Thesis for doctoral degree (Ph.D.) 2015

Long-term exposure to air pollution from road-traffic and cardiovascular disease with a focus on exposure modeling

Michal Korek

Karolinska Institutet

Thesis for doctoral degree (Ph.D.) 2015

Long-term exposure to air pollution from road-traffic and cardiovascular disease with a focus on exposure modeling

Michal Korek

Karolinska Institutet

From THE INSTITUTE OF ENVIRONMENTAL MEDICINE

Karolinska Institutet, Stockholm, Sweden

LONG-TERM EXPOSURE TO AIR POLLUTION FROM ROAD TRAFFIC AND CARDIOVASCULAR DISEASE WITH A FOCUS ON EXPOSURE MODELING

Michal Korek

Karolinska Institutet

Stockholm 2015

Published by Karolinska Institutet.

Printed by Eprint AB 2015

ISBN 978-91-7676-168-7

LONG-TERM EXPOSURE TO AIR POLLUTION FROM ROAD TRAFFIC AND CARDIOVASCULAR DISEASE WITH A FOCUS ON EXPOSURE MODELING

THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Michal Korek

Principal Supervisor:
Professor Göran Pershagen
Karolinska Institutet
Institute of Environmental Medicine

Co-supervisor(s):
Professor Tom Bellander
Karolinska Institutet
Institute of Environmental Medicine

Dr. Johanna Penell
University of Surrey-Guildford
Faculty of Health & Medical Sciences

Opponent:
Professor Ole Hertel
Aarhus University
Department of Environmental Science - Atmospheric chemistry and physics

Examination Board:
Professor Bo Burström
Karolinska Institutet
Department of Public Health Sciences

Professor Johan Frostegård
Karolinska Institutet
Institute of Environmental Medicine

Hans-Christen Hansson
Stockholms Universitet
Department of Environmental Science and Analytical Chemistry

To the vast frontier and the pioneers!

ABSTRACT

Air pollution is an important environmental health factor contributing to the burden of disease. From a public health point of view cardiovascular effects of long-term exposure are predominant, primarily coronary events and stroke. However, sub-types of disease have not been well investigated and few studies have been conducted in areas with lower air pollution levels. The role of timing of exposure is also unclear.

In epidemiological studies different types of models are used to estimate exposure of study participants. It is therefore important to understand if modeled levels are similar for different model types. Furthermore, there is a need to develop better modeling techniques, and it has been proposed to combine models into so called hybrid models.

The aim of this thesis was to investigate the relation between individual long-term air pollution exposure from road traffic and the risk of coronary events and stroke in an area with comparatively low exposure levels, while considering timing of exposure. Furthermore a comparison of dispersion modeling (DM) and land use regression (LUR) was done in several study areas and a hybrid model based on DM and LUR was developed for Stockholm.

From four cohorts in Stockholm County, 20070 individuals were followed for an average of 12 years. Information on covariates was available from questionnaires and interviews from the time of recruitment. Air pollution exposure from traffic was assessed at residential addresses during follow-up using dispersion modeled levels of nitrogen oxides (NO_x), as a marker of exhaust emissions, and particles with an aerodynamic diameter of <10 μm (PM_{10}), as a marker of road dust. A suggestive association between road traffic exposure at the recruitment address and cardiovascular disease incidence was seen. For NO_x the hazard ratio for stroke and coronary events per 20μg/m^3 was 1.16 (0.83 -1.61) and 1.02 (0.82-1.27), respectively. Corresponding hazard ratios for PM_{10} were 1.14 (0.68-1.90) and 1.14 (0.87-1.49), respectively, per 10μg/m^3. Results did not appear to be modified by covariates, disease sub-types or exposure time windows.

LUR models and DMs were compared in 4 to13 European study areas depending on the pollutant. At study addresses, the median Pearson correlation (range) for annual mean concentrations of NO_2, PM_{10} and $PM_{2.5}$ were: 0.75 (0.19–0.89), 0.39 (0.23–0.66) and 0.29 (0.22–0.81). A hybrid model was developed for Stockholm for 93 bi-weekly NO_x observations using DM estimates, LUR variables, stationary monitoring and individual meteorological factors. The hybrid model explained NO_x levels at monitoring stations better (R^2 =89%) than the LUR and DM models (R^2 =58% and R^2 =68%, respectively).

In conclusion, our results suggest an elevated risk of coronary events and stroke related to traffic air pollution exposures in Stockholm County, however, no modification by time window of exposure could be detected. On average, estimates from LUR and DMs correlate well for NO_2 but less so for particulates. To combine DM and LUR seems promising for increasing the quality of the exposure assessment.

LIST OF SCIENTIFIC PAPERS

I. **Korek M**, Bellander T, Lind T, Bottai M, Eneroth K, Caracciolo B, de Faire U, Fratiglioni L, Hilding A, Leander K, Magnusson PK, Pedersen NL, Östenson CG, Pershagen G, Penell J.Traffic-related air pollution exposure and incidence of stroke in four cohorts from Stockholm. Journal of Exposure Science and Environmental Epidemiology 2015;25: 517–523.

II. **Korek M**, Bellander T, Lind T, Bottai M, Eneroth K, Caracciolo B, de Faire U, Fratiglioni L, Hilding A, Leander K Magnusson PK, Pedersen NL, Ostenson CG, Pershagen G, Penell J. Long-term exposure to traffic-related air pollution and coronary events in four cohorts from Stockholm. Manuscript

III. de Hoogh K, **Korek M**, Vienneau D, Keuken M, Kukkonen J, Nieuwenhuijsen MJ, Badaloni C, Beelen R, Bolignano A, Cesaroni G, Pradas MC, Cyrys J, Douros J, Eeftens M, Forastiere F, Forsberg B, Fuks K, Gehring U, Gryparis A, Gulliver J, Hansell AL, Hoffmann B, Johansson C, Jonkers S, Kangas L, Katsouyanni K, Künzli N, Lanki T, Memmesheimer M, Moussiopoulos N, Modig L, Pershagen G, Probst-Hensch N, Schindler C, Schikowski T, Sugiri D, Teixidó O, Tsai MY, Yli-Tuomi T, Brunekreef B, Hoek G, Bellander T. Comparing land use regression and dispersion modelling to assess residential exposure to ambient air pollution for epidemiological studies. Environment International 2014;7: 382-92.

IV. **Korek M,** Johansson C, Svensson N, Lind T, Beelen R, Hoek G, Pershagen G, Bellander T. Can dispersion modeling of air pollution be enhanced by land use regression? Manuscript

RELATED PUBLICATIONS

Eeftens M, Beelen R, de Hoogh K, Bellander T, Cesaroni G, Cirach M, Declercq C, Dėdelė A, Dons E, de Nazelle A, Dimakopoulou K, Eriksen K, Falq G, Fischer P, Galassi C, Gražulevičienė R, Heinrich J, Hoffmann B, Jerrett M, Keidel D, **Korek M**, Lanki T, Lindley S, Madsen C, Mölter A, Nádor G, Nieuwenhuijsen M, Nonnemacher M, Pedeli X, Raaschou-Nielsen O, Patelarou E, Quass U, Ranzi A, Schindler C, Stempfelet M, Stephanou E, Sugiri D, Tsai MY, Yli-Tuomi T, Varró MJ, Vienneau D, Klot Sv, Wolf K, Brunekreef B, Hoek G. *Development of Land Use Regression models for PM(2.5), PM(2.5) absorbance, PM(10) and PM(coarse) in 20 European study areas; results of the ESCAPE project*. Environ Sci Technol. 2012 Oct 16;46(20):11195-205.

Wang M, Beelen R, Basagana X, Becker T, Cesaroni G, de Hoogh K, Dedele A, Declercq C, Dimakopoulou K, Eeftens M, Forastiere F, Galassi C, Gražulevičienė R, Hoffmann B, Heinrich J, Iakovides M, Künzli N, **Korek M**, Lindley S, Mölter A, Mosler G, Madsen C, Nieuwenhuijsen M, Phuleria H, Pedeli X, Raaschou-Nielsen O, Ranzi A, Stephanou E, Sugiri D, Stempfelet M, Tsai MY, Lanki T, Udvardy O, Varró MJ, Wolf K, Weinmayr G, Yli-Tuomi T, Hoek G, Brunekreef B *Evaluation of land use regression models for* NO_2 *and particulate matter in 20 European study areas: the ESCAPE project.* Environ Sci Technol. 2013 May 7;47(9):4357-64

Cesaroni G, Forastiere F, Stafoggia M, Andersen ZJ, Badaloni C, Beelen R, Caracciolo B, de Faire U, Erbel R, Eriksen KT, Fratiglioni L, Galassi C, Hampel R, Heier M, Hennig F, Hilding A, Hoffmann B, Houthuijs D, Jöckel KH, **Korek M**, Lanki T, Leander K, Magnusson PK, Migliore E, Ostenson CG, Overvad K, Pedersen NL, J JP, Penell J, Pershagen G, Pyko A, Raaschou-Nielsen O, Ranzi A, Ricceri F, Sacerdote C, Salomaa V, Swart W, Turunen AW, Vineis P, Weinmayr G, Wolf K, de Hoogh K, Hoek G, Brunekreef B, Peters *A. Long term exposure to ambient air pollution and incidence of acute coronary events: prospective cohort study and meta-analysis in 11 European cohorts from the ESCAPE Project.* BMJ. 2014 Jan 21;348:f7412.

Beelen R, Raaschou-Nielsen O, Stafoggia M, Andersen ZJ, Weinmayr G, Hoffmann B, Wolf K, Samoli E, Fischer P, Nieuwenhuijsen M, Vineis P, Xun WW, Katsouyanni K, Dimakopoulou K, Oudin A, Forsberg B, Modig L, Havulinna AS, Lanki T, Turunen A, Oftedal B, Nystad W, Nafstad P, De Faire U, Pedersen NL, Östenson CG, Fratiglioni L, Penell J, **Korek M**, Pershagen G, Eriksen KT, Overvad K, Ellermann T, Eeftens M, Peeters PH, Meliefste K, Wang M, Bueno-de-Mesquita B, Sugiri D, Krämer U, Heinrich J, de Hoogh K, Key T, Peters A, Hampel R, Concin H, Nagel G, Ineichen A, Schaffner E, Probst-Hensch N, Künzli N, Schindler C, Schikowski T, Adam M, Phuleria H, Vilier A, Clavel-Chapelon F, Declercq C, Grioni S, Krogh V, Tsai MY, Ricceri F, Sacerdote C, Galassi C, Migliore E, Ranzi A, Cesaroni G, Badaloni C, Forastiere F, Tamayo I, Amiano P, Dorronsoro M, Katsoulis M, Trichopoulou A, Brunekreef B, Hoek G. *Effects of long-term exposure to air*

pollution on natural-cause mortality: an analysis of 22 European cohorts within the multicentre ESCAPE project. Lancet. 2014 Mar 1;383(9919):785-95.

Wang M, Beelen R, Stafoggia M, Raaschou-Nielsen O, Andersen ZJ, Hoffmann B, Fischer P, Houthuijs D, Nieuwenhuijsen M, Weinmayr G, Vineis P, Xun WW, Dimakopoulou K, Samoli E, Laatikainen T, Lanki T, Turunen AW, Oftedal B, Schwarze P, Aamodt G, Penell J, De Faire U,**Korek M**, Leander K, Pershagen G, Pedersen NL, Östenson CG, Fratiglioni L, Eriksen KT, Sørensen M, Tjønneland A, Bueno-de-Mesquita B, Eeftens M, Bots ML, Meliefste K, Krämer U, Heinrich J, Sugiri D, Key T, de Hoogh K, Wolf K, Peters A, Cyrys J, Jaensch A, Concin H, Nagel G, Tsai MY, Phuleria H, Ineichen A, Künzli N, Probst-Hensch N, Schaffner E, Vilier A, Clavel-Chapelon F, Declerq C, Ricceri F, Sacerdote C, Marcon A, Galassi C, Migliore E, Ranzi A, Cesaroni G, Badaloni C, Forastiere F, Katsoulis M, Trichopoulou A, Keuken M, Jedynska A, Kooter IM, Kukkonen J, Sokhi RS, Brunekreef B, Katsouyanni K, Hoek G. *Long-term exposure to elemental constituents of particulate matter and cardiovascular mortality in 19 European cohorts: results from the ESCAPE and TRANSPHORM projects.*Environ Int. 2014 May;66:97-106.

Beelen R, Stafoggia M, Raaschou-Nielsen O, Andersen ZJ, Xun WW, Katsouyanni K, Dimakopoulou K, Brunekreef B, Weinmayr G, Hoffmann B, Wolf K, Samoli E, Houthuijs D, Nieuwenhuijsen M, Oudin A, Forsberg B, Olsson D, Salomaa V, Lanki T, Yli-Tuomi T, Oftedal B, Aamodt G, Nafstad P, De Faire U, Pedersen NL, Östenson CG, Fratiglioni L, Penell J, **Korek M**, Pyko A, Eriksen KT, Tjønneland A, Becker T, Eeftens M, Bots M, Meliefste K, Wang M, Bueno-de-Mesquita B, Sugiri D, Krämer U, Heinrich J, de Hoogh K, Key T, Peters A, Cyrys J, Concin H, Nagel G, Ineichen A, Schaffner E, Probst-Hensch N, Dratva J, Ducret-Stich R, Vilier A, Clavel-Chapelon F, Stempfelet M, Grioni S, Krogh V, Tsai MY, Marcon A, Ricceri F, Sacerdote C, Galassi C, Migliore E, Ranzi A, Cesaroni G, Badaloni C, Forastiere F, Tamayo I, Amiano P, Dorronsoro M, Katsoulis M, Trichopoulou A, Vineis P, Hoek G. *Long-term exposure to air pollution and cardiovascular mortality: an analysis of 22 European cohorts.* Epidemiology. 2014 May;25(3):368-78.

Stafoggia M, Cesaroni G, Peters A, Andersen ZJ, Badaloni C, Beelen R, Caracciolo B, Cyrys J, de Faire U, de Hoogh K, Eriksen KT, Fratiglioni L, Galassi C, Gigante B, Havulinna AS, Hennig F, Hilding A, Hoek G, Hoffmann B, Houthuijs D, **Korek M**, Lanki T, Leander K, Magnusson PK, Meisinger C, Migliore E, Overvad K, Ostenson CG, Pedersen NL, Pekkanen J, Penell J, Pershagen G, Pundt N, Pyko A, Raaschou-Nielsen O, Ranzi A, Ricceri F, Sacerdote C, Swart WJ, Turunen AW, Vineis P, Weimar C, Weinmayr G, Wolf K, Brunekreef B, Forastiere F. *Long-term exposure to ambient air pollution and incidence of cerebrovascular events: results from 11 European cohorts within the ESCAPE project.*Environ Health Perspect. 2014 Sep;122(9):919-25

In addition, co-authorship of 13 manuscripts on health effects of air pollution and air pollution modeling, primarily within the ESCAPE project, as responsible for the exposure assessment in cohorts from Stockholm.

