and sex. Very strong confounding, particularly by age, was observed in the associations of traffic-related pollution and subclinical atherosclerosis. This was not due to interaction by age. Atherosclerosis is a cumulative process that starts in childhood and age is its principal determinant. On the other hand, age was strongly correlated with 10-year NO_2, current NO_2 and obviously with the cumulated life-time exposure to air pollution. Adjustment covariates were selected based on the directed acyclic graph DAG, which is defined based on prior knowledge and associations observed in our data. A DAG is not a conceptual model of the mechanisms or pathways of the disease; it is a tool to infer on biases and facilitate the decision of which variables need to be adjusted for.

In a DAG only the common causes of any pair of variables in the diagram need to be drawn. The chosen DAG is a simplification of the temporal variability of exposure. In addition, we implicitly assumed that there were no *feedback loops* in the association. A feedback loop could occur if, for example, exposure to air pollution induced atherosclerosis and once ahterosclerosis progressed, people tried to avoid future exposure to air pollution. We further adjusted by variables that might be in the pathway of the association, nevertheless residual confounding might still remain.

Statistical adjustment might not be enough to control for established cardiovascular risk factors that were associated to NO_2. For example, NO_2 was inversely associated to several factors that increase the risk of atherosclerosis e.g. BMI, blood pressure and percentage of low education in the census tract. It is also directly associated to factors that decrease the risk of atherosclerosis such as education level as a proxy of socioeconomic status (SES). In case that education level was not a good enough indicator of SES, it would not capture all the variability of SES and adjustment would not be sufficient to control for such confounding.

As mentioned before (section 9.1.1) adjusting for noise and SES is recommended when using direct traffic measurements (HEI Panel on the Health Effects of Traffic-Related Air Pollution, 2010). Adjusting for noise is recommended because the effects of traffic are not only caused by the chemical pollutants but also by the noise they emit. In our case, this adjustment was done in a sensitivity analysis for a subset with available modeled noise from traffic and results did not change. Adjusting for socioeconomic characteristics is recommended particularly when using distance-only metrics as exposure markers; this recommendation was driven by the evidence from US studies where the social deprived are more likely to live close to traffic. As shown in paper IV (page 189) and Supplemental **Figure S4** of the same paper), the opposite is true in Girona. Adjustment by SES was done in all our analyses and it was confirmed a strong confounder and effect modifier of the association of air pollution and atherosclerosis. Given the different confounding patterns of socioeconomic status (SES) observed in the US and northern Europe compared to southern Europe, analyses of the association of air pollution and atherosclerosis should be presented stratifying by SES.

The exposure associated to the address of longest residence resulted as precise as the exposure associated to the last 10-year period

(derived as the time-weighted average of residential NO_2 extrapolated back in time). For study areas with similar mobility patterns and pollution ranges as the Mediterranean Spanish region, collecting the address of longest residence instead of the residence history might be sufficient.

As in most population-based studies, the possibility of a selection bias resulting in the observed associations as well as the lack of associations cannot be ruled out. If air pollution promotes atherosclerosis and if the consequences of atherosclerosis, e.g. a stroke or infarction, reduce the likelihood to participate in the study, one would expect an underestimation of the true – unknown- associations.

Reverse causation cannot be ruled out either as a possible explanation for our results. This could be given if people at increased risk of atherosclerosis move to live in polluted areas, because of age, SES (not necessarily low), accessibility (e.g. in downtown areas the stores, clinics, parks are at a walking distance), etc. One would need longitudinal studies taking into account IMT progression and changes in addresses. Only one longitudinal study of IMT progression has been published so far, confirming a faster progression of IMT in people with higher exposures to traffic proximity and $PM_{2.5}$ (Künzli et al., 2010), which makes reverse causation a less likely explanation. Last but not least it has to be re-emphasized that several animal studies have consistently shown that ambient PM and particularly UFP accelerates the development of atherosclerosis (Araujo et al., 2008; Sun et al., 2005).

9.4 Implications

9.4.1 For public health

The whole population is exposed to traffic-related air pollution, to a higher or lower extent. Our findings in a healthy population support that people living closer to streets with high traffic intensity are at increased risk of atherosclerosis due to traffic-related pollution. In the Mediterranean population, this risk is particularly higher for people with high education level and men older than 60. Thus, effect modification factors play an important role in the association between traffic related air pollution and atherosclerosis, but are not yet fully elucidated and might differ across populations.

Increased traffic intensity in the nearest street and in a 100m around the residence were associated to a higher degree of subclinical atherosclerosis. Reducing traffic levels particularly in areas of high population density are thus expected to have substantial health benefits. An alternative policy would be to replace the current vehicle fleet with "zero-emissions" vehicles to reduce pollution in such areas. To estimate the impact of such policies, a better understanding of the independent health effects of emissions from tail-pipe pollutants, resuspended dust and brake and tire wearing is needed.

Bus lines in the nearest street were the main predictors of residential outdoor heavy metals in Girona city. While public transportation is an alternative to reduce traffic intensity and related air pollution, the type of fuel used and composition and quantity of emissions generated by public transportation vehicles should be carefully evaluated and translated into policies.

Our results should be confirmed with other studies, and particularly including several markers of air pollution (traffic intensity, elemental carbon, UFP, heavy metals). If these results reflect a true effect it is important to translate those into measurements that are relevant for a person (e.g. risk of developing a CVD event); to quantify the risks at the population level that are attributable to the long-term exposure to air pollution, i.e. how prevalent is living near traffic and what would be the reduction in risk if those people were not exposed to high levels of traffic-related pollution; and to quantify the reduction in risk resulting from the shift of the population distribution of IMT or ABI by a given percent.

The preventive measures to reduce the exposure to air pollution that could be achieved at the individual level are rather limited. Thus there is need for public health measures addressed to the general population, and in particular to people living in urban areas with high traffic intensity.

9.4.2 For the scientific community: open research questions

Several gaps in the current knowledge on long-term exposure to traffic-related air pollution call for further research in the following areas:

- Validation studies of long-term traffic indicators -particularly traffic intensity within buffers- against personal exposure to pollutants derived from traffic are needed.

- Identification of the shortest monitoring time period needed to adequately capture the long-term distribution of the UFP remains unexplored.

- Verifying the association of other traffic markers with subclinical atherosclerosis and further investigation of the susceptibility factors that may modify those associations. The REGICOR-Air will continue to address this question for the Mediterranean population. It is planned to investigate the association between exposure to UFP based on short-term monitoring and subclinical atherosclerosis; as well as conduct statistical analyses of indoor heavy metals and indoor and outdoor PAH, also measured with the moss monitors and benzene, toluene, ethylbenzene and xylenes (BTEX) measured in parallel. With a similar methodology to the one used in this thesis the association of this markers of exposure to traffic-related pollution with subclinical atherosclerosis will be investigated.

- Other health effects of long-term exposure to traffic related pollution need to be investigated jointly, such as the effect of noise on hypertension and atherosclerosis. REGICOR-Air will investigate the prevalence and incidence of hypertension associated with traffic-related noise, independently of the effects of air pollution.

Results from these studies will shed more light on to health effects of traffic-related pollution as well as disentangle potential confounding between traffic-related pollutants and noise.

Banyoles Lake, Banyoles.

10 CONCLUSIONS

1. Mosses can be used to measure the distribution of metals at home outdoors within cities. In Girona city, heavy metals showed higher spatial variability than NO_2 and a significant percentage of its distribution was explained by the number of bus lines passing in the nearest street. For long-term exposure studies, heavy metals measured by mosses should be considered an alternative to PM and complementary to NO_2 monitoring.

2. The use of direct measures of traffic intensity is a very valuable approach to derive crucial information on traffic-related pollution. Those measures may serve as direct markers of exposure as well as key co-variates in models of traffic related pollutants. Traffic intensity road network databases can be derived based on short-term traffic count protocols.

3. Distribution of UFP can be modeled with fair performance based on short-term mobile monitoring. Complementing the approach with further strategies to address sources of temporal variation of UFP is likely to result in improved models as indicated by the good performance of the model for the subset of sites with one repeated measurement. Far shorter monitoring protocols than currently assumed to be needed provide very promising models at lower costs.