CONTENTS

LIST OF ABBREVIATIONS

CEANS	Cardiovascular Effects of Air pollution and Noise in Stockholm
CE	Coronary events
CVD	Cardiovascular disease
DM	Dispersion modeling
ESCAPE	European Study of Cohorts for Air Pollution Effects
EVA	Effects of Road Analysis
HR	Hazard ratio
ICD	International classification of diseases
LUR	Land use regression
MI	Myocardial infarction
MT	Meterological predictor
NVDB	National road database
NO_x	Nitrogen oxides
NO_2	Nitrogen dioxide
HNO_3	Nitric acid
O_3	Ozone
PM_{10}	Particles with an aerodynamic diameter of less than 10μg/m^3
$PM_{2.5}$	Particles with an aerodynamic diameter of less than 2.5μg/m^3
RMSE	Root mean square error
SALT	Screening across the Lifespan Twin study

SAMS	Small areas for market statistics
SDPP	Stockholm Diabetes Preventive Program
SES	Socioeconomic status
SIXTY	60 year old men and women from Stockholm
SMHI	Swedish Meteorological and Hydrological Institute
SNACK	Swedish National Study of Aging and Care in Kungsholmen
STAT	Predictor based on data from stationary monitoring
WHO	World Health Organization

1 INTRODUCTION

Exposure to ambient air pollution has since long been associated with adverse health effects. Early studies concerned extreme outbreaks such as the Meuse Valley fog in 1930 (Roholm 1937, Nemery, Hoet et al. 2001) and the London smog episode in 1952 (Wilkins 1954, Davis 2002), harvesting 12000 excess deaths during peak exposure and the following two months. The "new" era of air pollution epidemiology started with the land mark study called the Harvard six cites study, which indicated adverse health effects following long-term exposure at much lower air pollution levels than in the earlier smog episodes (Dockery, Pope et al. 1993). Today, the estimated health effects of exposure to ambient particulate air pollution globally are massive. In a recent assessment of several major risk factors influencing the global burden of disease, the joint effect of outdoor and indoor particulate air pollution in 2010 was estimated to cause between 5.3 and 7.9 million premature deaths. When considering outdoor ambient air pollution specifically the estimates ranged between 2.6 million to 4.4 million deaths (Lim, Vos et al. 2012). According to data these deaths were mainly related to ischemic heart disease and stroke, together responsible for 80% of all events (WHO 2014a). The corresponding yearly excess mortality due to outdoor air pollution in Europe was 482000 deaths. In addition, air pollution causes a number of non-lethal effects such as cardiovascular and respiratory diseases in adults as well as asthma and lung function disturbances in children.

1.1 AMBIENT AIR POLLUTION

Air pollution can be defined as substances in ambient air with negative effects on health and the environment. It is a dynamic mixture of particles and gases from multiple sources, which can be natural, such as dust storms, wild fires, pollen or sea spray, or anthropogenic (manmade) such as industrial activity, biomass burning and road traffic. Directly emitted pollutants are called "primary", while pollutants formed in the air from primary pollutants are called "secondary".

Nitrogen oxides (NO_x) are used in this study as a marker of vehicle exhaust emissions and are formed from nitrogen and oxygen in air by combustion at high temperatures. NO_x emitted into the ambient air consists of the primary pollutant nitric oxide (NO) which by the interaction with ozone and oxygen in air forms the secondary pollutant nitrogen dioxide (NO_2). Depending on the source significant amounts of NO_2 may also be emitted directly. The main source of NO_x emissions in urban areas are vehicles. Other gases directly emitted via combustion processes include sulphur oxides (SO_x) and carbon monoxide (CO). Another secondary pollutant is ozone (O_3) formed by reactions of NO_2 and volatile organic compounds, or naturally, in the presence of sunlight. Furthermore, in the presence of solar radiation NO_2 is also slowly converted into constituents of particulate matter via nitric acid (HNO_3) (WHO 2013).

Particulate matter (PM) is a complex mixture of substances in both solid and liquid form varying in size and composition. Main components include sulfate, nitrates, ammonia, sodium chloride, black carbon, mineral dust and water (Schwarze, Ovrevik et al. 2006, Kelly and Fussell 2012). PM is defined according to particle size, where different sizes have different aerodynamic properties as well as composition. Both size and composition are related to the characteristics of the source. Furthermore, the size of the particles influences transportation and deposition, both in the environment and within the human respiratory system. The largest inhalable particle fraction PM_{10} (PM with an aerodynamic diameter < 10 µm) includes all smaller PM sizes and may have both anthropogenic and biogenic sources. The coarse fraction $PM_{10-2.5}$ is mainly composed out of crustal materials and such particles are able to reach the upper bronchial tract. Another size fraction, $PM_{2.5}$ (PM with an aerodynamic diameter < 2.5 µm) or so called fine particles are primarily derived from combustion processes. These particles are able to deposit deep in the respiratory tract. The $PM_{2.5}$ size fraction also includes ultrafine particles or $PM_{0.1}$ which may penetrate into the alveoli and possibly reach into the circulatory system (Brunekreef and Holgate 2002, Brook, Rajagopalan et al. 2010, Chin 2015). In this thesis traffic-related PM_{10} is used as a proxy for road dust from road wear, and other grinding mechanisms involving breaks, clutches and tires.

1.2 EXPOSURE TO AIR POLLUTION

Air pollution exposure on a population or individual level depends on the combination of both pollutant characteristics and temporal patterns of emission sources and weather. For example, air pollution emissions from traffic vary with daily cycles in traffic intensity together with weather patterns which in turn are dependent on season. In addition, variations in solar radiation and temperature have an important role for the formation of air pollutants (Brook, Rajagopalan et al. 2010). The subsequent spread of pollution from the source is further dependent on the formation rate of the pollutant and its atmospheric lifetime (dependent on the size fraction and reactivity) together with meteorological effects such as wind speed, stability and direction but also infrastructural features able to both shield from and trap (accumulate) air pollutants. Thus, air pollution levels in cities may vary within meters (Briggs, de Hoogh et al. 2000, Durant, Ash et al. 2010). Personal air pollution exposure is dependent on the individual's life patterns such as choice of living area, which is often related to socioeconomic features (Filleul and Harrabi 2004). The type and geographic location of the work place, and the time and way of commuting add further variation in the exposure (Ozkaynak, Baxter et al. 2013). For example, variation in individual exposure has been observed when comparing walking, biking and commuting by car or buss using personal monitoring (Briggs, de Hoogh et al. 2008, de Nazelle, Fruin et al. 2012). Techniques based on individual exposure measurements are not easily extrapolated to epidemiological studies including large amounts of participants (Ragettli, Phuleria et al. 2015).

However, dispersion modeled exposure at home address and workplace level has been shown to capture daily individual exposure variation measured by personal monitoring (Bellander, Wichmann et al. 2012). Another important source is air pollution generated indoors. People spend most of the time indoors, especially chronically ill, very young or old individuals (WHO 2013). Studies indicate that ambient air pollution exposure is modified both by time spent indoors and seasonality in particle infiltration (Hänninen, Hoek et al. 2011). Globally, health risks related to indoor air pollution are overrepresented in the low and middle income countries and often related to burning of solid fuel and indoor cooking (WHO 2014b).

1.3 AIR POLLUTION MODELING

During the evolution of exposure modeling and air pollution risk studies, the increase in computational capacity together with developed software such as geographical information systems and constantly refined study designs have allowed us to trace health effects to types of pollution sources and to adjust for time-related variations in source strength and meteorology on different geographical levels. Early exposure assessment was conducted on large spatial scales (city) and exposure estimates attributed to a specific population were even based on a single monitoring site (Dockery, Pope et al. 1993).

Subsequent advancements included estimating exposure on much finer scales, capturing intra-urban (HEI 2010) as well as temporally resolved exposure variation (Bellander, Berglind et al. 2001, Johansson, Burman et al. 2009). Large intra-urban variations have been found in monitoring studies (Cyrys, Eeftens et al. 2012), which confirms the importance of techniques able to adjust for the small scale spatial variability in air pollution. Today, two alternative methods describing small scale variations in air pollution are dispersion modeling (DM) and land use regression modeling (LUR).

Dispersion models combine input data on emissions from point sources (for example chimneys), line sources (such as road traffic), area sources (e.g. port area), meteorological conditions including wind speed, direction and stability, solar radiation, temperature and topography (Bellander, Berglind et al. 2001, Jerrett, Arain et al. 2005). Emissions from traffic are calculated based on the combination of source intensities and source specific emission factors adjusting emission calculations for characteristics such as car fleet composition, types of fuels etc. Data are usually combined in a Gaussian plume equation where the geographical distribution and pollutant levels are calculated based on deterministic assumptions of the atmospheric dispersion of the pollutants (Bellander, Berglind et al. 2001). DM can calculate levels of air pollution at any time scale and for different geographical resolutions such as local and regional scales or at receptor points. Furthermore, dispersion modeling can include adjustments for street canyon effects leading to elevated pollution in streetscapes surrounded by buildings (Hertel, Berkowicz et al. 1991, Raaschou-Nielsen, Hertel et al. 2000). DMs are generally calibrated against monitored levels of air pollution (Bellander, Berglind et al. 2001).

In land use regression modeling, monitored pollution variability in an area is related to area characteristics potentially affecting pollution concentrations, commonly through a multiple linear regression technique. Measured air pollution levels are retrieved from stationary national or regional monitoring or by monitoring campaigns. In campaigns, monitors can be distributed to obtain data for specific scientific aims. Although monitoring is often conducted in shorter time periods, the observations may be adjusted for seasonal variation and other time trends using stationary data. Typical predictors of air pollution include different types of data on traffic, population and land use (vegetation and urbanization) retrieved from geographical information systems (GIS). Using GIS, usually the distance to a predictor (e.g. a road of a certain class) or the amount of predictor (population, land use etc.) within a buffer circle around each monitor is calculated. LUR models are calibrated by regressing predictor variables against monitored concentrations, and the final LUR model is established by inclusion of the predictors that together explain most of the exposure variability. Data on these predictors are then extracted for sites aimed for exposure assessment and used together with the established regression formula (Briggs, Collins et al. 1997, Jerrett, Arain et al. 2005). Future exposure modeling in the form of hybrid models that combine existing data and techniques have been proposed. Such models are anticipated to further minimize exposure error in epidemiological studies, and thus help to increase the accuracy and precision of exposure-response functions (Jerrett, Arain et al. 2005).

DM and LUR are commonly used for air pollution risk assessment but not often in combination. The models have different strengths and weaknesses. DM models are often calibrated against one or a few stationary monitoring sites. However, central site monitoring sites do not completely represent the population exposure, particularly for air pollutants with high spatial variability (Ozkaynak, Baxter et al. 2013). Therefore, a possible optimization of DM could be the inclusion of data from techniques based on larger monitoring networks in the same region. LUR is often based on dense monitoring campaigns. On the other hand, traditional LUR modeling does not include the interaction between meteorology and traffic (Wilton, Szpiro et al. 2010), which can be done with DM on any time scale. LUR models also have difficulties to capture elevated pollution levels due to street-canyon effects (Beelen, Hoek et al. 2008). Some street-canyon adjustments for LUR exist but are oversimplifications (Brauer, Hoek et al. 2003) and difficult to apply to larger amount of addresses or in cities with complex street compositions or does only enhance the LUR model marginally (Eeftens, Beekhuizen et al. 2013). For DM, street-canyon adjustments have been developed and used in epidemiological studies (Gruzieva, Bellander et al. 2012). Earlier combinations of LUR and meteorological factors as well as exposure estimates from simpler DM have been successful (Wilton, Szpiro et al. 2010). In general, model mixing leads to better resolution or a better coverage of factors relevant for the pollution concentrations (Ozkaynak, Baxter et al. 2013), which brings promise to future DM-LUR hybrid models.

1.4 CARDIOVASCULAR EFFECTS ASSOCIATED WITH AIR POLLUTION

Cardiovascular disease (CVD) is a general term for disorders of the heart and blood vessels. Types of CVD include coronary heart disease from complications due to depleted or limited supply of blood to the heart, rheumatic heart disease (heart valve damage due to rheumatic fever) and congenital heart disease (malformed heart structure), as well as cerebrovascular disease due to disturbed flow in blood vessels supplying the brain (Martinelli, Olivieri et al. 2013). A large amount of studies relating long-term air pollution exposure to various cardiovascular conditions and diseases have been published (Brook, Rajagopalan et al. 2010, Hoek, Krishnan et al. 2013, WHO 2013) Associations have been found (mainly for PM) with the development of ischemic heart disease, heart failure, cerebrovascular disease, atherosclerosis, hypertension, and cardiovascular mortality. Proposed pathophysiological mechanisms include systemic inflammation, oxidative stress, imbalance in the autonomic nervous system, endothelial dysfunction, vasoconstriction, thrombosis and epigenetic modifications (Newby, Mannucci et al. 2015). Epidemiological findings have been supported by experimental studies focusing on mechanisms. The pathway with best experimental support is a PM induced provocation of inflammatory response and oxidative stress in the lungs spilling over into systemic inflammation (Brook, Rajagopalan et al. 2010, Chin 2015).

When investigating coronary events (CE) in relation to long-term air pollution exposure most studies focused on mortality, mainly finding increased risks (Hoek, Krishnan et al. 2013). In comparison, non-fatal CE events have been studied less well even though these constitute the majority of the cases. The few cohort studies on the association between non-fatal CE and long-term PM_{10} exposure show varying results, and no association has generally been found for NO_2/NO_x exposure (Miller, Siscovick et al. 2007, Lipsett, Ostro et al. 2011, Puett, Hart et al. 2011, Atkinson, Carey et al. 2013, Cesaroni, Forastiere et al. 2014, Katsoulis, Dimakopoulou et al. 2014). Similar results were found in studies on myocardial infarction and long-term exposure to air pollution from traffic, suggesting stronger associations for fatal compared to non-fatal events (Miller, Siscovick et al. 2007, Puett, Schwartz et al. 2008, Rosenlund, Bellander et al. 2009). The evidence on the associations between cerebrovascular incidence or mortality and long-term air pollution exposure, is limited and conflicting (Nafstad, Haheim et al. 2004, Pope, Burnett et al. 2004, Miller, Siscovick et al. 2007, Andersen, Kristiansen et al. 2012, Stafoggia, Cesaroni et al. 2014). Still, studies on short-term exposure have reported associations both for mortality, and hospital admissions due to stroke (Brook, Rajagopalan et al. 2010), primarily in mid- and low income countries (Shah, Lee et al. 2015). Most studies on long-term air pollution exposure and cardiovascular events have been conducted in or included areas with pollution exposure exceeding the WHO guidelines, e.g. $20\mu g/m^3$ for PM_{10}. However, recent findings suggest a lack of a threshold level for the air pollution effect on cardiovascular events (Brook, Rajagopalan et al. 2010, Cesaroni, Forastiere et al. 2014). There is a need for more studies on health effects in regions with air pollution levels below current guidelines.

Short-term air pollution exposure is an established trigger of cardiovascular disease (Brook, Rajagopalan et al. 2010), however, the role of timing related to long term exposure is not well understood. An Irish intervention study found a decrease in cardiovascular mortality within a year after drastically decreased levels of black smoke (Clancy et al., 2002). When investigating cardiovascular mortality (Laden, Schwartz et al. 2006) and myocardial infarction (Zanobetti and Schwartz, 2007) in relation to different exposure time-periods, a stronger effect was found for exposure within a few years of the event compared to other time-periods. On the contrary, other studies failed to find time-windows of particular importance in relation to CE (Nafstad, Haheim et al. 2004, Puett, Hart et al. 2009, Puett, Hart et al. 2011, Chen, Goldberg et al. 2013). Similar results were found for the risk of stroke (Nafstad, Haheim et al. 2004, Puett, Hart et al. 2011, Chen, Zhang et al. 2013), however, most of the studies on cerebrovascular effects of air pollution did not assess exposure time windows.

1.5 AIMS

The main aims of this thesis were to investigate the association between air pollution and cardiovascular disease at the lower range of previously studied exposures, and to develop the exposure assessment methodology.

The specific aims were

- To investigate the effect of long term exposure to traffic-related air pollution on the risk of coronary events and stroke
- To assess the role of different exposure-time windows for the risk of cardiovascular disease related to air pollution exposure
- To investigate if the choice of exposure model (land use regression and dispersion models) has an effect on air pollution levels attributed to study populations
- To create a hybrid model by combining land use regression and dispersion modeling aimed at improving exposure assessment

2 MATERIAL AND METHODS

2.1 STUDY POPULATION

In paper I and II relating air pollution exposure to risk of cardiovascular disease, four cohorts located in Stockholm County, Sweden, were used. The cohorts were "the Stockholm Diabetes Preventive Program" (SDPP)," the cohort study of 60 year olds" (SIXTY), "the Screening Across the Lifespan Twin Study" (SALT) and " the Swedish National study of Aging and Care in Kungsholmen" (SNACK). Furthermore, the addresses of study subjects in these cohorts were also used in paper III, comparing modeled levels of air pollution from land use regression (LUR) and dispersion modeling (DM) in several European study areas. This paper was part of the multicenter collaboration project ESCAPE "The European Study of Cohorts for Air Pollution Effects ". In Paper IV, data linked to study subjects were not used.