4. Models based on a small number of sampling sites result in artificially high adjusted R^2 and leave-one-out cross-validation R^2 compared to the "true" prediction ability (validation in an independent set). The selection of predictor variables for the LUR model might result in biased health effect estimates. LUR models for complex urban settings should be based on a large number of measurement sites (>80) and the number of potential predictors variables should be restricted.

5. Long-term exposure to traffic-related air pollution is associated with carotid subclinical atherosclerosis with stronger associations in people with high education level and men above 60 years.

Blanes Photo: Josep Mares Huerta

REFERENCES

Abernethy, R., Brauer, M., Allen, R., 2011. Developing a land use regression model for ultrafine particle concentrations in Vancouver, Canada, 23rd Annual Conference of the International Society of Environmental Epidemiology (ISEE). Environ Health Perspect., Barcelona, Spain.

Adamo, P., Giordano, S., Vingiani, S., Castaldo Cobianchi, R., Violante, P., 2003. Trace element accumulation by moss and lichen exposed in bags in the city of Naples (Italy). Environ. Pollut. 122, 91-103.

Agarwal, S., 2009. The Association of Active and Passive Smoking With Peripheral Arterial Disease: Results From NHANES 1999-2004. Angiology 60, 335-345.

Alexeeff, S.E., Coull, B.A., Gryparis, A., Suh, H., Sparrow, D., Vokonas, P.S., Schwartz, J., 2011. Medium-Term Exposure to Traffic-Related Air Pollution and Markers of Inflammation and Endothelial Function. Environ Health Perspect 119, 481-486.

Allen, R.W., Criqui, M.H., Diez Roux, A.V., Allison, M., Shea, S., Detrano, R., Sheppard, L., Wong, N.D., Stukovsky, K.H., Kaufman, J.D., 2009. Fine Particulate Matter Air Pollution, Proximity to Traffic, and Aortic Atherosclerosis. Epidemiology 20, 254-264

Allison, M.A., Hiatt, W.R., Hirsch, A.T., Coll, J.R., Criqui, M.H., 2008. A High Ankle-Brachial Index Is Associated With Increased Cardiovascular Disease Morbidity and Lower Quality of Life. Journal of the American College of Cardiology 51, 1292-1298.

Angiolillo, D.J., Biasucci, L.M., Liuzzo, G., Crea, F., 2004. Inflammation in Acute Coronary Syndromes: Mechanisms and Clinical Implications. Revista Española de Cardiología 57, 433-446.

Anicic, M., Tasic, M., Frontasyeva, M.V., Tomasevic, M., Rajsic, S., Mijic, Z., Popovic, A., 2009a. Active moss biomonitoring of trace elements with Sphagnum girgensohnii moss bags in relation to atmospheric bulk deposition in Belgrade, Serbia. Environ. Pollut. 157, 673-679.

Anicic, M., Tomasevic, M., Tasic, M., Rajsic, S., Popovic, A., Frontasyeva, M.V., Lierhagen, S., Steinnes, E., 2009b. Monitoring of trace element atmospheric deposition using dry and wet moss bags: Accumulation capacity versus exposure time. J. Hazard. Mater. 171, 182-188.

Ankle Brachial Index Collaboration, 2008. Ankle Brachial Index Combined With Framingham Risk Score to Predict Cardiovascular Events and Mortality. JAMA: The Journal of the American Medical Association 300, 197-208.

Araujo, J.A., Barajas, B., Kleinman, M., Wang, X., Bennett, B.J., Gong, K.W., Navab, M., Harkema, J., Sioutas, C., Lusis, A.J., Nel, A.E., 2008. Ambient Particulate Pollutants in the Ultrafine Range Promote Early Atherosclerosis and Systemic Oxidative Stress. Circulation Research 102, 589-596.

Araujo, J.A., Nel, A.E., 2009. Particulate matter and atherosclerosis: role of particle size, composition and oxidative stress. Particle and Fibre Toxicology 6.

Basagaña, X., Rivera, M., Agis, D., Aguilera, I., Foraster, M., de Nazelle, A., Künzli, N., 2011. Effect of the number of measurement sites on Land Use Regression (LUR) models of air pollutants. Abstract 667, 23rd Annual Conference of the International Society of Environmental Epidemiology (ISEE). Environ Health Perspect., Barcelona, Spain.

Basagaña, X., Rivera, M., Aguilera, I., Agis, D., Bouso, L., Elosua, R., Foraster, M., de Nazelle, A., Nieuwenhuijsen, M., Vila, J., Künzli, N., In press. Effect of the number of measurement sites on Land Use Regression Models in estimating local air pollution. Atmospheric Environment.

Bauer, M., Moebus, S., Mohlenkamp, S., Dragano, N., Nonnemacher, M., Fuchsluger, M., Kessler, C., Jakobs, H., Memmesheimer, M., Erbel, R., Jockel, K.-H., Hoffmann, B., on behalf of the, H.N.R.S.I.G., 2010. Urban Particulate Matter Air Pollution Is Associated With Subclinical Atherosclerosis: Results From the HNR (Heinz Nixdorf Recall) Study. J Am Coll Cardiol 56, 1803-1808.

Baxter, L.K., Wright, R.J., Paciorek, C.J., Laden, F., Suh, H.H., Levy, J.I., 2009. Effects of exposure measurement error in the analysis of health effects from traffic-related air pollution. J Expos Sci Environ Epidemiol 20, 101-111.

Beelen, R., de Hoogh, K., Eeftens, M., Meliefste, K., Cirach, M., de Nazelle, A., Nieuwenhuijsen, M., Molter, A., Cyrys, J., Birk, M., Bellander, T., Sugiri, D., Tsai, M.-Y., Ineichen, A., et al., 2011. Estimating Long-term Exposure to Air Pollution in 38 Study Areas in Europe in a Harmonized Way Using Land Use Regression Modeling (ESCAPE Project). Epidemiology 22, S82.

Beelen, R., Hoek, G., Fischer, P., van den Brandt, P.A., Brunekreef, B., 2007a. Estimated long-term outdoor air pollution concentrations in a cohort study. Atmospheric Environment 41, 1343-1358.

Beelen, R., Hoek, G., van den Brandt, P.A., Goldbohm, R.A., Fischer, P., Schouten, L.J., Jerrett, M., Hughes, E., Armstrong, B., Brunekreef, B., 2007b. Long-Term Effects of Traffic-Related Air Pollution on Mortality in a Dutch Cohort (NLCS-AIR Study). Environ Health Perspect 116.

Bell, M.L., Ebisu, K., Peng, R.D., Samet, J.M., Dominici, F., 2009. Hospital Admissions and Chemical Composition of Fine Particle Air Pollution. Am. J. Respir. Crit. Care Med. 179, 1115-1120.

Berg, T., Røyset, O., Steinnes, E., 1995. Moss (Hylocomium splendens) used as biomonitor of atmospheric trace element deposition: Estimation of uptake efficiencies. Atmos. Environ. 29, 353-360.

Bloemen, H.J.T., Balvers, T.T.M., van Scheindelen, H.J., Lebret, E., Oosterlee, A., Drijver, M., 1993. The Benzene Research in South of Kennemerland, The Netherlands, Het benzeen onderzoek Zuid-Kennemerland, . National Institute of Public Health and the Environment (RIVM) and Municipal Health Service Haarlem (GGD), Bilthoven.

Borm, P.J.A., 2008. Future interactions in Particle Toxicology: the role of PFT. Particle and Fibre Toxicology 5.

Bots, M.L., Grobbee, D.E., 2002. Intima Media Thickness as a Surrogate Marker for Generalised Atherosclerosis. Cardiovascular Drugs and Therapy 16, 341-351.

Brauer, M., 2010. How Much, How Long, What, and Where. Proceedings of the American Thoracic Society *7*, 111-115.

Brauer, M., Hoek, G., van Vliet, P., Meliefste, K., Fischer, P., Gehring, U., Heinrich, J., Cyrys, J., Bellander, T., Lewne, M., Brunekreef, B., 2003. Estimating Long-Term Average Particulate Air Pollution Concentrations: Application of Traffic Indicators and Geographic Information Systems. Epidemiology 14, 228-239.

Briggs, D., 2005. The Role of Gis: Coping With Space (And Time) in Air Pollution Exposure Assessment. Journal of Toxicology and Environmental Health, Part A 68, 1243-1261.

Brook, R.D., Franklin, B., Cascio, W., Hong, Y., Howard, G., Lipsett, M., Luepker, R., Mittleman, M., Samet, J., Smith, S.C., Tager, I., 2004. Air Pollution and Cardiovascular Disease. Circulation 109, 2655-2671.

Brook, R.D., Rajagopalan, S., Pope, C.A., III, Brook, J.R., Bhatnagar, A., Diez-Roux, A.V., Holguin, F., Hong, Y., Luepker, R.V., Mittleman, M.A., Peters, A., Siscovick, D., Smith, S.C., Jr., Whitsel, L., et al., 2010. Particulate Matter Air Pollution and Cardiovascular Disease: An Update to the Scientific Statement From the American Heart Association. Circulation 121, 2331-2378.

Brunekreef, B., Forsberg, B., 2005. Epidemiological evidence of effects of coarse airborne particles on health. Eur. Respir. J. 26, 309-318.

Cao, T., Wang, M., An, L., Yu, Y., Lou, Y., Guo, S., Zuo, B., Liu, Y., Wu, J., Cao, Y., Zhu, Z., 2009. Air quality for metals and sulfur in Shanghai, China, determined with moss bags. Environ. Pollut. 157, 1270-1278.

Carnow, B.W., Lepper, M.H., Shekelle, R.B., Stamler, J., 1969. Chicago air pollution study. SO2 levels and acute illness in patients with chronic bronchopulmonary disease. Archives of Environmental Health 18, 768-776.

Carroll, R.J., Ruppert, D., Stefanski, L.A., Crainiceanu, C.M., 2006. Measurement error in nonlinear models a modern perspective, 2nd ed. Chapman & Hall/CRC, Boca Raton, FL.

Cavalcante, J.L., Lima, J.A.C., Redheuil, A., Al-Mallah, M.H., 2011. Aortic stiffness: Current understanding and future directions. Journal of the American College of Cardiology *57*, 1511-1522.

Cesaroni, G., Badaloni, C., Romano, V., Donato, E., Perucci, C., Forastiere, F., 2010. Socioeconomic position and health status of people who live near busy roads: the Rome Longitudinal Study (RoLS). Environmental Health 9, 41.

Coll, B., Feinstein, S.B., 2008. Carotid intima-media thickness measurements: techniques and clinical relevance. Current Atherosclerosis Reports 10, 444-450.

Chaix, B., Gustafsson, S., Jerrett, M., Kristersson, H.k., Lithman, T., Boalt, Ã.k., Merlo, J., 2006. Children's exposure to nitrogen dioxide in Sweden: investigating environmental injustice in an egalitarian country. Journal of Epidemiology and Community Health 60, 234-241.

Chen, L.C., Lippmann, M., 2009. Effects of Metals within Ambient Air Particulate Matter (PM) on Human Health. Inhal. Toxicol. 21, 1-31.

Davison, A.C., Hinkley, D.V., 1997. Bootstrap methods and their application. Cambridge University Press, Cambridge ; New York, NY, USA.

de Groot, E., van Leuven, S.I., Duivenvoorden, R., Meuwese, M.C., Akdim, F., Bots, M.L., Kastelein, J.J.P., 2008. Measurement of carotid intima-media thickness to assess progression and regression of atherosclerosis. Nat Clin Pract Cardiovasc Med 5, 280-288.

de Kok, T., Driece, H.A.L., Hogervorst, J.G.F., Briede, J.J., 2006. Toxicological assessment of ambient and traffic-related particulate matter: A review of recent studies. Mutation Research-Reviews in Mutation Research 613, 103-122.

Deltell, A., 2005. Elaboració del mapa acústic de trànsit de la ciutat de Girona (in Catalan). University of Girona. Spain.

Derksen, S., Keselman, H.J., 1992. Backward, forward and stepwise automated subset selection algorithms: Frequency of obtaining authentic and noise variables. British Journal of Mathematical and Statistical Psychology 42, 265-282.

DGT, 2009. Direccion General de Trafico. Spain.

Diapouli, E., Chaloulakou, A., Spyrellis, N., 2007. Levels of ultrafine particles in different microenvironments--implications to children exposure. Sci Total Environ 388, 128-136.

Diez Roux, A.V., Auchincloss, A.H., Franklin, T.G., Raghunathan, T., Barr, R.G., Kaufman, J., Astor, B., Keeler, J., 2008. Long-term Exposure to Ambient Particulate Matter and Prevalence of Subclinical Atherosclerosis in the Multi-Ethnic Study of Atherosclerosis. American Journal of Epidemiology 167, 667-675.

Dockery, D.W., 2009. Health Effects of Particulate Air Pollution. Ann. Epidemiol. 19, 257-263.

Dockery, D.W., Pope, C.A., Xu, X., Spengler, J.D., Ware, J.H., Fay, M.E., Ferris, B.G., Speizer, F.E., 1993. An Association between Air Pollution and Mortality in Six U.S. Cities. New England Journal of Medicine 329, 1753-1759.

Dockery, D.W., Spengler, J.D., 1981. Personal exposure to respirable particulates and sulfates. Journal of the Air Pollution Control Association 31, 153-159.

Dominici, F., McDermott, A., Daniels, D., 2003. Mortality among residents of 90 cities. Special Report: Revised Analyses of Time-Series Studies of Air Pollution and Health, 9-24.

Efron, B., Tibshirani, R., Tibshirani, R.J., 1993. An introduction to the bootstrap. Chapman & Hall, New York.

Elosua, R., Marrugat, J., Molina, L., Pons, S., Pujol, E., 1994. Validation of the Minnesota Leisure Time Physical Activity Questionnaire in Spanish Men. American Journal of Epidemiology 139, 1197-1209.

Erdogmus, B., Yazici, B., Annakkaya, A.N., Bilgin, C., Safak, A.A., Arbak, P., Tugay, A., 2006. Intima-media thickness of the common carotid artery in highway toll collectors. J Clin Ultrasound 34, 430-433.

Erdogmus, B., Yazici, B., Onder, H.I., Bilgin, C., Tugay, A., Buyukkaya, R., 2007. Investigation of the effect of exhaust particles on ocular blood flow in highway toll collectors using color Doppler ultrasonography. Acta Radiol 48, 223-227.

Fernandez, J.A., Aboal, J.R., Carballeira, A., 2004. Identification of pollution sources by means of moss bags. Ecotoxicol. Environ. Saf. 59, 76-83.

Fernández, J.Á., Aboal, J.R., Real, C., Carballeira, A., 2007. A new moss biomonitoring method for detecting sources of small scale pollution. Atmos. Environ. 41, 2098-2110.

Flack, V.F., Chang, P.C., 1987. Frequency of Selecting Noise Variables in Subset Regression Analysis: A Simulation Study. The American Statistician 41, 84-86.

Foraster, M., Deltell, A., Basagaña, X., Medina-Ramón, M., Aguilera, I., Bouso, L., Grau, M., Phuleria, H.C., Rivera, M., Slama, R., Sunyer, J., Targa, J., Künzli, N., 2011. Local determinants of road traffic noise levels versus determinants of air pollution levels in a Mediterranean city. Environmental Research 111, 177-183.

Gehring, U., Heinrich, J., Krämer, U., Grote, V., Hochadel, M., Sugiri, D., Kraft, M., Rauchfuss, K., Eberwein, H.G., Wichmann, H.E., 2006. Long-Term Exposure to Ambient Air Pollution and Cardiopulmonary Mortality in Women. Epidemiology 17, 545-551.

Goodman, G.T., Roberts, T.M., 1971. Plants and Soils as Indicators of Metals in the Air. Nature 231, 287-292.