2.1.1 The Stockholm Diabetes Preventive Program (SDPP)

From this population-based prospective study all men recruited into the cohort 1992-1994 (n=3128) and women recruited in 1996-1998 (4821) were used in paper I and II. The catchment area included five municipalities in Stockholm County (Sigtuna, Upplands Väsby, Värmdö, Upplands Bro and Tyresö) and the initial recruitment involved all citizens aged 35-56 years at the time of the initial questionnaire. Seventy-nine percent of the men and 85% of the women responded to a short postal questionnaire regarding their family history of diabetes (FHD), defined as diabetes in one first degree relative or two second degree relatives. From these, all respondents that reported FHD (n=5689) and 424 women with gestational diabetes as well as an age and sex adjusted sample of non FHD respondents (n=5921) were invited for a base-line survey. Respondents were excluded from follow-up if they had a diagnosis of diabetes, were of foreign origin or the information on FHD was unclear, leaving 7949 study subjects (Eriksson, Ekbom et al. 2008). Furthermore, for the purpose of study I and II, individuals should not have insufficient address information (therefore missing exposure) at any addresses or missing data on any of the confounders. The final number of participants from the SDPP study was 7451 in paper I and 7450 in paper II (Table 1).

2.1.2 The cohort study of 60 year olds (SIXTY)

The aim of the cohort study of 60 year olds (SIXTY) was to identify biological and socio-economical risk factors and predictors for cardiovascular disease (Wandell, Wajngot et al. 2007). An invitation was sent out to every third individual living in Stockholm County who had turned 60 years of age between August 1997 and March 1999. A total of 5460 participants, 2779 men and 2681 women, were invited with an overall participation rate of 77% (n=4228). At study entry a questionnaire was filled in.

For the purpose of study I and II, eligible individuals should not participate in any of the other cohorts, have insufficient address information/missing exposure data or missing data on any of the confounders, leaving 3697 study participants (Table 1).

2.1.3 The Screening Across the Lifespan Twin Study (SALT)

In the Screening Across the Lifespan Twin Study (SALT) all twins born in Sweden before 1958 were screened for the most common complex diseases, including cardiovascular diseases (Lichtenstein, Sullivan et al. 2006). Participants were recruited at two stages between 1998-2002 starting with twins over 65 years of age at the time of the interview and then younger twins. For all recruited participants a computer assisted telephone interview was conducted, including the collection of information on risk factors of cardiovascular disease. For individuals born 1886-1925, the response rate of eligible subjects (alive and living in Sweden) was 65 %. The response rate for twins born 1926-1958 was 74%, leaving 20839 male and 22186 female twins. For the purpose of study I and II, the SALT participants residing in Stockholm County at recruitment were included, resulting in 7043 subjects with an age range of 42-100 years at recruitment. After exclusions due to participation in earlier cohorts, having missing information on exposure or in any covariate the total number for analysis was 6006 in paper I and 6004 paper II (Table 1).

2.1.4 The Swedish National study of Aging and Care in Kungsholmen (SNACK)

The Swedish National study of Aging and Care in Kungsholmen (SNACK) is an ongoing longitudinal study including randomly sampled individuals that were >=60 years old between March 2001 and August 2004 living in Kungsholmen in Stockholm City (Lagergren, Fratiglioni et al. 2004, Santoni, Angleman et al. 2015). The cohort was set up to investigate various health processes associated with aging as well as to identify intervention strategies to improve health care in the elderly. Study participants were stratified for age and year of assessment and investigated in sub-cohorts (60, 66, 72, 78, 81, 84, 87, 90, 93, 96, and 99+ year of age). Information on confounders was collected through interviews, clinical examinations, cognitive assessment and examinations of physical function. Out of 5111invited individuals, 521 were excluded due to, death before study entry, deafness, being non-traceable, had moved from Kungsholmen or were non-Swedish speakers. 1227 declined participation leaving 3363 study subjects (a participation rate of 73%) of 60-104 years of age. The final number of eligible study subjects in papers I and II were 2916 and 2917, respectively (Table 1).

Table 1. Number of individuals included in the analysis of four cohorts from Stockholm County

Cohort	SDPP	SIXTY	SALT	SNACK
Number of recruited	7949	4232	7043	3363
Reasons for exclusion				
Participation in more than one cohort[a]	0	8	159	78
Missing exposure data for time-window[b]	111	47	168	17
Missing data on covariates	387	480	710	352
Total number in analysis (% of number recruited)	7451 (94%)	3697 (87%)	6006 (85%)	2916 (87%)

[a] Subjects are included in the first cohort into which they were recruited.

[b] Subjects may have moved in and out from the study area in time periods earlier than study entry, leading to missing exposure in time-windows and therefore exclusion.

Note: Similar exclusions were made in paper II with the final number of participants of 7450 in SDPP, 3697 in SIXTY, 6004 in SALT and 2917 in SNACK.

2.1.5 The European Study of Cohorts for Air Pollution Effects

This multi-center study abbreviated ESCAPE was designed to investigate the effects of long-term air pollution exposure on diverse health outcomes by the use of existing cohorts across Europe (www.escapeproject.eu). For the purpose of study III, 13 European cohorts were selected based on accessibility of data from both dispersion and land use regression modeling. The cohorts were located in Umeå, Stockholm, Helsinki-Vantaa, Bradford, London, The Netherlands, the Ruhr area, Basel, Geneva, Lugano, Rome, Barcelona and Athens. Across participating cohorts, the number of residential addresses ranged between 39409 in Stockholm to 737 in Geneva. The original purpose of the cohorts varied but all were used for studies relating air pollution to different health outcomes.

2.2 EXPOSURE ASSESSMENT

In papers I, II and III long-term air pollution exposure levels from traffic were calculated as annual means at subject-specific residential addresses (at study entry and during follow-up) using a dispersion modeling (DM) system described previously (Bellander, Berglind et al. 2001). Briefly, for each participant in the four cohorts, a residential history covering 1991 to 2010 was retrieved from the Swedish tax authorities. Address history earlier than 1991 was also known for participants if they lived on the same address before this year. Ninety percent of the addresses were directly transferable into geographical coordinates by matching with databases at the Swedish Mapping Cadastral and Land Registration Authority. A remaining 9% were geocoded manually while 1% of the addresses were non-traceable.

For every geocoded address, annual mean levels of traffic-derived PM_{10} (marker of road dust) and NO_x (marker of tail pipe emissions) were modeled from 1987 to the end of follow-up (2011) using the Airviro Air Quality Management system (SMHI, Norrköping, Sweden http://airviro.smhi.se). Information on emission sources (local road traffic) were retrieved from an inventory provided by the Stockholm and Uppsala County Air Quality Management Association. This inventory includes a map covering 90% of the trafficked roads in the form of road links (Johansson, Hadenius et al. 1999). Every road link contains information on traffic related data such as traffic intensities the share of heavy traffic and speed limits. The inventory is updated yearly since 1993, although traffic counts were not updated yearly for all streets. In study I, II and III, the emission inventory for the year 2004 was used. However, for calculations of NO_x the traffic intensities were re-scaled during follow-up using annual data on average traffic intensities in central parts of Stockholm. To calculate pollution levels at the road links of the road map, the EVA (Effects at Road Analysis) model of the Swedish Transport Administration was used. In the EVA model, the emission inventory data were combined with emission factors for tail pipe (NO_x) and road wear (PM_{10}) and levels were estimated during follow-up. For the NO_x calculations emission factors updated every five years from 1990 to 2010 were used. For PM_{10} the levels of road wear were assumed stable over the time period and emission scenarios were used from the year 2004. The described exposure scenarios were adjusted for car fleet composition, share of diesel cars and the composition of the vehicle fleet in terms of European emission standards (Euro classification) for different years. Non-exhaust PM, including road, break and tire wear particles, was also included in the model. The emission concentrations at road links, together with wind fields (attained from a wind model based on local climatology) were used in a Gaussian air quality dispersion model. The model calculated the meteorological spread of annual mean NO_x and PM_{10} with a spatial resolution of 25x25 to 500x500 meters depending on area type (city, urban, rural) in Stockholm County. For addresses where pollution levels were influenced by a street canyon effect, a contribution was calculated using the SMHI-Airviro street canyon model (http://airviro.smhi.se).

In paper III the difference in estimated levels of various air pollutants from DM and LUR models were explored. For this purpose, study area specific DMs and LUR models were used. Three different types of DMs were used: the Gaussian plume model (used in 10 study areas), Eulerian or chemical transport models (2 study areas) and the computational fluid dynamic model (1 area).

The development of land use regression models was harmonized over study areas within the ESCAPE project and detailed information on LUR model development has been documented earlier (Eeftens, Beelen et al. 2012, Beelen, Hoek et al. 2013, de Hoogh, Wang et al. 2013). Briefly, a monitoring campaign was initiated in the different study areas between 2008 and 2011. Ogawa badges were used to measure NO_2 and NO_x levels and Harvard Impactors were used for monitoring of particulate matter of different sizes (Cyrys, Eeftens et al. 2012). The number of monitors used to measure the pollutant levels in study III were 20-40 for NO_2 in each study area and 13-34 monitors for PM_{10}, $PM_{2.5}$ and $PM_{2.5}$ absorbance. Monitors were positioned to include regional, urban and traffic sites with more densely distributed monitoring in areas with more small scale pollution variability. For the purpose of the ESCAPE project annual air pollution levels were constructed and the yearly means were estimated by sampling each monitoring site for two weeks in the cold, warm and an intermediate season, respectively. Furthermore, a reference site was used to adjust the measurements for temporal variation over the year. Predefined area specific predictor data was collected from both central (ESCAPE) and local sources, and included data on traffic, land use, water, population density and terrain. These data were then used to describe local and urban sources of air pollution around every monitoring site as, buffer areas with radii of (25-5000 m), proximity to trafficked roads adjusted for road type, traffic intensities on roads adjusted for road type and type of vehicle or by combining the adjusted traffic intensities with the length of road segments within buffer areas (traffic load) using geographical information systems. The monitored annual means were then combined with the predictor data in a least square regression model using a forward stepwise selection procedure. Calibrated models were then used to describe air pollution levels at geocoded addresses of study participants within each study area.

In paper IV two dispersion models were used together with land use regression methodology to create a DM-LUR hybrid model. The two DMs used the "the national road database" (NVDB) emission inventory developed by the Swedish transport administration. Emission factors for NO_x levels from local traffic were calculated in this study using the Handbook Emission Factors for Road Transport database (http://www.hbefa.net/e/index.html), providing data on emissions for different categories of vehicles and for different traffic situations. Furthermore, average vehicle intensities were adjusted for month of year, type of day, hour of day and speed limits. The DMs were used to calculate daily emissions for 31 of the monitoring sites originally selected in the ESCAPE study. These sites were selected to be within the domain of the models and to reflect NO_x levels in urban and traffic settings. At these sites hourly NO_x levels were calculated and summed into bi-weekly averages corresponding to the 2-weeks of monitoring in the ESCAPE study. The DM estimates were

complemented with rural background NO_x concentrations to better correlate with observed NO_x levels. In the analyses the bi-weekly means of NO_x were used, giving 93 NO_x observations in total. At urban sites or street sites with an open street configuration, the same Gaussian model and wind model described for paper I and II was used (Korek, Bellander et al. 2015). However, the resolution of the Gauss model in this study was 500m. To describe NO_x levels at monitoring sites in street canyon scenarios, the SIMAIR-road model was used (Omstedt, Andersson et al. 2011). Meteorological data were supplied by the MESAN system based on all available measurement stations, radars and satellites combined with a background field forecast (Häggmark, Ivarsson et al. 2000). LUR modeling was performed using least square multiple linear regression as described earlier for paper III. In an intermediate step between modeling air pollution using DM and LUR separately and the Hybrid model, the 93 NO_x observations were also used to calibrate a LUR model for which temporally resolved meteorological data (used in dispersion modeling) were offered as well as stationary NO_x measurements representing urban and traffic variations over the measurement year. This "meteorological" LUR model was then compared to a multiple linear regression model including DM estimates together with meteorology and stationary data. The hybrid model was developed based on all predictor data including DM estimates (spatio-temporal data), stationary monitoring of NO_x levels and meteorology (temporal data) and LUR data (spatial).

2.3 DEFINITION OF HEALTH OUTCOMES

In study I and II information on coronary (CE) events for the period February 1964 to December 2011 was gathered from the National Hospital Discharge Registry and the National Cause of Death Registry. In the National Hospital Discharge Registry CE was defined as "Acute Myocardial Infarction or "Other acute and sub-acute forms of ischemic heart disease" using the international classification of disease (ICD) codes (ICD9: 410; 411; ICD10: I21, I23, 120.0, I24 while "Ischemic heart disease" was defined using (ICD9: 410-414; ICD10: I20-25) in the Cause of Death Registry. The same registers were used to retrieve data on stroke events as: hospitalizations with principal diagnosis of ischemic stroke (ICD9: 433; 434; ICD10: I63), hemorrhagic stroke (ICD9: 431; ICD10: I61), unspecified stroke (ICD9: 436; ICD10: I64) and out-of-hospital deaths from cerebrovascular diseases (ICD9: 431–436; ICD10: I61-I64).

Stroke and CE events occurring after recruitment to the respective cohorts were included in the analyses, whereas earlier events were used to classify later events as non-incident. Both stroke and CE events were classified as fatal if the person passed away within 28 days after disease onset.

2.4 STATISTICAL ANALYSIS

2.4.1 Association between air pollution and cardiovascular disease

In all papers, air pollutant concentrations were calculated in micrograms per cubic meter ($\mu g/m^3$). In paper I and II the effect estimates were assessed using increments of $20\mu g/m^3$ for NO_x and $10\mu g/m^3$ for PM_{10}. In paper I - III environmental exposure to air pollution from traffic was calculated as annual means of NO_x and PM_{10} for subject specific entry year and in paper I and II also within time-windows during follow-up. Time-windows were constructed as 0-2 2-4 4-6 and 6-10 years of cumulative NO_x and PM_{10} exposure, prior to time periods of 6 months during follow-up and adjusted for change of address. CE and stroke outcomes (paper I and II) were coded as dichotomous variables (0/1). Only study subjects without missing data in covariates both at study entry (confounders/exposure), or during pro- and retrospective follow -up (exposure) were included.

To assess the relation between long-term exposure to ambient air pollution and the risk of CE or stroke, Cox proportional hazard regression models were used (paper I and II). Hazard ratios (HRs) and 95% confidence intervals (CIs) were first calculated for each cohort separately and then effect estimates were meta-analyzed. Study subjects were assumed under risk from the time of enrolment into the study until occurrence of an event under study (CE/stroke), death due to other causes, emigration to an address for which air pollution exposure was not defined or end of follow-up (31 December 2011). Age divided into 6 month periods was used as the underlying time variable. Calendar time was adjusted for in the analysis, using 5 year periods. The annual mean levels at recruitment and during the time windows of NO_x and PM_{10} exposure were then related to these 6 month risk periods.

The final models were adjusted for a set of confounders selected a priori. The hazard risk ratio for CE and stroke was adjusted for gender, education level, smoking status, smoking intensity among current smokers and socioeconomic index. The index was based on current or last (if retired) profession and categorized into low (blue collar worker), medium (low and intermediate level white collar worker, and self-employed) and high (high-level white-collar worker). Data on potential covariates were included in cohort specific analyses if available for at least two cohorts with at least 80% non-missing observations. For the SALT cohort, alcohol consumption and occupational status were not available and for SNACK physical activity was not included. The proportional-hazard assumption was tested for all categorical covariates. If any variable in the individual cohort models violated this assumption, the effect estimates of that model were compared with a stratified Cox analysis for that cohort and covariate (Nafstad, Haheim et al. 2004).

The effects estimates in each cohort were combined in a random effect meta-analysis (DerSimonian and Laird 1986). Statistical heterogeneity was evaluated by use of the I^2 statistic (Higgins, Thompson et al. 2003). Effect modification was investigated based on cohort specific analyses which were also combined by meta-analysis. Potential effect modification was investigated for gender, smoking (never smoker ever smoker), diabetes

(yes/no) and hypertension (yes/no) defined as ≥ 140mmHg systolic or ≥90mmHg diastolic blood pressure, or intake of blood pressure-lowering medication, or in the SALT cohort, on self-reported data on hypertension. Information on family history of diabetes was not available for SALT and SNACK cohorts, limiting the effect modification analyses to the SDPP and SIXTY cohorts. In paper II potential effect modification by re-location (defined as never moving compared to ever moving) was also investigated.