Grau, M., Marrugat, J., 2008. Risk Functions and the Primary Prevention of Cardiovascular Disease. Revista Española de Cardiología (English Edition) 61, 404-416.

Grau, M., Subirana, I., Elosua, R., Solanas, P., Ramos, R., Masiá, R., Cordon, F., Sala, J., Juvinyá , D., Cerezo, C., Fitó, M., Vila, J., Covas, M.I., Marrugat, J., et al., 2007. Trends in cardiovascular risk factor prevalence (1995-2000-2005) in northeastern Spain. European Journal of Cardiovascular Prevention & Rehabilitation 14, 653-659. Erratum in: Eur J Cardiovasc Prev Rehabil, 2008 Aug;2015(2004):2502.

Greenland, S., 1993. Basic problems in interaction assessment. Environ Health Perspect 101.

Guallar-Castillón, P., Rodriguez-Artalejo, F., Tormo, M.J., Sánchez, M.J., Rodriguez, L., Quirós, J.R., Navarro, C., Molina, E., MartÃnez, C., Marín, P., Lopez-Garcia, E., Larrañaga, N., Huerta, J.M., Dorronsoro, M., et al., 2011. Major dietary patterns and risk of coronary heart disease in middle-

aged persons from a Mediterranean country: The EPIC-Spain cohort study. Nutrition, Metabolism and Cardiovascular Diseases.

Gunier, R.B., Hertz, A., von Behren, J., Reynolds, P., 2003. Traffic density in California: Socioeconomic and ethnic differences among potentially exposed children. J Expo Anal Environ Epidemiol 13, 240-246.

Hagler, G.S.W., Baldauf, R.W., Thoma, E.D., Long, T.R., Snow, R.F., Kinsey, J.S., Oudejans, L., Gullett, B.K., 2009. Ultrafine particles near a major roadway in Raleigh, North Carolina: Downwind attenuation and correlation with traffic-related pollutants. Atmospheric Environment 43, 1229-1234.

Hansson, G.r.K., 2005. Inflammation, Atherosclerosis, and Coronary Artery Disease. New England Journal of Medicine 352, 1685-1695.

Harmens, H., 2004. Report of the Kalamata-meeting, UNECE convention on long-range transboundary air pollution, working group on effects, February 2004, Kalamata, Greece.

Harrell, F.E., 2001. Regression modeling strategies : with applications to linear models, logistic regression, and survival analysis. Springer, New York.

Harrell, F.E., Jr., Lee, K.L., Mark, D.B., 1996. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med 15, 361-387.

Harrison, R.M., Deacon, A.R., 1998. Spatial correlation of automatic air quality monitoring at urban background sites: Implications for network design. Environ. Technol. 19, 121-132.

Hastie, T., Tibshirani, R., Friedman, J.H., 2001. The elements of statistical learning : data mining, inference, and prediction. Springer, New York.

Havard, S., Deguen, S.v., Zmirou-Navier, D., Schillinger, C., Bard, D., 2009. Traffic-Related Air Pollution and Socioeconomic Status: A Spatial Autocorrelation Study to Assess Environmental Equity on a Small-Area Scale. Epidemiology 20, 223-230

Hawkins, D.M., 2004. The problem of overfitting. Journal of Chemical Information and Computer Sciences 44, 1-12.

Hawkins, D.M., Basak, S.C., D., M., 2003. Assessing model fit by cross-validation. Journal of Chemical Information and Computer Sciences 43, 579-586.

He, Y., Lam, T.H., Jiang, B., Wang, J., Sai, X., Fan, L., Li, X., Qin, Y., Hu, F.B., 2008. Passive Smoking and Risk of Peripheral Arterial Disease and Ischemic Stroke in Chinese Women Who Never Smoked. Circulation 118, 1535-1540.

HEI Panel on the Health Effects of Traffic-Related Air Pollution, 2010. Traffic-Related Air Pollution: A Critical Review of the Literature on Emissions, Exposure, and Health Effects. HEI Special Report 17. Health Effects Institute, Boston, MA.

Henderson Y, H.H., Teague MC, Prince AL, Wunderlich RM, 1921. Physiological effects of automobile exhaust gas and standards of ventilation for brief exposures. J Ind Hyg III(3):79-92.

Hobbs, S.D., Wilmink, A.B.M., Adam, D.J., Bradbury, A.W., 2005. Assessment of smoking status in patients with peripheral arterial disease. Journal of Vascular Surgery 41, 451-456.

Hoek, G., Beelen, R., de Hoogh, K., Vienneau, D., Gulliver, J., Fischer, P., Briggs, D., 2008a. A review of land-use regression models to assess spatial variation of outdoor air pollution. Atmospheric Environment 42, 7561–7578.

Hoek, G., Beelen, R., Kos, G., Dijkema, M., van der Zee, S.C., Fischer, P.H., Brunekreef, B., 2010. Land use regression model for ultrafine particles in Amsterdam. Environ Sci Technol 45, 622-628.

Hoek, G., Boogaard, H., Knol, A., de Hartog, J., Slottje, P., Ayres, J.G., Borm, P., Brunekreef, B., Donaldson, K., Forastiere, F., Holgate, S., Kreyling, W.G., Nemery, B., Pekkanen, J., et al., 2009. Concentration Response Functions for Ultrafine Particles and All-Cause Mortality and Hospital Admissions: Results of a European Expert Panel Elicitation. Environmental Science & Technology 44, 476-482.

Hoek, G., Brunekreef, B., Goldbohm, S., Fischer, P., van den Brandt, P.A., 2002. Association between mortality and indicators of traffic-related air pollution in the Netherlands: a cohort study. The Lancet 360, 1203-1209.

Hoek, G., Kos, G., Harrison, R., de Hartog, J., Meliefste, K., ten Brink, H., Katsouyanni, K., Karakatsani, A., Lianou, M., Kotronarou, A., Kavouras, I., Pekkanen, J., Vallius, M., Kulmala, M., et al., 2008b. Indoor-outdoor relationships of particle number and mass in four European cities. Atmospheric Environment 42, 156-169.

Hoffmann, B., Moebus, S., Dragano, N., Stang, A., Möhlenkamp, S., Schmermund, A., Memmesheimer, M., Bröcker-Preuss, M., Mann, K., Erbel, R., Jöckel, K.H., 2009a. Chronic residential exposure to particulate matter air pollution and systemic inflammatory markers. Environ Health Perspect 117, 1302-1308.

Hoffmann, B., Moebus, S., Kröger, K., Stang, A., Möhlenkamp, S., Dragano, N., Schmermund, A., Memmesheimer, M., Erbel, R., Jöckel, K.-H., 2009b. Residential Exposure to Urban Air Pollution, Ankle-Brachial Index, and Peripheral Arterial Disease. Epidemiology 20, 280-288.

Hoffmann, B., Moebus, S., Möhlenkamp, S., Stang, A., Lehmann, N., Dragano, N., Schmermund, A., Memmesheimer, M., Mann, K., Erbel, R., Jöckel, K.H., for the Heinz Nixdorf Recall Study Investigative, G., 2007. Residential Exposure to Traffic Is Associated With Coronary Atherosclerosis. Circulation 116, 489-496.

Holewijn, S., den Heijer, M., Stalenhoef, A.F.H., de Graaf, J., 2010. Non-invasive measurements of atherosclerosis (NIMA): Current evidence and future perspectives. Netherlands Journal of Medicine 68, 388-399.

Holland, D.M., 1983. Carbon monoxide levels in microenvironment types of four U.S. cities. Environment International 9, 369-377.

Horie, Y., Stern, A.C., 1976. Analysis of Population Exposure to Air Pollution in the New York-New Jersey-Connecticut Tri-State Region, Report No. EPA-450/3-76-027, in: Agency, U.S.E.P. (Ed.), Research Triangle Park, NC.

Iannuzzi, A., Verga, M.C., Renis, M., Schiavo, A., Salvatore, V., Santoriello, C., Pazzano, D., Licenziati, M.R., Polverino, M., 2010. Air pollution and carotid arterial stiffness in children. Cardiol Young 20, 186-190.

Ibald-Mulli, A., Wichmann, H.E., Kreyling, W., Peters, A., 2002. Epidemiological evidence on health effects of ultrafine particles. J Aerosol Med 15, 189-201.

Iglseder, B., Cip, P., Malaimare, L., Ladurner, G., Paulweber, B., 2005. The Metabolic Syndrome Is a Stronger Risk Factor for Early Carotid Atherosclerosis in Women Than in Men. Stroke 36, 1212-1217.

INE, 2004. Instituto Nacional de Estadística. Population and Housing Census 2001.

INE, 2011. Instituto Nacional de Estadística. Defunciones según la Causa de Muerte. Año 2009.

Isaksson, A., Wallman, M., Göransson, H., Gustafsson, M.G., 2008. Cross-validation and bootstrapping are unreliable in small sample classification. Pattern Recognition Letters 29, 1960-1965.

Jerrett, M., Arain, A., Kanaroglou, P., Beckerman, B., Potoglou, D., Sahsuvaroglu, T., Morrison, J., Giovis, C., 2004. A review and evaluation of intraurban air pollution exposure models. J Expo Anal Environ Epidemiol 15, 185-204.

Jerrett, M., Burnett, R.T., Kanaroglou, P., Eyles, J., Finkelstein, N., Giovis, C., Brook, J.R., 2001. A GIS - environmental justice analysis of particulate air pollution in Hamilton, Canada. Environment and Planning A 33, 955-973.

Johnson, M., Isakov, V., Touma, J.S., Mukerjee, S., Özkaynak, H., 2010. Evaluation of land-use regression models used to predict air quality concentrations in an urban area. Atmospheric Environment 44, 3660-3668.

Kanaroglou, P.S., Jerrett, M., Morrison, J., Beckerman, B., Arain, M.A., Gilbert, N.L., Brook, J.R., 2005. Establishing an air pollution monitoring network for intra-urban population exposure assessment: A location-allocation approach. Atmospheric Environment 39, 2399-2409.

Kassomenos, P., Karakitsios, S., Papaloukas, C., 2006. Estimation of daily traffic emissions in a South-European urban agglomeration during a workday. Evaluation of several "what if" scenarios. Science of the Total Environment 370, 480-490.

Katsouyanni, K., Touloumi, G., Samoli, E., Gryparis, A., Le Tertre, A., Monopolis, Y., Rossi, G., Zmirou, D., Ballester, F., Boumghar, A., Anderson, H.R., Wojtyniak, B., Paldy, A., Braunstein, R., et al., 2001. Confounding and effect modification in the short-term effects of ambient particles on total mortality: Results from 29 European cities within the APHEA2 project. Epidemiology 12, 521-531.

Krämer, U., Koch, T., Ranft, U., Ring, J., Behrendt, H., 2000. Traffic-Related Air Pollution Is Associated with Atopy in Children Living in Urban Areas. Epidemiology 11, 64-70.

Künzli, N., Jerrett, M., Garcia-Esteban, R., Basagaña, X., Beckermann, B., Gilliland, F., Medina, M., Peters, J., Hodis, H.N., Mack, W.J., 2010. Ambient Air

Pollution and the Progression of Atherosclerosis in Adults. PLoS ONE 5(2):, e9096.

Künzli, N., Jerrett, M., Mack, W.J., Beckerman, B., LaBree, L., Gilliland, F., Thomas, D., Peters, J., Hodis, H.N., 2005a. Ambient air pollution and atherosclerosis in Los Angeles. Environmental health perspectives 113, 201-206.

Künzli, N., Lurmann, F., Segal, M., Ngo, L., Balmes, J., Tager, I.B., 1997. Association between Lifetime Ambient Ozone Exposure and Pulmonary Function in College Freshmenâ€"Results of a Pilot Study. Environmental Research 72, 8-23.

Künzli, N., Perez, L., von Klot, S., Baldassarre, D., Bauer, M., Basagana, X., Breton, C., Dratva, J., Elosua, R., de Faire, U., Fuks, K., de Groot, E., Marrugat, J., Penell, J., et al., 2011. Investigating Air Pollution and Atherosclerosis in Humans: Concepts and Outlook. Progress in Cardiovascular Diseases 53, 334-343.

Künzli, N., Tager, I.B., 2005b. Air pollution: from lung to heart. Swiss Med Wkly 135, 697-702.

Lambert, W.E., Samet, J.M., Stidley, C.A., 1992. Classification of residential exposure to nitrogen dioxide. Atmospheric Environment - Part A General Topics 26 A, 2185-2192.

Larson, T., Henderson, S.B., Brauer, M., 2009. Mobile Monitoring of Particle Light Absorption Coefficient in an Urban Area as a Basis for Land Use Regression. Environmental Science & Technology 43, 4672-4678.

Lenters, V., Uiterwaal, C.S., Beelen, R., Bots, M.L., Fischer, P., Brunekreef, B., Hoek, G., 2010a. Long-Term Exposure to Air Pollution and Vascular Damage in Young Adults. Epidemiology 21, 512-520

Lewné, M., Cyrys, J., Meliefste, K., Hoek, G., Brauer, M., Fischer, P., Gehring, U., Heinrich, J., Brunekreef, B., Bellander, T., 2004. Spatial variation in nitrogen dioxide in three European areas. Science of The Total Environment 332, 217-230.

Li, N., Sioutas, C., Cho, A., Schmitz, D., Misra, C., Sempf, J., Wang, M., Oberley, T., Froines, J., Nel, A., 2003. Ultrafine particulate pollutants induce oxidative stress and mitochondrial damage. Environ. Health Perspect. 111, 455-460.

Lipsett, M.J., Tsai, F.C., Roger, L., Woo, M., Ostro, B.D., 2006. Coarse particles and heart rate variability among older adults with coronary artery disease in the Coachella Valley, California. Environ. Health Perspect. 114, 1215-1220.

Lonn, E.M., Gerstein, H.C., Sheridan, P., Smith, S., Diaz, R., Mohan, V., Bosch, J., Yusuf, S., Dagenais, G.R., Investigators, D.a.S., 2009. Effect of Ramipril and of Rosiglitazone on Carotid Intima-Media Thickness in People With Impaired Glucose Tolerance or Impaired Fasting Glucose: STARR (STudy of Atherosclerosis with Ramipril and Rosiglitazone). J Am Coll Cardiol 53, 2028-2035.

Lorenz, M.W., Markus, H.S., Bots, M.L., Rosvall, M., Sitzer, M., 2007. Prediction of Clinical Cardiovascular Events With Carotid Intima-Media Thickness. Circulation 115, 459-467.

Lusis, A.J., 2000. Atherosclerosis. Nature 407, 233-241.

MacMahon, S., Sharpe, N., Gamble, G., Clague, A., Ni Mhurchu, C., Clark, T., Hart, H., Scott, J., White, H., Group, P.-C.R., 2000. Randomized, placebo-controlled trial of the angiotensin-converting enzyme inhibitor, ramipril, in patients with coronary or other occlusive arterial disease. J Am Coll Cardiol 36, 438-443.

Makholm, M.M., Mladenoff, D.J., 2005. Efficacy of a biomonitoring (moss bag) technique for determining element deposition trends on a mid-range (375 km) scale. Environ. Monit. Assess. 104, 1-18.

Mancini, G.B.J., Dahlöf, B., Díez, J., 2004. Surrogate Markers for Cardiovascular Disease. Circulation 109, IV-22-IV-30.

Martínez-González, M.A., García-López, M., Bes-Rastrollo, M., Toledo, E., Martínez-Lapiscina, E.H., Delgado-Rodriguez, M., Vazquez, Z., Benito, S., Beunza, J.J., 2011. Mediterranean diet and the incidence of cardiovascular disease: A Spanish cohort. Nutrition, Metabolism and Cardiovascular Diseases 21, 237-244.

Masiá, R., Pena, A., Marrugat, J., Sala, J., Vila, J., Pavesi, M., Covas, M., Aubó, C., Elosua, R., 1998. High prevalence of cardiovascular risk factors in Gerona, Spain, a province with low myocardial infarction incidence. REGICOR Investigators. Journal of Epidemiology and Community Health 52, 707-715.