Sensitivity analyses in paper I were performed to investigate the association between traffic derived air pollution and incident cases of stroke after study enrollment, non-fatal stroke and ischemic stroke by meta-analyzing cohort specific effect estimates. For paper II coronary events were restricted to non-fatal cases, incident cases after study enrolment and myocardial infarction in sensitivity analyses. For both studies the effect of contextual confounding was assessed by including a contextual socioeconomic status (SES) variable in the form of mean income at "neighborhood level" to the fully adjusted model." Neighborhood" was defined as SAMS (Small Areas for market statistics) areas, containing approximately 1000 inhabitants with similar SES characteristics.

Both for paper I and II a linear trend between exposure time windows and the events was assessed by using cohort specific effect estimates derived from time windows as a dependent variable and the complementary time window intervals as categorical explanatory variables and then combining the data in a meta-regression model.

2.4.2 Comparison of land use regression and dispersion model

The comparison of estimated levels of air pollution from DM and LUR models was done on address levels. For this purpose Pearson and Spearman correlation coefficients were calculated and the relation was visualized in scatterplots. In epidemiological studies, modeled air pollution exposure is often categorized to relax the assumption of linearity between exposure and outcome. The DM and LUR estimates were therefore categorized into quintiles for which level of agreement was calculated using the Kappa coefficient. Also Bland–Altman plots were constructed, in particular to test if the difference between LUR and DM estimates depended on the absolute concentrations. DM performance was further compared to monitored concentrations at the ESCAPE monitoring sites by calculation of correlation coefficients and in scatterplots.

2.4.3 Development of a hybrid model

In paper IV, model performances were assessed and compared using the proportion of explained variability statistic (R^2), the root mean square error (RMSE) as well as the best visual fit. Model transferability within the study area was assessed using leave one out cross validation (Cyrys, Eeftens et al. 2012).To determine if the fit of the hybrid model compared to the other models was significantly better (i.e. if the additional predictors in the hybrid model compared to other models actually affected the model fit) while considering clustering in the data, the Wald test was used. Finally, the degree of association between monitored NO_x and separate predictors were calculated as partial R^2.

All statistical analyses were performed with STATA Statistical Software (Release 10-11.1; StataCorp, College Station, Texas USA).

2.5 ETHICAL CONSIDERATIONS

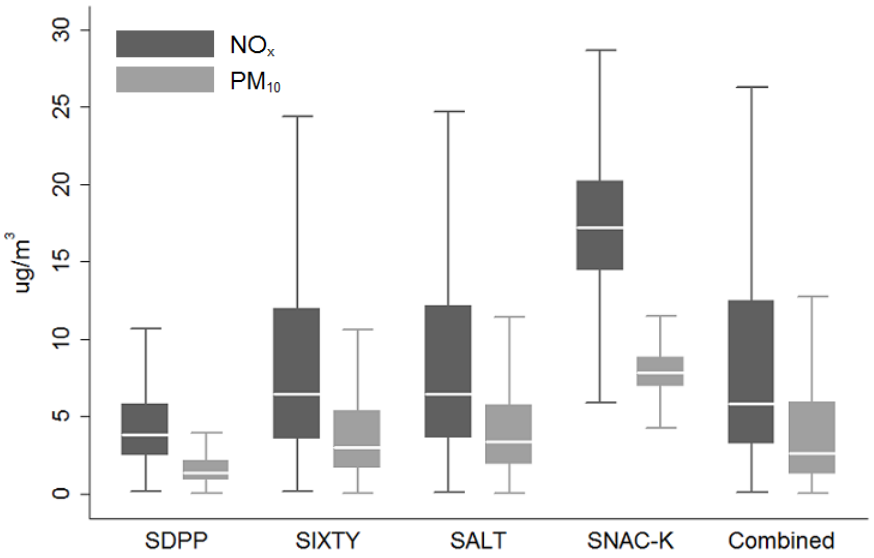
The use of individual data from the four cohorts in Stockholm County, was approved by the Ethics Committee of Karolinska Institutet, Stockholm, Sweden. All cohort studies included in paper III were approved by local ethics committees.

3 RESULTS

3.1 AIR POLLUTION LEVELS IN STOCKHOLM COUNTY

At the cohort specific base-line addresses, dispersion modeled NO_x and PM_{10} was found to vary between the recruitment areas of the four cohorts (Figure 1). Exposure levels were similar in the SIXTY and SALT cohorts, with participants recruited across the whole county. The lowest levels were found in the SDPP cohort, recruited primarily from suburban and rural areas while the highest levels were found in the SNACK cohort from an area in central Stockholm. A similar inter-cohort variability was found in the time-window estimates in paper I and II. During the study period PM_{10} levels stayed relatively constant while some reduction in NO_x levels could be seen (data not shown). Furthermore, the estimated NO_x and PM_{10} levels were found to be highly correlated, Pearson correlation coefficient (R = 0.75-0.9).

Figure 1 Traffic-derived NO_x and PM_{10} ($\mu g/m^3$) levels modeled at study entry addresses in four cohorts from Stockholm



Notes: Box plots are defined by the median (white middle line) and the lower and upper quartiles (box edges) defining the inter quartile range (IQR). The vertical lines (whiskers) are indicating the minimum and maximum range (1.5x IQR) excluding outliers.

3.2 TRAFFIC RELATED AIR-POLLUTION EXPOSURE AND CARDIOVASCULAR DISEASE

In paper I, from 22587 initial cohort participants, 20070 subjects were included in the final analysis. Across cohorts, 6-13% of the participants were excluded, primarily because of incomplete data on covariates or air pollution exposure. During the 238731 person-years at risk, 868 subjects were diagnosed with a stroke, including 89% first ever cases after study entry. Across cohorts the number of stroke events was: SDPP = 130, SIXTY= 160, SALT= 314 and SNACK = 264. The hazard ratio (HR) for cohort specific stroke related to study entry address exposure, ranged between 0.84 and 1.78 for an increment of 20µg/m^3 of NO_x (Figure 2). The combined HR was 1.16 (0.83–1.61). The estimated HR for PM_{10} exposure was similar with a combined HR of 1.14 (0.68–1.90) per 10µg/m^3.

The meta-analyzed effect estimates from cohort specific time-windows did not reveal a clear trend or particularly important exposure periods (Figure 3). The results were similar for both NO_x and PM_{10}, but the wide confidence intervals hampered interpretation.

Moderate heterogeneity was found in both the meta-analysis of stroke risk related to study entry addresses exposure (Higgin's I^2 statistic: 53.7% for NO_x and 66.9% for PM_{10}) and to exposure in time-windows (I^2 = 35.4% to 67.0% for NO_x and 58.3% to 67.0% for PM_{10}).

Figure 2. Exposure at recruitment from road traffic NO_x (per 20µg/m^3) and PM_{10} (per10µg/m^3) and adjusted hazard ratio (HR) of stroke, in four cohorts in Stockholm County, separately and combined.

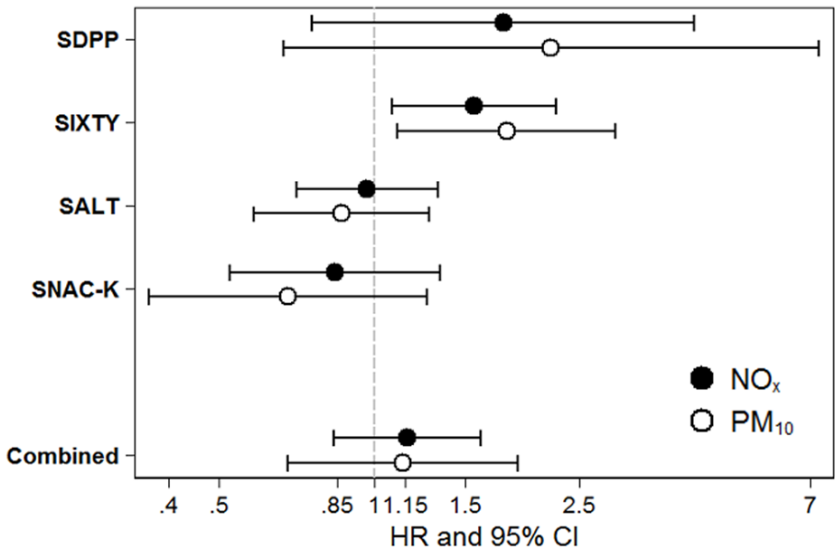
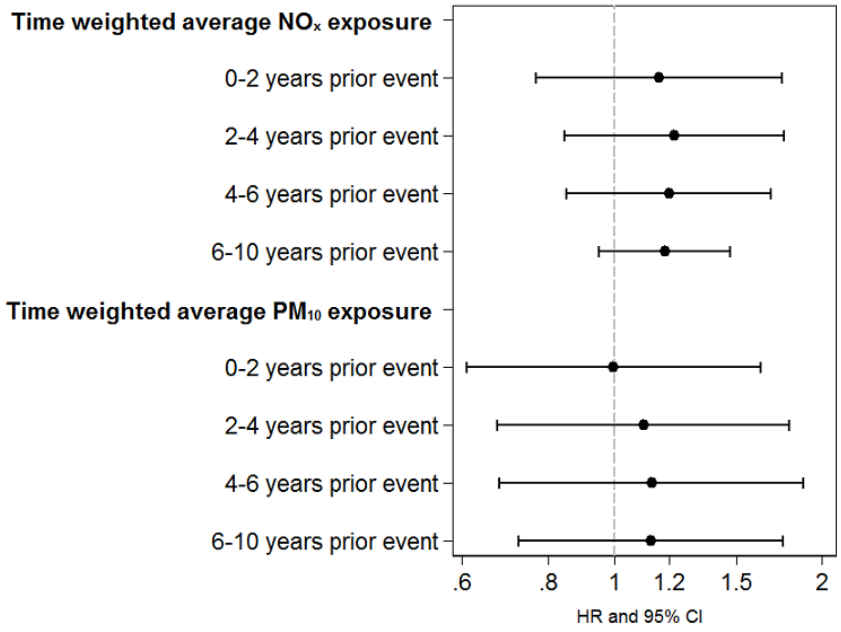


Figure 3. Adjusted Hazard ratios (HR) of stroke, in relation to time-window exposure to NO_x (per 20µg/m^3) and PM_{10} (per 10µg/m^3) from road traffic in a meta-analysis of four cohorts from Stockholm County



In paper II, the 20068 eligible subjects contributed with 237723 person-years at risk and 913 coronary events (CE). The distribution of events across cohorts was 206 each in SDPP, SIXTY, and SNACK, and 295 in the SALT cohort. The cohort-specific HR for CE related to a 20µg/m³ increase in road-traffic NO_x exposure at enrolment addresses ranged between 0.72 and 1.21 when adjusting for covariates (Figure 4). Meta-analysis showed HR of 1.02 (0.82- 1.27). The cohort specific HR for PM_{10} ranged between 0.97 and 1.49 per 10µg/m³ for the different cohorts with a combined HR of 1.14 (0.78- 1.49).

No clear effect modification was found due to hypertension, gender, diabetes status, smoking status or between ever movers and never movers during follow-up (Figure 5). Similar results were found in paper I. Furthermore, for both paper I and paper II, sensitivity analysis did not indicate an association of air pollution with types of stroke or CE, respectively, and no exposure time windows of particular importance were found. However compared to paper I heterogeneity was not detected in paper II by the Higgins I^2 (I^2 = 0.0%) in any the meta-analyzed effect estimates.

Figure 4. Exposure at recruitment from road traffic NO_x (per 20µg/m^3) and PM_{10} (per10µg/m^3) and adjusted hazard ratio (HR) of coronary events, in four cohorts in Stockholm County, separately and combined

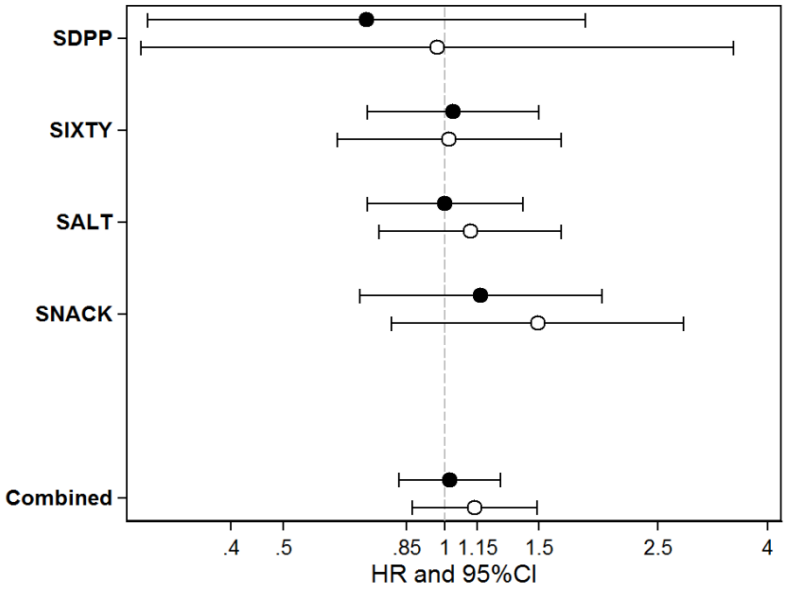
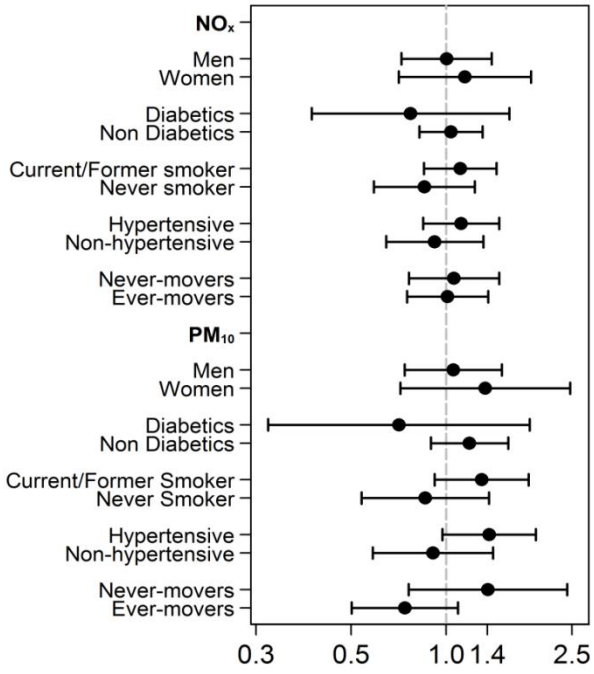


Figure 5. Effect modification by gender, smoking, diabetes, hypertension and relocation during follow-up of the association between NO_x or PM_{10} at recruitment and coronary events in a meta-analysis of four cohorts from Stockholm County



3.3 COMPARISON OF DISPERSION MODELING AND LAND USE REGRESSION

The comparison of dispersion modeled (DM) and land use regression (LUR) modeled NO_2 levels at address sites was done in 13 study areas including 112971 addresses. The range of the correlations between LUR and DM estimates was wide, Pearson *R* 0.19 - 0.89, although for most study areas the correlation was above 0.65 (Table 2). The percentage of agreement within quintiles ranged from 24% to 62% with Kappa statistics ranging from 0.005 to 0.52. In general, the LUR models estimated NO_2 concentrations slightly higher compared to DMs. The size of this discrepancy (absolute difference) could be related to the resolution of the dispersion models. When comparing estimated levels of PM_{10}, 7 study areas and 69591 addresses were used. The correlation between models was found to be lower (Table 1), and the difference in estimated levels larger, compared to model performances for NO_2. Pearson correlations differed across study areas ranging from 0.23 to0.66 and the percentage of agreement by quintiles ranged from 25 to 55%.

For the comparison of estimated levels of fine particles ($PM_{2.5}$), four study regions and 28159 addresses were used. In one area (the Netherlands) a high correlation (Pearson $R = 0.81$) was found while the remaining study areas showed low correlation and low agreement between the estimates.