McDermott, M.M., Liu, K., Criqui, M.H., Ruth, K., Goff, D., Saad, M.F., Wu, C., Homma, S., Sharrett, A.R., 2005. Ankle-Brachial Index and Subclinical Cardiac and Carotid Disease. American Journal of Epidemiology 162, 33-41.

McGill, H.C., Jr., McMahan, C.A., 1998. Determinants of atherosclerosis in the young. Pathobiological Determinants of Atherosclerosis in Youth (PDAY) Research Group. Am J Cardiol 82, 30T-36T.

Mendez, M.A., Popkin, B.M., Buckland, G., Schroder, H., Amiano, P., Barricarte, A., Huerta, J.-M., Quirós, J.R., Sánchez, M.-J., González, C.A., 2011. Alternative Methods of Accounting for Underreporting and Overreporting When Measuring Dietary Intake-Obesity Relations. American Journal of Epidemiology 173, 448-458.

Mimoun, L., Massin, P., Steg, G., 2009. Retinal microvascularisation abnormalities and cardiovascular risk. Archives of Cardiovascular Diseases 102, 449-456.

Molinaro, A.M., Simon, R., Pfeiffer, R.M., 2005. Prediction error estimation: a comparison of resampling methods. Bioinformatics 21, 3301-3307.

Monn, C., 2001. Exposure assessment of air pollutants: a review on spatial heterogeneity and indoor/outdoor/personal exposure to suspended particulate matter, nitrogen dioxide and ozone. Atmos. Environ. 35, 1-32.

Monn, C., Brandli, O., Schindler, C., Ackermann-Liebrich, U., Leuenberger, P., 1998. Personal exposure to nitrogen dioxide in Switzerland. Science of The Total Environment 215, 243-251.

Morello-Frosch, R., Jesdale, B.M., 2005. Separate and Unequal: Residential Segregation and Estimated Cancer Risks Associated with Ambient Air Toxics in U.S. Metropolitan Areas. Environ Health Perspect 114.

Moreno, T., Querol, X., Alastuey, A., Viana, M., Salvador, P., de la Campa, A.S., Artinano, B., de la Rosa, J., Gibbons, W., 2006. Variations in atmospheric PM trace metal content in Spanish towns: Illustrating the chemical complexity of the inorganic urban aerosol cocktail. Atmos. Environ. 40, 6791-6803.

Morris, S.C., 1981. Personal monitoring of air pollution exposures. Environment International 5, 69-72.

Moschandreas, D.J., Winchester, J.W., Nelson, J.W., Burton, R.M., 1979. Fine particle residential indoor air pollution. Atmospheric Environment (1967) 13, 1413-1418.

Namekata, T., Carnow, B.W., Reda, D.J., O'Farrell, E., Marselle, J.R., 1979. Model for measuring the health impact from changing levels of ambient air pollution: mortality study. Final report (Chicago).

Naszradi, T., Badacsonyi, A., Keresztenyi, I., Podar, D., Csintalan, Z., Tuba, Z., 2007. Comparison of two metal surveys by moss Tortula ruralis in Budapest, Hungary. Environ. Monit. Assess. 134, 279-285.

Nawrot, T.S., Perez, L., Künzli, N., Munters, E., Nemery, B., 2011. Public health importance of triggers of myocardial infarction: a comparative risk assessment. The Lancet 377, 732-740.

Nel, A.E., Diaz-Sanchez, D., Li, N., 2001. The role of particulate pollutants in pulmonary inflammation and asthma: evidence for the involvement of organic chemicals and oxidative stress. Curr. Opin. Pulm. Med. 7, 20-26.

O'Neill, M.S., Diez-Roux, A.V., Auchincloss, A.H., Shen, M., Lima, J.A., Polak, J.F., Barr, R.G., Kaufman, J., Jacobs, D.R., Jr., 2011. Long-term exposure to airborne particles and arterial stiffness: the Multi-Ethnic Study of Atherosclerosis (MESA). Environ Health Perspect 119, 844-851.

Oosterlee, A., Drijver, M., Lebret, E., Brunekreef, B., 1996. Chronic respiratory symptoms in children and adults living along streets with high traffic density. Occupational and Environmental Medicine 53, 241-247.

Ostro, B., Tobias, A., Querol, X., Alastuey, A., Amato, F., Pey, J., Pérez, N., Sunyer, J., 2011. The Effects of Particulate Matter Sources on Daily Mortality: A Case-Crossover Study of Barcelona, Spain. Environ Health Perspect 119.

Ott, W., Eliassen, R., 1973. A survey technique for determining the representativeness of urban air monitoring stations with respect to carbon monoxide. Journal of the Air Pollution Control Association 23, 685-690.

Ott, W.R., 1982. Concepts of human exposure to air pollution. Environment International 7, 179-196.

Paatero, P., Aalto, P., Picciotto, S., Bellander, T., Castaño, G., Cattani, G., Cyrys, J., Kulmala, M., Lanki, T., Nyberg, F., Pekkanen, J., Peters, A., Sunyer, J., Forastiere, F., 2005. Estimating time series of aerosol particle number concentrations in the five HEAPSS cities on the basis of measured air pollution and meteorological variables. Atmospheric Environment 39, 2261-2273.

Pérez, C., 2001. El sistema estadístico SAS. Prentice Hall, Madrid.

Perez, L., Tobias, A., Querol, X., Künzli, N., Pey, J., Alastuey, A., Viana, M., Valero, N., Gonzalez-Cabre, M., Sunyer, J., 2008. Coarse particles from Saharan dust and daily mortality. Epidemiology 19, 800-807.

Peters, A., Ruckerl, R., Cyrys, J., 2011. Lessons from air pollution epidemiology for studies of engineered nanomaterials. J Occup Environ Med 53, S8-S13.

Peters, A., Wichmann, H.E., Tuch, T., Heinrich, J., Heyder, J., 1997. Respiratory effects are associated with the number of ultrafine particles. Am. J. Respir. Crit. Care Med. 155, 1376-1383.

Pey, J., Querol, X., Alastuey, A., Rodríguez, S., Putaud, J.P., Van Dingenen, R., 2009. Source apportionment of urban fine and ultra-fine particle number concentration in a Western Mediterranean city. Atmospheric Environment 43, 4407-4415.

Phuleria, H.C., Corradi, E., Ineichen, A., Künzli, N., Liu, L.-J.S., 2011. Spatiotemporal variation of particle number concentration in urban area of Basel, Switzerland and relationship with mass measurements, 15th ETH-Conference on Combustion Generated Nanoparticles. June 26th – 29th, 2011.

Pletcher, M.J., Tice, J.A., Pignone, M., Browner, W.S., 2004. Using the coronary artery calcium score to predict coronary heart disease events: A systematic review and meta-analysis. Archives of Internal Medicine 164, 1285-1292.

PMU-Girona, 2007. Pla de mobilitat urbana PMU del municipi de Girona - Urban mobility plan for Girona.

Polak, J.F., 2009. Carotid Intima-Media Thickness: An Early Marker of Cardiovascular Disease. Ultrasound Quarterly 25, 55-61.

Politis M, P.C., Lekkas TD: , 2008. Ultrafine particles (UFP) and health effects. Dangerous. Like no other PM? Review and analysis. Global NEST Journal 10, 439-452

Pope, C.A., 3rd, Burnett, R.T., Thun, M.J., Calle, E.E., Krewski, D., Ito, K., Thurston, G.D., 2002a. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. JAMA 287, 1132-1141.

Pope, C.A., Burnett, R.T., Thun, M.J., Calle, E.E., Krewski, D., Ito, K., Thurston, G.D., 2002b. Lung Cancer, Cardiopulmonary Mortality, and Long-term Exposure to Fine Particulate Air Pollution. JAMA: The Journal of the American Medical Association 287, 1132-1141.

R Development Core Team, 2010. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. http://www.R-project.org/, Vienna, Austria.