Table 2 A comparison and descriptive statistics of study area specific LUR and DM model performance at recruitment addresses for 13 European cohorts

Comparison of LUR with DM								
		Continuos: DM = Constant + Slope x LUR					Quintiles	
Study Area	N[a]	Spearman's Rho	Pearson R	Constant	Slope	RMSE	Agreement (%)[b]	Kappa
NO_2								
Umeå region, SE[c]	4575	0.782	0.792	5.17	0.93	2.63	48.3	0.352
Stockholm County, SE[c]	39409	0.791	0.856	-1.98	0.93	2.46	48.9	0.361
Helsinki-Vantaa region, FI[c]	5871	0.762	0.745	2.01	0.52	2.34	43.7	0.297
Bradford, UK[c]	20919	0.820	0.667	-1.62	0.86	3.06	49.2	0.365
London, UK[c]	7089	0.836	0.798	8.55	0.70	4.05	55.2	0.441
Netherlands[c]	7295	0.901	0.891	-2.37	1.13	3.70	61.8	0.523
Ruhr Area, DE[d]	4809	0.428	0.389	28.45	0.30	3.51	31.0	0.138
Basel, SU[c]	1118	0.771	0.768	11.11	0.65	2.71	48.9	0.362
Geneva, SU[c]	737	0.708	0.657	21.73	0.36	2.84	41.4	0.267
Lugano, SU[c]	1090	0.773	0.819	20.43	0.37	1.97	50.2	0.377
Rome, IT[d]	10157	0.406	0.386	33.35	0.36	7.65	29.4	0.120
Barcelona, ES[c]	8402	0.687	0.688	21.41	0.59	8.84	43.3	0.292
Athens, GR[d]	1500	0.207	0.188	42.86	0.10	6.35	23.9	0.005
All	112971							

PM_{10}								
Stockholm County, SE[c]	39409	0.378	0.367	6.83	0.29	2.82	31.2	0.140
London, UK[c]	7089	0.554	0.517	17.94	0.22	.646	55.2	0.441
Netherlands[c]	7295	0.625	0.556	-4.88	1.16	1.91	42.0	0.275
Ruhr Area, DE[d]	4809	0.328	0.346	5.97	0.43	2.18	24.8	0.060
Lugano, SU[c]	1087	0.575	0.659	13.87	0.43	1.25	39.8	0.248
Barcelona, ES[c]	8402	0.495	0.393	24.14	0.35	2.62	33.1	0.163
Athens, GR[d]	1500	0.272	0.233	24.70	0.046	2.36	26.5	0.080
All	69591							

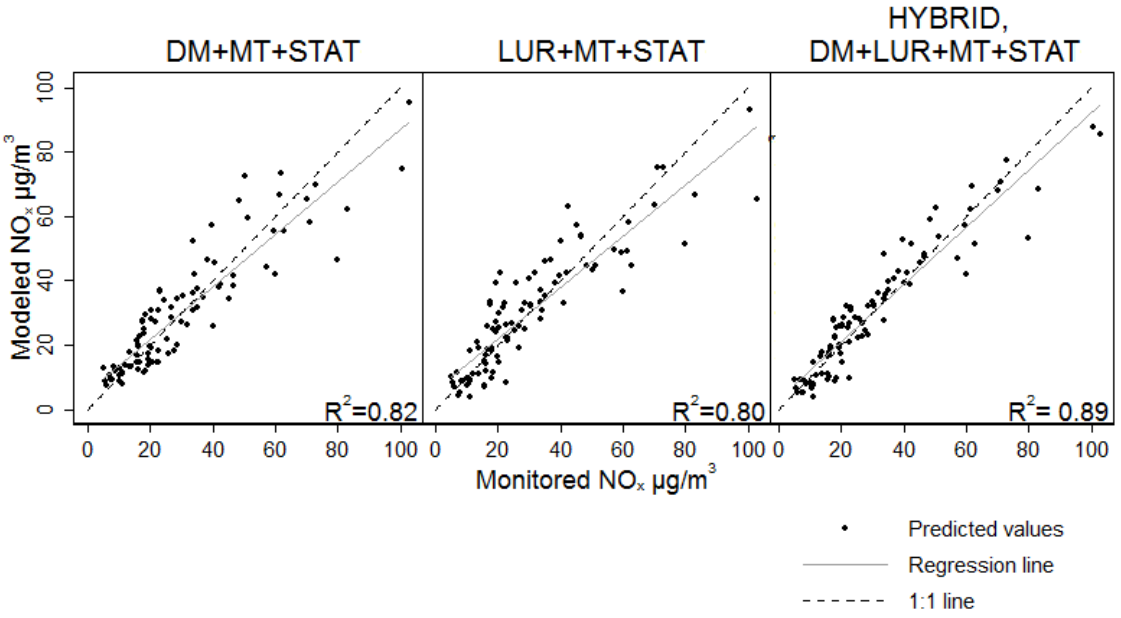
$PM_{2.5}$								
Helsinki-Vantaa region, FI[c]	5871	0.215	0.252	7.85	0.093	.370	25.8	0.073
Netherlands[c]	7295	0.879	0.812	-20.40	2.23	0.41	50.4	0.380
Ruhr Area, DE[d]	4809	0.391	0.327	8.21	0.35	1.12	28.0	0.100
Rome, IT[d]	10544	0.252	0.223	16.03	0.19	1.53	26.5	0.081
All	28159							

[a]Number of residential addresses in the participating cohorts
[b]Percentage of residential addresses falling in the same quintile
[c]Spatial resolution of DM estimates <= 100x100m
[d]Spatial resolution of DM estimates >= 500x500m

3.4 COMBINING DISPERSION MODELING AND LAND USE REGRESSION MODELING

Compared to DM and LUR models separately and when optimized by meteorological (MT) and stationary (STAT) monitoring data (DM+MT+STAT, LUR+MT+STAT), the hybrid model for NO_x including dispersion modeled estimates, LUR variables, meteorological data and the urban background contribution (DM+LUR+MT+STAT) performed the best. The hybrid model captured 89% of the variance in the monitored concentrations (R^2 = 0.89) while having the lowest model RMSE value (Table 3). The predicted NO_x estimates were also more accurate across the whole exposure range (Figure 6). Furthermore, the observed better fit of the hybrid model compared to the other models was statistically significant (p< 0.01.). The predictors included in the hybrid model were the DM estimates, traffic intensity on the nearest street, population density within 100m, global radiation and urban minus rural background NO_x. Predictors in the DM+MT+STAT and LUR+MT+STAT models were similar to the predictors in the hybrid model. All models were robust according to the leave one out cross validation test, suggesting a loss in explained variance of 2-3% if applied at sites other than the training set.

Figure 6. Comparison of model specific NO_x predictions from three modeling scenarios: Dispersion modeling with additional information on global radiation, Land use regression modeling including global radiation and a hybrid model including dispersion modeling, global radiation and LUR components.



DM = Airviro Gauss and SIMAIR road dispersion model
LUR = land use regression model
STAT= stationary monitoring, delta urban NO_x (urban –rural)
MT =meterological variables, (global radiation)

Table 3. Performance evaluation and model structures of the DM, LUR and hybrid model explaining observed levels of NO_x

	Multivariate linear regression			LOOCV	
Model	intercept + (slope x predictors)	R^{2a}	RMSE [b]	R^{2a}	RMSE [b]
DM	9.67 +(1.14x DM)	0.68	12.05	0.66	12.4
LUR	10.12917 +(0.004 x population density 300m) +(.001 x traffic intensity in the nearest street)	0.58	13.90	0.55	14.2
DM+MT+STAT	9.29+ (1.10 x DM)+ (-0.059x global radiation) +(0.70 x Delta Urban NO_x (urban-rural))	0.82	9.14	0.80	9.5
LUR+MT+STAT	1.00+(1.40x Delta Urban NO_x (urban-rural))+(0.001xtraffic intensity on nearest street) + (0.025xpopulation density 100m)+ (-0.046xglobal radiation)	0.80	9.7	0.77	10.1
HYBRID	2,92+ (0.67 x DM)+ (-0.054xglobrad)+ (.0008 x traffic intensity on nearest street)+(0.015x population density 100m) +(0.99 x Delta Urban NO_x (urban-rural)	0.89	7.15	0.87	7.6

DM = dispersion modeled NO_x estimates, MT = meteorological predictors, final models included levels of global radiation from continuous monitoring, STAT = NO_x levels from continuous monitoring, final models included Delta Urban NO_x (urban-rural), LUR = land use regression data, final models included population density (calculated within buffers with specified radii) and traffic intensity. LOOCV = leave one out cross validation

[a] Coefficient of determination

[b] Root mean square error

4 DISCUSSION

4.1 AIR POLLUTION AND CARDIOVASCULAR DISEASE

In paper I and II, we found suggestive evidence of an association between long-term exposure to traffic-related air pollution and the incidence of stroke and coronary events (CE) in a region with comparatively low levels of air pollution. The hazard ratio for stroke per 20μg/m^3 of NO_x and per 10μg/m^3 of PM_{10} were similar, 1.16 (0.83 -1.61) and 1.14 (0.68-1.90), respectively. For CE for the same increments a more convincing effect was suggested for PM_{10}: 1.14 (0.87-1.49) compared to NO_x: 1.02(0.82-1.27). However, in both studies the confidence intervals were wide and overlapping, indicating a low precision of the risk estimates.

For CE, main results were in line with other studies relating incident CE events to long-term exposure to NO_x or NO_2, in particular with the ESCAPE study in which 15% of the CE events were from cohorts used in this thesis (Cesaroni, Forastiere et al. 2014). However, both increased risks (Nafstad, Haheim et al. 2004, Atkinson, Carey et al. 2013, Katsoulis, Dimakopoulou et al. 2014) and no associations have been reported(Lipsett, Ostro et al. 2011, Atkinson, Carey et al. 2013, Cesaroni, Forastiere et al. 2014). Two studies reported comparatively high risk estimates compared to ours, in the UK the risk of heart failure was 1.09 (1.05–1.14) per 3μg/m^3 increase in PM_{10} (Atkinson, Carey et al. 2013) and in Greece the risk of IHD was1.41 (0.91–2.17) per10μg/m^3 of PM_{10} (Katsoulis, Dimakopoulou et al. 2014). In both these study areas the PM_{10} levels were considerably higher than in our study which potentially contributed to the higher risk estimates. On the other hand the risk ratio related to PM_{10} exposure in our study was somewhat larger compared to three studies from the US reporting excess risk related to PM_{10} or coarse particulate for any kind of CE (Puett, Hart et al. 2009, Lipsett, Ostro et al. 2011, Puett, Hart et al. 2011). The difference in estimated risk may be related to differences in particle composition in the US and Europe, where diesel emissions constitute a larger portion of overall air pollution from traffic. The varying dose and composition of air pollutants across cohorts exemplifies some of the challenges when comparing epidemiological studies.

Long-term air pollution exposure has mostly been related to fatal CE rather than to incident or non-fatal CE, both for NO_2/NO_x and particulate exposure. (Nafstad, Haheim et al. 2004, Pope, Burnett et al. 2004, Miller, Siscovick et al. 2007, Puett, Schwartz et al. 2008, Beelen, Hoek et al. 2009, Yorifuji, Kashima et al. 2010, Lipsett, Ostro et al. 2011, Puett, Hart et al. 2011, Crouse, Peters et al. 2012, Raaschou-Nielsen, Andersen et al. 2012, Cesaroni, Badaloni et al. 2013, Beelen, Raaschou-Nielsen et al. 2014). Increased risks for CE mortality were shown in Stockholm in an earlier case-control study for both PM_{10} and NO_2, especially for out of hospital deaths (Rosenlund, Berglind et al. 2006). Similar results were found in a case-control study relating NO_2 to incident MI and a time-series study associating ultrafine particles, PM_{10} and gases to the risk of fatal/non-fatal coronary events in Rome (Stafoggia,

Picciotto et al. 2005, Rosenlund, Picciotto et al. 2008). The evidence suggests that air pollution exposure affects etiological pathways leading to mortality to a greater extent than those contributing to development of non-fatal events. Due to few fatal events of CE in our cohorts the risk estimates in this group were uncertain.

For the relation between long-term air pollution exposure and stroke, studies have presented mixed results. Both borderline and statistically significant effects were found for fatal stroke (Andersen, Kristiansen et al. 2012) as well as stroke sub-types (ischemic or hemorrhagic stroke) (Yorifuji, Kashima et al. 2013) and NO_2. Suggestive evidence was also found in the ESCAPE study for incidence of cerebrovascular events for $PM_{2.5,}$ PM_{10} and coarse particles but not for NO_2 or NO_x (Stafoggia, Cesaroni et al. 2014). On the other hand, studies from England (Atkinson, Carey et al. 2013), Oslo (Nafstad, Haheim et al. 2004) and North America (Pope, Burnett et al. 2004, Puett, Hart et al. 2011, Chen, Goldberg et al. 2013) did not see any elevated stroke risks associated with long-term air pollution exposure. One of the differences between the positive and negative studies was in the type of modeling used. Studies indicating an increased risk were based on exposure modeling with a finer spatial resolution (DM and LUR) which may be particularly important for air pollutants with high spatial variation. For example, when relating monitored $PM_{2.5}$ and NO_2 levels from the nearest urban background monitor to the risk of cerebrovascular disease in women in the US, larger effect estimates were found for $PM_{2.5}$ for within-city exposure differences than for between-city differences, but no effects of NO_2 (Miller, Siscovick et al. 2007). Notably, our study was based on the partial contribution of NO_x and PM_{10} from road traffic only, most other studies estimated total levels. In the Stockholm region, the small-scale spatial differences in residential levels of both PM_{10} and NO_x are dominated by road traffic emissions (Täppefur 2011), enhancing comparability with other studies based on high geographical resolution.

In a sensitivity analysis we did not find a difference between PM_{10} or NO_x exposure and the risk of ischemic stroke compared to the main analysis including all stroke types. This result is consistent with one other European studies on long-term effects of air pollution on stroke (Andersen, Kristiansen et al. 2012)and a US study (Puett, Hart et al. 2011) and with most short term studies. (Wellenius, Schwartz et al. 2005, Chan, Chuang et al. 2006, Andersen, Olsen et al. 2010, Oudin, Stromberg et al. 2010, Vidale, Bonanomi et al. 2010, Andersen, Kristiansen et al. 2012). However, studies from Asia linked air pollution to both ischemic and hemorrhagic stroke (Yamazaki, Nitta et al. 2007, Yorifuji, Kashima et al. 2013). In general, hemorrhagic stroke is less common than ischemic stroke, hampering statistical precision and detection of risk. A recent meta-analysis of 20 studies on stroke found that for long-term PM_{10} exposure, elevated risks were indicated in both North America, Europe and Asia but Asian studies showed a high degree of heterogeneity (Scheers, Jacobs et al. 2015).

We did not find effect modification by gender, diabetes status, smoking status or hypertension for the risk of stroke, however, confidence intervals were wide. The same was seen for CE. Some studies have reported stronger associations for stroke in women, but the

data are not consistent (Miller, Siscovick et al. 2007). For CE we also investigated the modifying effect of changing address during follow-up. The point estimates suggested a stronger risk for individuals living at the same address but confidence intervals included much uncertainty. In a recent cohort study on air pollution effects in the Ruhr area, effect estimates for the risk of a stroke or a coronary event were elevated after excluding individuals moving within 5 years prior study entry, which is in line with our findings (Hoffmann, Weinmayr et al. 2015).

When comparing the model adjusted for the full set of risk factors to the crude model only adjusted for age and sex, we did not find indications of strong confounding of the association between NO_x or PM_{10} and stroke or CE risk. In our main analyses we included data on several known individual risk factors and the effect of contextual confounding was investigated by adding mean income at neighborhood level to the fully adjusted model. This had no clear effect on the risk estimates in any of our studies. Several earlier investigations also reported no major effect of potential individual or contextual confounders on associations between long-term exposure to air pollutants and cardiovascular disease or mortality (Beelen, Raaschou-Nielsen et al. 2014, Cesaroni, Forastiere et al. 2014, Beelen, Hoek et al. 2015). Reanalysis of early prospective cohort studies in the US also demonstrated robustness in the risk estimates when controlling for covariates such as age, sex, race, smoking, alcohol use, marital status, education, body mass, occupational exposures and diet. The same studies reported little contextual confounding on a neighborhood level (Pope and Burnett 2007). However, it should be noted that confounding is study base specific and that low excess risks are observed (and expected) for air pollution exposure, which indicates that careful control of confounding is crucial.