Raaschou-Nielsen, O., Hertel, O., Vignati, E., Berkowicz, R., Jensen, S.S., Larsen, V.B., Lohse, C., Olsen, J.H., 2000. An air pollution model for use in epidemiological studies: evaluation with measured levels of nitrogen dioxide and benzene. J Expo Anal Environ Epidemiol 10, 4-14.

Ramos, R., Quesada, M., Solanas, P., Subirana, I., Sala, J., Vila, J., MasiÃ¡, R., Cerezo, C., Elosua, R., Grau, M., Cordan, F., Juvinyá , D., Fitó, M., Isabel Covas, M., et al., 2009. Prevalence of Symptomatic and Asymptomatic Peripheral Arterial Disease and the Value of the Ankle-brachial Index to Stratify Cardiovascular Risk. European Journal of Vascular and Endovascular Surgery 38, 305-311.

Ratcliffe, J.M., 1975. An evaluation of the use of biological indicators in an atmospheric lead survey. Atmos. Environ. 9, 623-629.

Reimann, C., Niskavaara, H., Kashulina, G., Filzmoser, P., Boyd, R., Volden, T., Tomilina, O., Bogatyrev, I., 2001. Critical remarks on the use of terrestrial moss (Hylocomium splendens and Pleurozium schreberi) for monitoring of airborne pollution. Environ Pollut 113, 41-57.

Rencher, A.C., Pun, F.C., 1980. Inflation of R2 in Best Subset Regression. Technometrics 22, 49-53.

Revkin, J.H., Shear, C.L., Pouleur, H.G., Ryder, S.W., Orloff, D.G., 2007. Biomarkers in the prevention and treatment of atherosclerosis: Need, validation, and future. Pharmacological Reviews 59, 40-53.

Rühling, A., Steinnes, E., 1998. Atmospheric Heavy Metal Deposition in Europe 1995-1996. Nordic Council of ministers, Copenhagen, 1998, Copenhagen.

Rühling, A., Tyler, G., 1970. Sorption and Retention of Heavy Metals in Woodland Moss Hylocomium-Splendens (Hedw) Br Et Sch. Oikos 21, 92-&.

Ryan, P.H., LeMasters, G.K., 2007a. A review of land-use regression models for characterizing intraurban air pollution exposure. Inhal Toxicol 19 Suppl 1, 127-133.

Ryan, P.H., Lemasters, G.K., 2007b. A review of land-use regression models for characterizing intraurban air pollution exposure. Inhalation Toxicology 19, 127-133.

Sahsuvaroglu, T., Su, J.G., Brook, J., Burnett, R., Loeb, M., Jerrett, M., 2009. Predicting Personal Nitrogen Dioxide Exposure in an Elderly Population: Integrating Residential Indoor and Outdoor Measurements, Fixed-Site Ambient Pollution Concentrations, Modeled Pollutant Levels, and Timeâ€“Activity Patterns. Journal of Toxicology and Environmental Health, Part A 72, 1520-1533.

Sayers, R.R., Fieldner, A.C., Yant, W.P., Thomas, B.G.H., McConnell, W.J., United, S., 1924. Exhaust gases from engines using ethyl gasoline. U.S. Dept. of the Interior, Bureau of Mines, [Washington, D.C.].

Schröder, H., Marrugat, J., Vila, J., Covas, M.I., Elosua, R., 2004. Adherence to the Traditional Mediterranean Diet Is Inversely Associated with Body Mass Index and Obesity in a Spanish Population. The Journal of Nutrition 134, 3355-3361.

Setton, E., Marshall, J.D., Brauer, M., Lundquist, K.R., Hystad, P., Keller, P., Cloutier-Fisher, D., 2011. The impact of daily mobility on exposure to traffic-related air pollution and health effect estimates. J Expos Sci Environ Epidemiol 21, 42-48.

Sidorenko, V.F., Fel'dman, I.G., Kirillov, G.P., Maksimova, V.A., 1972. Method of modelling in the study of the distribution of automobile exhaust in a residential district. Metod modelirovaniia pri issledovanii rasprostraneniia vybrosov avtotransporta v zhiloÇ zastroÇke. 37, 9-12.

Simkhovich, B.Z., Kleinman, M.T., Kloner, R.A., 2008. Air Pollution and Cardiovascular Injury: Epidemiology, Toxicology, and Mechanisms. J. Am. Coll. Cardiol. 52, 719-726.

Sinisi, S.E., van der Laan, M.J., 2004. Loss-based cross-validated Deletion/Substitution/Addition algorithms in estimation. Technical Report 143, Division of Biostatistics, University of California, Berkeley. URL www.bepress.com/ucbbiostat/paper143/.

Sorensen, M., Loft, S., Andersen, H.V., Raaschou-Nielsen, O., Skovgaard, L.T., Knudsen, L.E., Nielsen, I.V., Hertel, O., 2005. Personal exposure to PM2.5, black smoke and NO2 in Copenhagen: relationship to bedroom and outdoor concentrations covering seasonal variation. J Expo Anal Environ Epidemiol 15, 413-422.

Spengler, J., Schwab, M., Ryan, P.B., Colome, S., Wilson, A.L., Billick, I., Becker, E., 1994. Personal exposure to nitrogen dioxide in the Los Angeles Basin. Air Waste 44, 39-47.

Sterman, J.D., 2002. All models are wrong: reflections on becoming a systems scientist. System Dynamics Review 18, 501-531.

Study Group on Environmental Monitoring, 1977. Environmental Monitoring., National Academy of Sciences, Washington, DC.

Suglia, S.F., Gryparis, A., Wright, R.O., Schwartz, J., Wright, R.J., 2008. Association of Black Carbon with Cognition among Children in a Prospective Birth Cohort Study. American Journal of Epidemiology 167, 280-286.

Sun, Q., Wang, A., Jin, X., Natanzon, A., Duquaine, D., Brook, R.D., Aguinaldo, J.-G.S., Fayad, Z.A., Fuster, V., Lippmann, M., Chen, L.C., Rajagopalan, S., 2005. Long-term Air Pollution Exposure and Acceleration of Atherosclerosis and Vascular Inflammation in an Animal Model. JAMA: The Journal of the American Medical Association 294, 3003-3010.

Szpiro, A.A., Paciorek, C.J., Sheppard, L., 2011a. Does more accurate exposure prediction necessarily improve health effect estimates? Epidemiology 22, 680-685.

Szpiro, A.A., Sheppard, L., Lumley, T., 2011b. Efficient measurement error correction with spatially misaligned data. Biostatistics 12, 610-623.

Targa, J., 2008. "Avaluació de la qualitat de l'aire a la ciutat de Girona – Dades final juny 2007 al juliol 2008", in: Environment, A.E.a. (Ed.), Didcot.

Targa J., Loader A., and The Defra working group on harmonisation of diffusion tubes, 2008. Diffusion tubes for ambient NO2 monitoring: practical guidance for laboratories and users. AEA Energy and Environment.

The New York Academy of Medicine, 1931. Effect of Air Pollution on Health: Report of the Committee on Public Health Relations of The New York Academy of Medicine. Bull N Y Acad Med 7, 751-775.

Tretiach, M., Adamo, P., Bargagli, R., Baruffo, L., Carletti, L., Crisafulli, P., Giordano, S., Modenesi, P., Orlando, S., Pittao, E., 2007. Lichen and moss bags as monitoring devices in urban areas. Part I: influence of exposure on sample vitality. Environ. Pollut. 146, 380-391.

UMAT, 2008. Servidor de Cartografia. Girona City Council.

Urban mobility plan for Girona, 2007. http://www.ajgirona.cat/mobilitat/plamobilitat.html. Girona City Council.

Van Roosbroeck, S., Jacobs, J., Janssen, N.A.H., Hoek, G., Brunekreef, B., 2006. Long-term Personal Exposure to Traffic-related Air Pollution Among School Children: A Validation Study. Epidemiology 17, S36-S37.

Van Roosbroeck, S., Jacobs, J., Janssen, N.A.H., Oldenwening, M., Hoek, G., Brunekreef, B., 2007. Long-term personal exposure to PM2.5, soot and NOx in children attending schools located near busy roads, a validation study. Atmospheric Environment 41, 3381-3394.