In this thesis heterogeneity refers to the amount of variation in the risk estimates between cohorts that is not due to chance. For stroke, but not for CE, there was some heterogeneity between the cohorts. The same pattern was observed in ESCAPE where heterogeneity was more apparent for stroke (Stafoggia, Cesaroni et al. 2014) compared to CE (Cesaroni, Forastiere et al. 2014). The authors suggested age to be a major source of heterogeneity. In a recent review, heterogeneity was found between cohort studies relating long-term air pollution exposure to cardio-respiratory mortality (Hoek, Krishnan et al. 2013). Differences in particle composition, infiltration of particles indoors, characteristics of the populations and methodological differences in exposure assessment or confounder control were suggested as sources of heterogeneity. The cohort specific study participants in our studies differed in some respects. The catchment area for SDPP with the youngest participants at entry date was sub-urban which resulted in lower exposure levels compared to the other cohorts. The modeling or exposure assessment for this cohort was also based on overall lower resolution (100x100m) compared to other cohorts. However, the choice of resolution was based on assumptions of small scale variability related to the complexity of infrastructure and variations in traffic scenarios and the lower resolution does not have to result in significantly worse exposure attribution. This population also differed in the percent with family history of diabetes (50% of the cohort participants had FHD). In our sensitivity analysis we did not find

excess risks among FHD although, this could only be examined in two of the cohorts. The analysis in the SALT cohort could not be adjusted for occupation status or alcohol consumption which may have led to poorer confounding control, while the SNACK cohort participants were considerably older, had the overall highest exposure levels and lived in a comparatively small area in central Stockholm. For both SDPP and SNACK exposure ranges were narrow which could affect the power of the study. Another potential reason for heterogeneity was the follow-up length which differed between the cohorts. Overall, it is unclear how the differences between the cohorts contributed to the observed heterogeneity.

The role of exposure timing was difficult to assess because of insufficient power. No indications of a time period of exposure modifying the associations were found for CE or for stroke. This is in line with most cohort studies on cardiovascular disease or mortality with follow-up time up to two decades (Nafstad, Haheim et al. 2004, Puett, Hart et al. 2011, Chen, Goldberg et al. 2013, Chen, Zhang et al. 2013). Still most studies did not investigate the role of exposure timing.

4.2 AIR POLLUTION MODELING

The agreement between DM and LUR in paper III was fairly high for NO_2. Out of 13 study areas 7 had a Pearson $R > 0.70$. The agreement was lower for PM_{10} with a median R of 0.39 and $PM_{2.5}$ with a median R of 0.25. The factor that contributed the most to the agreement between the methods for NO_2 was probably data reflecting small-scale traffic variation. Overall, modeling of particulate matter was found to be more difficult compared to NO_2. There are many potential reasons for this, mainly related to the coverage of the PM contribution from other sources, calculations of formation processes of the particles and resuspension. The Gaussian models included background concentrations from stationary monitoring. This was, however, not enough to avoid under-predictions of PM in all DM models. In LUR modeling a source of misclassification may lay in the assumption of a linear relation between air pollution. Furthermore, monitored levels and the following predictor extraction are dependent on the positioning of monitors within the study area. Potential predictors may be missed if important emission sources influencing the measured exposure are not represented

At address sites the model correlation varied depending on study area and the main factor was the resolution of the DM models. Eluerian and computational fluid dynamics models had a less fine spatial resolution (1x1 km to 500×500 m) compared to Gaussian models, for which the resolution at address sites was 500 m or less. The differences in resolution are transferred to the efficacy of capturing small-scale variations in NO_2 and PM_{10}, and there was a lower agreement between more crude DMs and ESCAPE measurements. Similar results were seen when a LUR model developed for national scale monitoring (and using crude map data) and a city specific DM model were applied to 18 independent local monitors in a city (Beelen et al. 2010).

In another study, spatial resolution was found important with regard to modeling at different scales (neighborhood/ urban scale) (Marshal et al. 2008). Other factors affecting modeling includes the area size and area complexity.

In areas where the address-level agreement was high between DM and LUR, both models also correlated well with measurements. Ultimately, we could not determine which model was better, since there was no golden standard to compare model performance with. Furthermore, DMs and chemical transport models are often used for different epidemiological study designs compared to the LUR models such as when studying temporally or spatio-temporally defined exposure trends. In addition, the modeled mean exposure in large grid sizes by DMs, which were estimated around the LUR receptor points, should technically not give the same exposure concentration as the LUR receptors. This could even indicate a bad estimation by one or both models. However, for the sake of comparability across epidemiological studies, the air pollution variations on similar scales within an area should be fairly similar. In earlier studies differences in risk estimates have been seen between studies using DM and LUR for the same cohort addresses. In the ESCAPE study on CE risk (Cesaroni, Forastiere et al. 2014) and stroke risk (Stafoggia, Cesaroni et al. 2014) using LUR models to calculate mean annual exposure, the HR for the cohorts used in this study differed compared to the HR in this thesis using DM. These differences could probably not be explained only by differences in follow-up time and adjustment for confounders. Similar effects were found in a French study, when comparing DM and LUR for the same addresses (Sellier, Galineau et al. 2014). However, another study comparing DM and LUR found a strong correlation for NO_2 (83%) and PM (79%) between the models at both monitoring and address sites (Cyrys, Hochadel et al. 2005). In paper IV we explored the possibility to optimize exposure modeling by creating a hybrid model based on LUR and DM. We found our hybrid model to perform better than DM and LUR separately even when optimizing these models using meteorology and stationary monitoring. The hybrid model was also found to describe the exposure scenario well for the whole concentration range. This result is similar to another hybrid LUR-DM model that also performed better compared to DM and LUR (Wilton, Szpiro et al. 2010). However, to our knowledge no other study incorporating estimated DM concentrations in LUR methodology has been published.

Most of the observed spatial and temporal variance in the monitored data was explained by the DM (adjusted for background levels of NO_x), but the model tended to overestimate the lower NO_x levels and underestimate high levels. Other DM models have shown similar results and suggested explanations include difficulties to model complex traffic situations like heavily trafficked and congested roads (Dijkema, Gehring et al. 2011). Furthermore, emission factors used in DM have been seen to underestimate real-life traffic exposure in several studies (Franco, Kousoulidou et al. 2013).

The inclusion of meteorology and stationary data on urban NO_x levels improved the DM model and narrowed the discrepancy in the lowest and the highest NO_x range. Meteorology was represented by the predictor global radiation which explained more of the variance in monitored NO_x compared to other tested meteorological variables. Similar predictors were included in the LUR model which also included predictor data on traffic intensity on the nearest street and population density within 100 m. Both these variables are typical in LUR modeling (Ryan and LeMasters 2007). The difference in explained variability by these optimized DM and LUR models was small compared to the hybrid model (R^2 = 0.82, 0.80 and 0.89, respectively). This is promising for exposure modeling in areas for which only one model type is available. Several earlier studies have included meteorological components and spatially resolved data in their LUR modeling with success, however, global radiation was not used (Arain, Blair et al. 2007, Gryparis, Coull et al. 2007, Maynard, Coull et al. 2007, Ainslie, Steyn et al. 2008, Mavko, Tang et al. 2008, Su, Brauer et al. 2008). We further hypothesized that the NO_x predictors included in the hybrid model also indicated a potential for optimization of the DM model. As such, global radiation may be a marker of deficiencies in the description of mixing processes in the planetary boundary layer, which is of particular interest in estimation for sub-year periods. Discrepancies between DM modeled and monitored NO_x has been suggested to relate to such problems (Eneroth, Johansson et al. 2006). The predictor traffic intensity at the nearest street may reflect missing coverage of very near traffic or less well calibrated emission factors related to the car fleet by the DM (Franco, Kousoulidou et al. 2013). Population density has earlier been described as a marker of variation in traffic sources and home heating related to urban and rural areas (Brauer, Hoek et al. 2003), while the inclusion of urban levels of NO_x suggests that even an hourly calculated DM model fails to capture the NO_x temporal variability in the city.

It should also be noted that all exposure attribution in this thesis is based on outdoor levels at place of residency, ignoring e.g. the effects of time spent indoors and in traffic. Typical Stockholm homes seem, however, to offer little protection against outdoor air pollution (Wichmann, Lind et al. 2010) and individual exposure to NO_2 in Stockholm has been found to correlate with modeled levels (Bellander, Wichmann et al. 2012).

5 CONCLUSIONS

Based on the presented studies, and considering earlier evidence, it may be concluded that:

- There was suggestive evidence of an association between long-term exposure to air pollution from road traffic, using NO_x and PM_{10} as markers, and incidence of coronary events, primarily ischemic heart disease, as well as stroke. The risk estimates were in line with those found in other studies, although the air pollution levels were lower.

- Some heterogeneity was indicated in risk estimates for stroke between the four included cohorts but it is not clear why this occurred. On the other hand, no heterogeneity was observed in the risk estimates for coronary events.

- No time window of exposure under analysis appeared to be of particular importance for the association between traffic-related air pollution and cardiovascular disease, however, these results were uncertain because of a low statistical power in the analyses.

- Dispersion models and land use regression applied in several urban areas across Europe generated highly correlated estimates for NO_2, but correlations for PM_{10} and $PM_{2.5}$ were only moderate, with large variability between areas. The results indicate that both techniques may be useful for individual exposure assessment in epidemiological studies, but that especially for PM they may produce different risk estimates.

- A hybrid model combining dispersion modeling and land use regression, as well as some other features, performed better than either model alone in predicting average NO_x concentrations within a metropolitan area. This indicates that there is a potential for improvement in currently used methods for exposure assessment.

6 SVENSK SAMMANFATTNING

Luftföroreningar är en viktig miljöfaktor som påverkar den globala sjukdomsförekomsten. Från folkhälsosynpunkt är hjärt- kärlsjukdomar knutna till långtidsexponering för luftföroreningar den viktigaste effekten. Det finns dock begränsad information gällande sambandet med olika typer av hjärt-kärlsjukdomar, som hjärnblödning och hjärtinfarkt, samt sub-typer av dessa. Det är även oklart vilken roll som exponeringstidpunkten spelar samt om hälsorisker i områden med förhållandevis låga exponeringsnivåer, t ex under rådande gränsvärden.

I studier av hälsorisker på befolkningsnivå används exponeringsmodeller som med hög geografisk upplösning kan uppskatta luftföroreningsnivåer. Med hjälp av sådan information och adresshistorik kan exponeringsprofiler för studiepersoner beräknas. Då olika typer av modeller används i olika studier är det viktigt att undersöka om modellerna beräknar föroreningshalter på ett likvärdigt sätt. Bättre beräkningsmetoder behöver även utvecklas, t ex genom kombination av befintliga modelleringstekniker.

Syftet med denna avhandling var att undersöka relationen mellan långtidsexponering för trafikrelaterade luftföroreningar (från avgaser och vägslitage) och risken för hjärnblödning eller hjärtinfarkt i ett område med förhållandevis låga halter av föroreningar. Vidare undersöktes betydelsen av exponeringstidpunkt. Dessutom jämfördes två vanliga metoder för beräkning av luftföroreningsnivåer, dispersionsmodeller (DM) och Land use regression-modeller (LUR), i flera städer i Europa och en hybridmodell baserad på DM och LUR metodik skapades för Stockholm.

I den epidemiologiska analysen ingick 20070 individer från fyra kohorter i Stockholms län som följdes upp under 12 år i genomsnitt. Information gällande olika riskfaktorer för hjärt-kärlsjukdom inhämtades via enkäter och telefonintervjuer vid tiden för rekrytering till respektive kohort. En dispersionsmodell användes för att beräkna fordonsrelaterade luftföroreningar (kväveoxider och partiklar) på alla adresser under uppföljningstiden för undersökningsindividerna och information om hälsoutfall hämtades ur nationella register. Resultaten visade tendenser till samband mellan fordonsrelaterad luftföroreningsexponering och risken för hjärnblödning respektive hjärtinfarkt efter justering för andra faktorer som påverkar risken som exempelvis ålder, kön, rökning och diabetes. Vi kunde inte påvisa någon skillnad i risk i förhållande till exponeringstillfälle under uppföljningstiden.

LUR och DM jämfördes genom att båda modellerna estimerade luftföroreningshalter för samma adresser i 4 till 13 olika europeiska städer/regioner beroende på vilken förorening som studerades. Resultaten visade på en bra överensstämmelse för beräknade halter av kvävedioxid men sämre för partiklar. Vår hybridmodell föreföll prestera bättre än LUR och DM var för sig och gav även idéer för hur dessa två modelleringssätt kan förbättras.

Sammanfattningsvis antyder resultaten att trafikrelaterade föroreningar är associerade med förhöjda risker för hjärnblödning och hjärtinfarkt. Inga statistiskt säkra riskökningar

påvisades dock, vilket kan ha att göra med de förhållandevis låga luftföroreningsnivåerna. De två undersökta typerna av modeller för beräkning av luftföroreningsnivåer stämde bra överens gällande beräkningar kvävedioxid men inte för partiklar, vilket kan ha betydelse för tolkningen av resultat från epidemiologiska undersökningar. Vår hybridmodell tyder på att det finns en betydande förbättringspotential i metodiken för skattning av exponering i epidemiologiska studier.

7 ACKNOWLEDGEMENTS

I would like to take the opportunity to express my gratitude to those whom in one way or another, retrospectively or in present time, made important (and sometimes crucial) contributions to this thesis. Thank you;

The participants of the four cohorts in Stockholm County studied in this thesis.

Göran Pershagen, my main supervisor who is always an inspiration, a role model and a reliable and supportive captain in any kind of weather. Thank you so much for letting me in on the adventure.

Tom Bellander, my co-supervisor who asks the right questions regardless of the topic, follows them up with new ideas on that topic and finds time to explain it all if needed.

Johanna Penell, my second co-supervisor, thanks for all the hands on support based on epi-examples from a veterinary book, the "Veronica Maggio energy boosts" and good times at and in-between conferences.

Jenny Selander, my mentor whose mere presence makes everything as it should be, let us schedule another lunch meeting very soon!

The co-writers of the papers included in this thesis, your ideas and contributions have been truly invaluable.

Ulla Stenius, **Anders Ahlbom** current and previous chairmen of the Institute of Environmental Medicine during this thesis, thank you for governing a great scientific environment making us all strive for excellence!

All ESCAPE project participants across land and sea for productive workshops, interesting meetings, fantastic papers and good times in general. Special thanks **to Sara Nilsson** and **Anders Lundin** for being instrumental in the measurement campaign and **Rob Beelen, Kees de Hoogh** and **Gerard Hoek** whom I also worked close with. To the work package leaders and the steering committee, thanks for steering this colossus of an undertaking (is any more air pollution science even needed?). Related to this, I am very grateful to my supervisor who made me part of the project to begin with.

Christer Johansson, **Kristina Eneroth** and **Boel Lövenheim** for all kinds of dispersion modeling and support but also for the GIS help during a GIS workshop far away.

Claes-Göran Östenson, **Agneta Hilding**, **Ulf de Faire**, **Karin Leander**, **Nancy L. Pedersen**, **Patrik K.E Magnusson**, **Laura Fratiglioni**, **Barbara Caracciolo**, thank you for collecting and maintaining the lovely material in SDPP, SIXTY, SALT and the SNACK cohort.

The magicians, I mean statisticians, **Tomas Lind** and **Niklas Andersson** you are awesome!

Anna Bergström, for helping out in general and for discussions (answering questions) around the fika table.