Van Roosbroeck, S., Hoek, G., Meliefste, K., Janssen, N.A.H., Brunekreef, B., 2008. Validity of Residential Traffic Intensity as an Estimate of Long-Term Personal Exposure to Traffic-Related Air Pollution among Adults. Environmental Science & Technology 42, 1337-1344.

Vasconcelos, M.T.S.D., Tavares, H.M.F., 1998. Atmospheric metal pollution (Cr, Cu, Fe, Mn, Ni, Pb and Zn) in Oporto city derived from results for low-volume aerosol samplers and for the moss Sphagnum auriculatum bioindicator. The Science of The Total Environment 212, 11-20.

Viskari, E.L., Rekilä, R., Roy, S., Lehto, O., Ruuskanen, J., Kärenlampi, L., 1997. Airborne pollutants along a roadside: Assessment using snow analyses and moss bags. Environ. Pollut. 97, 153-160.

Wang, H.K., Chen, C.H., Huang, C., Fu, L.X., 2008. On-road vehicle emission inventory and its uncertainty analysis for Shanghai, China. Sci. Total Environ. 398, 60-67.

Wang, Y., Hopke, P.K., Chalupa, D.C., Utell, M.J., 2010. Long-term study of urban ultrafine particles and other pollutants. Atmospheric Environment In Press doi:10.1016/j.atmosenv.2010.08.022.

Wattanakit, K., Folsom, A.R., Duprez, D.A., Weatherley, B.D., Hirsch, A.T., 2007. Clinical significance of a high ankle-brachial index: Insights from the Atherosclerosis Risk in Communities (ARIC) Study. Atherosclerosis 190, 459-464.

Wheeler, B.W., Ben-Shlomo, Y., 2005. Environmental equity, air quality, socioeconomic status, and respiratory health: a linkage analysis of routine data from the Health Survey for England. Journal of Epidemiology and Community Health 59, 948-954.

Whipple, M.O., Lewis, T.T., Sutton-Tyrrell, K., Matthews, K.A., Barinas-Mitchell, E., Powell, L.H., Everson-Rose, S.A., 2009. Hopelessness, Depressive Symptoms, and Carotid Atherosclerosis in Women. Stroke 40, 3166-3172.

WHO, 2006. Health risks of particulate matter from long-range transboundary air pollution. , European Centre for Environment and Health, Bonn Office. World Health Organization, Europe.

Wittig, A., Engenhart-Cabillic, R., 2011. Cardiac side effects of conventional and particle radiotherapy in cancer patients. Herz 36, 311-324.

Wood, S.N., 2003. Thin plate regression splines. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 65, 95-114.

World Health Organization, 2006. Air quality guidelines : global update 2005 : particulate matter, ozone, nitrogen dioxide, and sulfur dioxide. World Health Organization, Copenhagen, Denmark

World Health Organization, 2008. The global burden of disease: 2004 update., Geneva.

World Health Organization, 2011. Global Atlas on Cardiovascular Disease Prevention and Control. Mendis S, Puska P, Norrving. B editors., Geneva 2011.

Wu, C.-f., Kuo, I.C., Su, T.-C., Li, Y.-R., Lin, L.-Y., Chan, C.-C., Hsu, S.-C., 2010. Effects of Personal Exposure to Particulate Matter and Ozone on Arterial Stiffness and Heart Rate Variability in Healthy Adults. American Journal of Epidemiology 171, 1299-1309.

Yeatts, K., Svendsen, E., Creason, J., Alexis, N., Herbst, M., Scott, J., Kupper, L., Williams, R., Neas, L., Cascio, W., Devlin, R.B., Peden, D.B., 2007. Coarse particulate matter (PM2.5-10) affects heart rate variability, blood lipids, and circulating eosinophils in adults with asthma. Environ. Health Perspect. 115, 709-714.

Zechmeister, H.G., Dullinger, S., Hohenwallner, D., Riss, A., Hanus-Illnar, A., Scharf, S., 2006a. Pilot study on road traffic emissions (PAHs, heavy metals) measured by using mosses in a tunnel experiment in Vienna, Austria. Environmental Science and Pollution Research 13, 398-405.

Zechmeister, H.G., Grodzinska, K., Szarek-Lukaszewska, G., B.A. Markert, A.M.B.a.H.G.Z., 2003. Chapter 10 Bryophytes, Trace Metals and other Contaminants in the Environment. Elsevier, pp. 329-375.

Zechmeister, H.G., Hohenwallner, D., 2006b. A comparison of biomonitoring methods for the estimation of atmospheric pollutants in an industrial town in Austria. Environ. Monit. Assess. 117, 245-259.

Zhang, Y., Stedman, D.H., Bishop, G.A., Guenther, P.L., Beaton, S.P., 1995. Worldwide on-Road Vehicle Exhaust Emissions Study by Remote-Sensing. Environ. Sci. Technol. 29, 2286-2294.

Zhou, Y., Levy, J., 2007. Factors influencing the spatial extent of mobile source air pollution impacts: a meta-analysis. BMC Public Health 7, 89.

Zhu, Y., Kuhn, T., Mayo, P., Hinds, W.C., 2006. Comparison of Daytime and Nighttime Concentration Profiles and Size Distributions of Ultrafine Particles near a Major Highway. Environmental Science & Technology 40, 2531-2536.

Zhu, Y., Zhang, Y., Ling, W., Feng, D., Wei, X., Yang, C., Ma, J., 2011. Fruit Consumption Is Associated with Lower Carotid Intima-Media Thickness and C-Reactive Protein Levels in Patients with Type 2 Diabetes Mellitus. Journal of the American Dietetic Association 111, 1536-1542.

Appendix 1. Air quality standards of the European Union.

Compound	Limit value	WHO Recommended Value	EU Value	Entry into force	Comments
Particulate matter (PM_{10})	Annual average	20 µg/m3 †	40 µg/m3	2005	
	Daily average	50 µg/m3 †	50 µg/m3	2005	May be exceeded up to 35 days a year
Fine particles ($PM_{2.5}$)	Annual average	10 µg/m3 †	25 µg/m3	2010	
	Daily average	25 µg/m3 †			Exposure concentration obligation
	3-year average		20 µg/m3*	2015	
Nitrogen dioxide (NO_2)	Annual average	40 µg/m3 †	40 µg/m3	2010	
	Hourly average	200 µg/m3 †	200 µg/m3	2010	May be exceeded up to 18 hours a year
Ozone (O_3)	8-hour average	120 µg/m3 ‡			
	Eight-hour average	100 µg/m3 †	120 µg/m3	2010	May be exceeded up to 25 days a year **
Sulphur dioxide (SO_2)	Annual average	50 µg/m3 ‡			
	Daily average	20 µg/m3 †	125 µg/m3	2005	May be exceeded up to three days a year
	Hourly average		350 µg/m3	2005	May be exceeded up to 24 hours a year
Carbon monoxide (CO)	10-minute average	500 µg/m3†			
	15-minute average	100 mg/m3 ‡			
	30-minute average	60 mg/m3 ‡			
	Hourly average	30 mg/m3 ‡			
	Eight-hour average	10 mg/m3 ‡	10 mg/m3	2005	
Lead (Pb)	Annual average	0.5 µg/m3 ‡	0.5 µg/m3	2005	
Benzene (C_6H_6)	Annual average		5 µg/m3	2010	
Arsenic (As)	Annual average		6 ng/m3	2013	
Cadmium (Cd)	Annual average	5 ng/m3 ‡	5 ng/m3	2013	
Nickel (Ni)	Annual average		20 ng/m3	2013	
Benzo[a]pyrene	Annual average		1 ng/m3	2013	

* AEI: Average exposure indicator

** As an average over the three preceding years

Source: http://ec.europa.eu/environment/air/quality/standards.htm

Appendix 2. KLZ file of Study Area

The file Girona.KLZ is located in the website:

http://www.mediafire.com/?yt1b6t3h3lpdsdr

This file can be viewed in Google Earth.

Onyar River, Girona Photo: Albert Marin (www.graph.cat)