The current and pre-current inhabitants of corridor C6, for being fun and surprising and for shared scientific endeavors. This non-exhaustive list in no particular order includes, **Charlotta Eriksson**, **Olena Gruzieva**, **Anna Gref** (high five, ni är bäst), **Jesse Thacher**, **Andrei Pyko** (mentally and to some extent geographically we have traveled "to infinity and beyond"), **Erika Schultz**, **Sandra Ekström**, **Jessica Magnusson**, **Simon Kebede Merid**, **Alva Wallas**, **Petter Ljungman**, **Jennifer Protudjer**, (thanks for positive modifications of day-to-day life at neighbour-room level), **Anne-Sophie Merritt**, **Antonis Georgellis**, **André Lauber**, **Mare Löhmus Sundström**, **Auriba Raza**, **Laura Thomas**, **Bettina Julin**, **Andrea Discacciati**, **Niclas Håkansson**, **Susanne Rautiainen Lagerström**, **Nicola Orsini**, **Jinjin Zheng Selin** and many others for being part in similar and in many individually different things!

Erik Melén, thank you for saving my arm during working hours, it did come in handy later on.

The journal club members.

The teachers at IMM (thank you **Matteo Bottai** also for being a co-writer).

My childhood friends whom went to work early in life to provide us all with beer while I was studying, and to the gang of university rascals that spread out after graduation and flourished in their own right.

The **Walle**-family for caring, for the support and for great vacations obviously.

And finally, for the love and all that comes with that, thank you **Ewa Korek**, **Janusz Korek** **Magdalena Korek** and **Ellen Larsdotter Walle** you are truly extraordinary and that's a fact!

Aha, cześć Babcie (**Helena Korek** i **Janina Kondusz**) i dziadki (**Stefan Korek** i **Marian Kondusz**)!! Napisalem książke

jak się czuję Sara?

8 REFERENCES

Ainslie, B., D. G. Steyn, J. Su, M. Buzzelli, M. Brauer, T. Larson and M. Rucker (2008). "A source area model incorporating simplified atmospheric dispersion and advection at fine scale for population air pollutant exposure assessment." Atmospheric Environment **42**(10): 2394-2404.

Andersen, Z. J., L. C. Kristiansen, K. K. Andersen, T. S. Olsen, M. Hvidberg, S. S. Jensen, M. Ketzel, S. Loft, M. Sorensen, A. Tjonneland, K. Overvad and O. Raaschou-Nielsen (2012). "Stroke and long-term exposure to outdoor air pollution from nitrogen dioxide: a cohort study." Stroke **43**(2): 320-325.

Andersen, Z. J., T. S. Olsen, K. K. Andersen, S. Loft, M. Ketzel and O. Raaschou-Nielsen (2010). "Association between short-term exposure to ultrafine particles and hospital admissions for stroke in Copenhagen, Denmark." Eur Heart J **31**(16): 2034-2040.

Arain, M. A., R. Blair, N. Finkelstein, J. R. Brook, T. Sahsuvaroglu, B. Beckerman, L. Zhang and M. Jerrett (2007). "The use of wind fields in a land use regression model to predict air pollution concentrations for health exposure studies." Atmospheric Environment **41**(16): 3453-3464.

Atkinson, R. W., I. M. Carey, A. J. Kent, T. P. van Staa, H. R. Anderson and D. G. Cook (2013). "Long-term exposure to outdoor air pollution and incidence of cardiovascular diseases." Epidemiology **24**(1): 44-53.

Beelen, R., G. Hoek, D. Houthuijs, P. A. van den Brandt, R. A. Goldbohm, P. Fischer, L. J. Schouten, B. Armstrong and B. Brunekreef (2009). "The joint association of air pollution and noise from road traffic with cardiovascular mortality in a cohort study." Occup Environ Med **66**(4): 243-250.

Beelen, R., G. Hoek, O. Raaschou-Nielsen, M. Stafoggia, Z. J. Andersen, G. Weinmayr, B. Hoffmann, K. Wolf, E. Samoli, P. H. Fischer, M. J. Nieuwenhuijsen, W. W. Xun, K. Katsouyanni, K. Dimakopoulou, A. Marcon, E. Vartiainen, T. Lanki, T. Yli-Tuomi, B. Oftedal, P. E. Schwarze, P. Nafstad, U. De Faire, N. L. Pedersen, C. G. Ostenson, L. Fratiglioni, J. Penell, M. Korek, G. Pershagen, K. T. Eriksen, K. Overvad, M. Sorensen, M. Eeftens, P. H. Peeters, K. Meliefste, M. Wang, H. B. Bueno-de-Mesquita, D. Sugiri, U. Kramer, J. Heinrich, K. de Hoogh, T. Key, A. Peters, R. Hampel, H. Concin, G. Nagel, A. Jaensch, A. Ineichen, M. Y. Tsai, E. Schaffner, N. M. Probst-Hensch, C. Schindler, M. S. Ragettli, A. Vilier, F. Clavel-Chapelon, C. Declercq, F. Ricceri, C. Sacerdote, C. Galassi, E. Migliore, A. Ranzi, G. Cesaroni, C. Badaloni, F. Forastiere, M. Katsoulis, A. Trichopoulou, M. Keuken, A. Jedynska, I. M. Kooter, J. Kukkonen, R. S. Sokhi, P. Vineis and B. Brunekreef (2015). "Natural-Cause Mortality and Long-Term Exposure to Particle Components: An Analysis of 19 European Cohorts within the Multi-Center ESCAPE Project." Environ Health Perspect **123**(6): 525-533.

Beelen, R., G. Hoek, P. A. van den Brandt, R. A. Goldbohm, P. Fischer, L. J. Schouten, M. Jerrett, E. Hughes, B. Armstrong and B. Brunekreef (2008). "Long-term effects of traffic-related air pollution on mortality in a Dutch cohort (NLCS-AIR study)." Environ Health Perspect **116**(2): 196-202.

Beelen, R., G. Hoek, D. Vienneau, M. Eeftens, K. Dimakopoulou, X. Pedeli, M.-Y. Tsai, N. Künzli, T. Schikowski and A. Marcon (2013). "Development of NO 2 and NO x land use regression models for estimating air pollution exposure in 36 study areas in Europe–the ESCAPE project." Atmospheric Environment **72**: 10-23.

Beelen, R., O. Raaschou-Nielsen, M. Stafoggia, Z. J. Andersen, G. Weinmayr, B. Hoffmann, K. Wolf, E. Samoli, P. Fischer, M. Nieuwenhuijsen, P. Vineis, W. W. Xun, K. Katsouyanni, K. Dimakopoulou, A. Oudin, B. Forsberg, L. Modig, A. S. Havulinna, T. Lanki, A. Turunen, B. Oftedal, W. Nystad, P. Nafstad, U. De Faire, N. L. Pedersen, C. G. Ostenson, L. Fratiglioni, J. Penell, M. Korek, G. Pershagen, K. T. Eriksen, K. Overvad, T. Ellermann, M. Eeftens, P. H. Peeters, K. Meliefste, M. Wang, B. Bueno-de-Mesquita, D. Sugiri, U. Kramer, J. Heinrich, K. de Hoogh, T. Key, A. Peters, R. Hampel, H. Concin, G. Nagel, A. Ineichen, E. Schaffner, N. Probst-Hensch, N. Kunzli, C. Schindler, T. Schikowski, M. Adam, H. Phuleria, A. Vilier, F. Clavel-Chapelon, C. Declercq, S. Grioni, V. Krogh, M. Y. Tsai, F. Ricceri, C. Sacerdote, C. Galassi, E. Migliore, A. Ranzi, G. Cesaroni, C. Badaloni, F. Forastiere, I. Tamayo, P. Amiano, M. Dorronsoro, M. Katsoulis, A. Trichopoulou, B. Brunekreef and G. Hoek (2014). "Effects of long-term exposure to air pollution on natural-cause mortality: an analysis of 22 European cohorts within the multicentre ESCAPE project." Lancet **383**(9919): 785-795.

Bellander, T., N. Berglind, P. Gustavsson, T. Jonson, F. Nyberg, G. Pershagen and L. Jarup (2001). "Using geographic information systems to assess individual historical exposure to air pollution from traffic and house heating in Stockholm." Environ Health Perspect **109**(6): 633-639.

Bellander, T., J. Wichmann and T. Lind (2012). "Individual exposure to NO2 in relation to spatial and temporal exposure indices in Stockholm, Sweden: the INDEX study." PLoS One **7**(6): e39536.

Brauer, M., G. Hoek, P. van Vliet, K. Meliefste, P. Fischer, U. Gehring, J. Heinrich, J. Cyrys, T. Bellander, M. Lewne and B. Brunekreef (2003). "Estimating long-term average particulate air pollution concentrations: application of traffic indicators and geographic information systems." Epidemiology **14**(2): 228-239.

Briggs, D. J., S. Collins, P. Elliott, P. Fischer, S. Kingham, E. Lebret, K. Pryl, H. van Reeuwijk, K. Smallbone and A. Van Der Veen (1997). "Mapping urban air pollution using GIS: a regression-based approach." International Journal of Geographical Information Science **11**(7): 699-718.

Briggs, D. J., C. de Hoogh, J. Guiliver, J. Wills, P. Elliott, S. Kingham and K. Smallbone (2000). "A regression-based method for mapping traffic-related air pollution: application and testing in four contrasting urban environments." Science of the Total Environment **253**(1-3): 151-167.

Briggs, D. J., K. de Hoogh, C. Morris and J. Gulliver (2008). "Effects of travel mode on exposures to particulate air pollution." Environ Int **34**(1): 12-22.

Brook, R. D., S. Rajagopalan, C. A. Pope, 3rd, J. R. Brook, A. Bhatnagar, A. V. Diez-Roux, F. Holguin, Y. Hong, R. V. Luepker, M. A. Mittleman, A. Peters, D. Siscovick, S. C. Smith, Jr., L. Whitsel and J. D. Kaufman (2010). "Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the American Heart Association." Circulation **121**(21): 2331-2378.

Brunekreef, B. and S. T. Holgate (2002). "Air pollution and health." Lancet **360**(9341): 1233-1242.

Cesaroni, G., C. Badaloni, C. Gariazzo, M. Stafoggia, R. Sozzi, M. Davoli and F. Forastiere (2013). "Long-term exposure to urban air pollution and mortality in a cohort of more than a million adults in Rome." Environ Health Perspect **121**(3): 324-331.

Cesaroni, G., F. Forastiere, M. Stafoggia, Z. J. Andersen, C. Badaloni, R. Beelen, B. Caracciolo, U. de Faire, R. Erbel, K. T. Eriksen, L. Fratiglioni, C. Galassi, R. Hampel, M.

Heier, F. Hennig, A. Hilding, B. Hoffmann, D. Houthuijs, K. H. Jockel, M. Korek, T. Lanki, K. Leander, P. K. Magnusson, E. Migliore, C. G. Ostenson, K. Overvad, N. L. Pedersen, J. P. J, J. Penell, G. Pershagen, A. Pyko, O. Raaschou-Nielsen, A. Ranzi, F. Ricceri, C. Sacerdote, V. Salomaa, W. Swart, A. W. Turunen, P. Vineis, G. Weinmayr, K. Wolf, K. de Hoogh, G. Hoek, B. Brunekreef and A. Peters (2014). "Long term exposure to ambient air pollution and incidence of acute coronary events: prospective cohort study and meta-analysis in 11 European cohorts from the ESCAPE Project." BMJ **348**: f7412.

Chan, C. C., K. J. Chuang, L. C. Chien, W. J. Chen and W. T. Chang (2006). "Urban air pollution and emergency admissions for cerebrovascular diseases in Taipei, Taiwan." Eur Heart J **27**(10): 1238-1244.

Chen, H., M. S. Goldberg, R. T. Burnett, M. Jerrett, A. J. Wheeler and P. J. Villeneuve (2013). "Long-term exposure to traffic-related air pollution and cardiovascular mortality." Epidemiology **24**(1): 35-43.

Chen, R., Y. Zhang, C. Yang, Z. Zhao, X. Xu and H. Kan (2013). "Acute effect of ambient air pollution on stroke mortality in the China air pollution and health effects study." Stroke **44**(4): 954-960.

Chin, M. T. (2015). "Basic mechanisms for adverse cardiovascular events associated with air pollution." Heart **101**(4): 253-256.

Crouse, D. L., P. A. Peters, A. van Donkelaar, M. S. Goldberg, P. J. Villeneuve, O. Brion, S. Khan, D. O. Atari, M. Jerrett, C. A. Pope, M. Brauer, J. R. Brook, R. V. Martin, D. Stieb and R. T. Burnett (2012). "Risk of nonaccidental and cardiovascular mortality in relation to long-term exposure to low concentrations of fine particulate matter: a Canadian national-level cohort study." Environ Health Perspect **120**(5): 708-714.

Cyrys, J., M. Eeftens, J. Heinrich, C. Ampe, A. Armengaud, R. Beelen, T. Bellander, T. Beregszaszi, M. Birk, G. Cesaroni, M. Cirach, K. de Hoogh, A. De Nazelle, F. de Vocht, C. Declercq, A. Dėdelė, K. Dimakopoulou, K. Eriksen, C. Galassi, R. Grąulevičienė, G. Grivas, O. Gruzieva, A. H. Gustafsson, B. Hoffmann, M. Iakovides, A. Ineichen, U. Krämer, T. Lanki, P. Lozano, C. Madsen, K. Meliefste, L. Modig, A. Mölter, G. Mosler, M. Nieuwenhuijsen, M. Nonnemacher, M. Oldenwening, A. Peters, S. Pontet, N. Probst-Hensch, U. Quass, O. Raaschou-Nielsen, A. Ranzi, D. Sugiri, E. G. Stephanou, P. Taimisto, M.-Y. Tsai, É. Vaskövi, S. Villani, M. Wang, B. Brunekreef and G. Hoek (2012). "Variation of NO2 and NOx concentrations between and within 36 European study areas: Results from the ESCAPE study." Atmospheric Environment **62**: 374-390.

Cyrys, J., M. Hochadel, U. Gehring, G. Hoek, V. Diegmann, B. Brunekreef and J. Heinrich (2005). "GIS-based estimation of exposure to particulate matter and NO2 in an urban area: stochastic versus dispersion modeling." Environmental Health Perspectives: 987-992.

Davis, D. L. (2002). "A look back at the London smog of 1952 and the half century since." Environmental health perspectives **110**(12): A734.

de Hoogh, K., M. Wang, M. Adam, C. Badaloni, R. Beelen, M. Birk, G. Cesaroni, M. Cirach, C. Declercq and A. Dedele (2013). "Development of land use regression models for particle composition in twenty study areas in Europe." Environmental science & technology **47**(11): 5778-5786.

de Nazelle, A., S. Fruin, D. Westerdahl, D. Martinez, A. Ripoll, N. Kubesch and M. Nieuwenhuijsen (2012). "A travel mode comparison of commuters' exposures to air pollutants in Barcelona." Atmospheric Environment **59**: 151-159.

DerSimonian, R. and N. Laird (1986). "Meta-analysis in clinical trials." Control Clin Trials **7**(3): 177-188.

Dijkema, M. B., U. Gehring, R. T. van Strien, S. C. van der Zee, P. Fischer, G. Hoek and B. Brunekreef (2011). "A comparison of different approaches to estimate small-scale spatial variation in outdoor NO(2) concentrations." Environ Health Perspect **119**(5): 670-675.

Dockery, D. W., C. A. Pope, 3rd, X. Xu, J. D. Spengler, J. H. Ware, M. E. Fay, B. G. Ferris, Jr. and F. E. Speizer (1993). "An association between air pollution and mortality in six U.S. cities." N Engl J Med **329**(24): 1753-1759.

Durant, J. L., C. A. Ash, E. C. Wood, S. C. Herndon, J. T. Jayne, W. B. Knighton, M. R. Canagaratna, J. B. Trull, D. Brugge, W. Zamore and C. E. Kolb (2010). "Short-term variation in near-highway air pollutant gradients on a winter morning." Atmos Chem Phys **10**(2): 5599-5626.

Eeftens, M., J. Beekhuizen, R. Beelen, M. Wang, R. Vermeulen, B. Brunekreef, A. Huss and G. Hoek (2013). "Quantifying urban street configuration for improvements in air pollution models." Atmospheric Environment **72**: 1-9.

Eeftens, M., R. Beelen, K. de Hoogh, T. Bellander, G. Cesaroni, M. Cirach, C. Declercq, A. Dedele, E. Dons, A. de Nazelle, K. Dimakopoulou, K. Eriksen, G. Falq, P. Fischer, C. Galassi, R. Grazuleviciene, J. Heinrich, B. Hoffmann, M. Jerrett, D. Keidel, M. Korek, T. Lanki, S. Lindley, C. Madsen, A. Molter, G. Nador, M. Nieuwenhuijsen, M. Nonnemacher, X. Pedeli, O. Raaschou-Nielsen, E. Patelarou, U. Quass, A. Ranzi, C. Schindler, M. Stempfelet, E. Stephanou, D. Sugiri, M. Y. Tsai, T. Yli-Tuomi, M. J. Varro, D. Vienneau, S. Klot, K. Wolf, B. Brunekreef and G. Hoek (2012). "Development of Land Use Regression models for PM(2.5), PM(2.5) absorbance, PM(10) and PM(coarse) in 20 European study areas; results of the ESCAPE project." Environ Sci Technol **46**(20): 11195-11205.

Eneroth, K., C. Johansson and T. Bellander (2006). "Exposure comparison between measurements and calculations based on dispersion modelling (EXPOSE)." Stockholm Environment and Health Protection Administration.

Eriksson, A. K., A. Ekbom, F. Granath, A. Hilding, S. Efendic and C. G. Ostenson (2008). "Psychological distress and risk of pre-diabetes and Type 2 diabetes in a prospective study of Swedish middle-aged men and women." Diabet Med **25**(7): 834-842.

Filleul, L. and I. Harrabi (2004). "Do socioeconomic conditions reflect a high exposure to air pollution or more sensitive health conditions?" Journal of epidemiology and community health **58**(9): 802-802.

Franco, V., M. Kousoulidou, M. Muntean, L. Ntziachristos, S. Hausberger and P. Dilara (2013). "Road vehicle emission factors development: A review." Atmospheric Environment **70**: 84-97.

Gruzieva, O., T. Bellander, K. Eneroth, I. Kull, E. Melen, E. Nordling, M. van Hage, M. Wickman, V. Moskalenko, O. Hulchiy and G. Pershagen (2012). "Traffic-related air pollution and development of allergic sensitization in children during the first 8 years of life." J Allergy Clin Immunol **129**(1): 240-246.

Gryparis, A., B. A. Coull, J. Schwartz and H. H. Suh (2007). "Semiparametric latent variable regression models for spatiotemporal modelling of mobile source particles in the greater Boston area." Journal of the Royal Statistical Society Series C-Applied Statistics **56**: 183-209.

Hāggmark, L., K. I. Ivarsson, S. Gollvik and P. O. Olofsson (2000). "Mesan, an operational mesoscale analysis system." Tellus A **52**(1): 2-20.

HEI (2010). "Traffic-Related Air Pollution: A Critical Review of the Literature on Emissions, Exposure, and Health Effects." HEI Special Report 17Health Effects Institute, Boston (2010) [Available: http://pubs.healtheffects.org/getfile.php?u=553 [accessed 8 June 2015]].

Hertel, O., R. Berkowicz and S. Larssen (1991). The operational street pollution model (OSPM). Air Pollution Modeling and Its Application VIII, Springer**:** 741-750.

Higgins, J. P., S. G. Thompson, J. J. Deeks and D. G. Altman (2003). "Measuring inconsistency in meta-analyses." BMJ **327**(7414): 557-560.

Hoek, G., R. M. Krishnan, R. Beelen, A. Peters, B. Ostro, B. Brunekreef and J. D. Kaufman (2013). "Long-term air pollution exposure and cardio- respiratory mortality: a review." Environ Health **12**(1): 43.

Hoffmann, B., G. Weinmayr, F. Hennig, K. Fuks, S. Moebus, C. Weimar, N. Dragano, D. M. Hermann, H. Kalsch, A. A. Mahabadi, R. Erbel and K. H. Jockel (2015). "Air quality, stroke, and coronary events: results of the Heinz Nixdorf Recall Study from the Ruhr Region." Dtsch Arztebl Int **112**(12): 195-201.

Hänninen, O., G. Hoek, S. Mallone, E. Chellini, K. Katsouyanni, C. Gariazzo, G. Cattani, A. Marconi, P. Molnár and T. Bellander (2011). "Seasonal patterns of outdoor PM infiltration into indoor environments: review and meta-analysis of available studies from different climatological zones in Europe." Air Quality, Atmosphere & Health **4**(3-4): 221-233.

Jerrett, M., A. Arain, P. Kanaroglou, B. Beckerman, D. Potoglou, T. Sahsuvaroglu, J. Morrison and C. Giovis (2005). "A review and evaluation of intraurban air pollution exposure models." J Expo Anal Environ Epidemiol **15**(2): 185-204.

Johansson, C., L. Burman and B. Forsberg (2009). "The effects of congestions tax on air quality and health. ." Atmos. Environ. 43, 4843-4854.

Johansson, C., A. Hadenius, P.-Å. Johansson and T. Jonson (1999). "SHAPE: THE STOCKHOLM STUDY ON HEALTH EFFECTS OF AIR POLLUTION AND THEIR ECONOMIC CONSEQUENCES. PART 1: NO2 AND PARTICULATE MATTER IN STOCKHOLM. CONCENTRATIONS AND POPULATION EXPOSURE." VAEGVERKET PUBLIKATION(1999: 41).

Katsoulis, M., K. Dimakopoulou, X. Pedeli, D. Trichopoulos, A. Gryparis, A. Trichopoulou and K. Katsouyanni (2014). "Long-term exposure to traffic-related air pollution and cardiovascular health in a Greek cohort study." Sci Total Environ **490**: 934-940.

Kelly, F. J. and J. C. Fussell (2012). "Size, source and chemical composition as determinants of toxicity attributable to ambient particulate matter." Atmospheric environment **60**: 504-526.

Korek, M. J., T. D. Bellander, T. Lind, M. Bottai, K. M. Eneroth, B. Caracciolo, U. H. de Faire, L. Fratiglioni, A. Hilding, K. Leander, P. K. Magnusson, N. L. Pedersen, C. G. Ostenson, G. Pershagen and J. C. Penell (2015). "Traffic-related air pollution exposure and incidence of stroke in four cohorts from Stockholm." J Expo Sci Environ Epidemiol.

Laden, F., J. Schwartz, F. E. Speizer and D. W. Dockery (2006). "Reduction in fine particulate air pollution and mortality: Extended follow-up of the Harvard Six Cities study." Am J Respir Crit Care Med **173**(6): 667-672.

Lagergren, M., L. Fratiglioni, I. R. Hallberg, J. Berglund, S. Elmståhl, B. Hagberg, G. Holst, M. Rennemark, B.-M. Sjolund and M. Thorslund (2004). "A longitudinal study integrating population, care and social services data. The Swedish National study on Aging and Care (SNAC)." Aging clinical and experimental research **16**(2): 158-168.

Lichtenstein, P., P. F. Sullivan, S. Cnattingius, M. Gatz, S. Johansson, E. Carlstrom, C. Bjork, M. Svartengren, A. Wolk, L. Klareskog, U. de Faire, M. Schalling, J. Palmgren and N. L. Pedersen (2006). "The Swedish Twin Registry in the third millennium: an update." Twin Res Hum Genet **9**(6): 875-882.

Lim, S. S., T. Vos, A. D. Flaxman, G. Danaei, K. Shibuya, H. Adair-Rohani, M. Amann, H. R. Anderson, K. G. Andrews, M. Aryee, C. Atkinson, L. J. Bacchus, A. N. Bahalim, K. Balakrishnan, J. Balmes, S. Barker-Collo, A. Baxter, M. L. Bell, J. D. Blore, F. Blyth, C. Bonner, G. Borges, R. Bourne, M. Boussinesq, M. Brauer, P. Brooks, N. G. Bruce, B. Brunekreef, C. Bryan-Hancock, C. Bucello, R. Buchbinder, F. Bull, R. T. Burnett, T. E. Byers, B. Calabria, J. Carapetis, E. Carnahan, Z. Chafe, F. Charlson, H. Chen, J. S. Chen, A. T. Cheng, J. C. Child, A. Cohen, K. E. Colson, B. C. Cowie, S. Darby, S. Darling, A. Davis, L. Degenhardt, F. Dentener, D. C. Des Jarlais, K. Devries, M. Dherani, E. L. Ding, E. R. Dorsey, T. Driscoll, K. Edmond, S. E. Ali, R. E. Engell, P. J. Erwin, S. Fahimi, G. Falder, F. Farzadfar, A. Ferrari, M. M. Finucane, S. Flaxman, F. G. Fowkes, G. Freedman, M. K. Freeman, E. Gakidou, S. Ghosh, E. Giovannucci, G. Gmel, K. Graham, R. Grainger, B. Grant, D. Gunnell, H. R. Gutierrez, W. Hall, H. W. Hoek, A. Hogan, H. D. Hosgood, 3rd, D. Hoy, H. Hu, B. J. Hubbell, S. J. Hutchings, S. E. Ibeanusi, G. L. Jacklyn, R. Jasrasaria, J. B. Jonas, H. Kan, J. A. Kanis, N. Kassebaum, N. Kawakami, Y. H. Khang, S. Khatibzadeh, J. P. Khoo, C. Kok, F. Laden, R. Lalloo, Q. Lan, T. Lathlean, J. L. Leasher, J. Leigh, Y. Li, J. K. Lin, S. E. Lipshultz, S. London, R. Lozano, Y. Lu, J. Mak, R. Malekzadeh, L. Mallinger, W. Marcenes, L. March, R. Marks, R. Martin, P. McGale, J. McGrath, S. Mehta, G. A. Mensah, T. R. Merriman, R. Micha, C. Michaud, V. Mishra, K. Mohd Hanafiah, A. A. Mokdad, L. Morawska, D. Mozaffarian, T. Murphy, M. Naghavi, B. Neal, P. K. Nelson, J. M. Nolla, R. Norman, C. Olives, S. B. Omer, J. Orchard, R. Osborne, B. Ostro, A. Page, K. D. Pandey, C. D. Parry, E. Passmore, J. Patra, N. Pearce, P. M. Pelizzari, M. Petzold, M. R. Phillips, D. Pope, C. A. Pope, 3rd, J. Powles, M. Rao, H. Razavi, E. A. Rehfuess, J. T. Rehm, B. Ritz, F. P. Rivara, T. Roberts, C. Robinson, J. A. Rodriguez-Portales, I. Romieu, R. Room, L. C. Rosenfeld, A. Roy, L. Rushton, J. A. Salomon, U. Sampson, L. Sanchez-Riera, E. Sanman, A. Sapkota, S. Seedat, P. Shi, K. Shield, R. Shivakoti, G. M. Singh, D. A. Sleet, E. Smith, K. R. Smith, N. J. Stapelberg, K. Steenland, H. Stockl, L. J. Stovner, K. Straif, L. Straney, G. D. Thurston, J. H. Tran, R. Van Dingenen, A. van Donkelaar, J. L. Veerman, L. Vijayakumar, R. Weintraub, M. M. Weissman, R. A. White, H. Whiteford, S. T. Wiersma, J. D. Wilkinson, H. C. Williams, W. Williams, N. Wilson, A. D. Woolf, P. Yip, J. M. Zielinski, A. D. Lopez, C. J. Murray, M. Ezzati, M. A. AlMazroa and Z. A. Memish (2012). "A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010." Lancet **380**(9859): 2224-2260.

Lipsett, M. J., B. D. Ostro, P. Reynolds, D. Goldberg, A. Hertz, M. Jerrett, D. F. Smith, C. Garcia, E. T. Chang and L. Bernstein (2011). "Long-term exposure to air pollution and cardiorespiratory disease in the California teachers study cohort." Am J Respir Crit Care Med **184**(7): 828-835.

Martinelli, N., O. Olivieri and D. Girelli (2013). "Air particulate matter and cardiovascular disease: a narrative review." Eur J Intern Med **24**(4): 295-302.

Mavko, M. E., B. Tang and L. A. George (2008). "A sub-neighborhood scale land use regression model for predicting NO(2)." Science of the Total Environment **398**(1-3): 68-75.

Maynard, D., B. A. Coull, A. Gryparis and J. Schwartz (2007). "Mortality risk associated with short-term exposure to traffic particles and sulfates." Environ Health Perspect **115**(5): 751-755.

Miller, K. A., D. S. Siscovick, L. Sheppard, K. Shepherd, J. H. Sullivan, G. L. Anderson and J. D. Kaufman (2007). "Long-term exposure to air pollution and incidence of cardiovascular events in women." N Engl J Med **356**(5): 447-458.

Nafstad, P., L. L. Haheim, T. Wisloff, F. Gram, B. Oftedal, I. Holme, I. Hjermann and P. Leren (2004). "Urban air pollution and mortality in a cohort of Norwegian men." Environ Health Perspect **112**(5): 610-615.

Nemery, B., P. H. Hoet and A. Nemmar (2001). "The Meuse Valley fog of 1930: an air pollution disaster." Lancet **357**(9257): 704-708.

Newby, D. E., P. M. Mannucci, G. S. Tell, A. A. Baccarelli, R. D. Brook, K. Donaldson, F. Forastiere, M. Franchini, O. H. Franco, I. Graham, G. Hoek, B. Hoffmann, M. F. Hoylaerts, N. Kunzli, N. Mills, J. Pekkanen, A. Peters, M. F. Piepoli, S. Rajagopalan, R. F. Storey, E. A. f. C. P. Esc Working Group on Thrombosis, Rehabilitation and E. S. C. H. F. Association (2015). "Expert position paper on air pollution and cardiovascular disease." Eur Heart J **36**(2): 83-93b.

Omstedt, G., S. Andersson, L. Gidhagen and L. Robertson (2011). "Evaluation of new model tools for meeting the targets of the EU Air Quality Directive: a case study on the studded tyre use in Sweden." International Journal of Environment and Pollution **47**(1-4): 79-96.

Oudin, A., U. Stromberg, K. Jakobsson, E. Stroh and J. Bjork (2010). "Estimation of short-term effects of air pollution on stroke hospital admissions in southern Sweden." Neuroepidemiology **34**(3): 131-142.

Ozkaynak, H., L. K. Baxter, K. L. Dionisio and J. Burke (2013). "Air pollution exposure prediction approaches used in air pollution epidemiology studies." J Expo Sci Environ Epidemiol **23**(6): 566-572.

Pope, C. A., 3rd and R. T. Burnett (2007). "Confounding in air pollution epidemiology: the broader context." Epidemiology **18**(4): 424-426; discussion 427-428.

Pope, C. A., 3rd, R. T. Burnett, G. D. Thurston, M. J. Thun, E. E. Calle, D. Krewski and J. J. Godleski (2004). "Cardiovascular mortality and long-term exposure to particulate air pollution: epidemiological evidence of general pathophysiological pathways of disease." Circulation **109**(1): 71-77.

Puett, R. C., J. E. Hart, H. Suh, M. Mittleman and F. Laden (2011). "Particulate matter exposures, mortality, and cardiovascular disease in the health professionals follow-up study." Environ Health Perspect **119**(8): 1130-1135.

Puett, R. C., J. E. Hart, J. D. Yanosky, C. Paciorek, J. Schwartz, H. Suh, F. E. Speizer and F. Laden (2009). "Chronic fine and coarse particulate exposure, mortality, and coronary heart disease in the Nurses' Health Study." Environ Health Perspect **117**(11): 1697-1701.

Puett, R. C., J. Schwartz, J. E. Hart, J. D. Yanosky, F. E. Speizer, H. Suh, C. J. Paciorek, L. M. Neas and F. Laden (2008). "Chronic particulate exposure, mortality, and coronary heart disease in the nurses' health study." Am J Epidemiol **168**(10): 1161-1168.

Raaschou-Nielsen, O., Z. J. Andersen, S. S. Jensen, M. Ketzel, M. Sorensen, J. Hansen, S. Loft, A. Tjonneland and K. Overvad (2012). "Traffic air pollution and mortality from cardiovascular disease and all causes: a Danish cohort study." Environ Health **11**: 60.

Raaschou-Nielsen, O., O. Hertel, E. Vignati, R. Berkowicz, S. S. Jensen, V. B. Larsen, C. Lohse and J. H. Olsen (2000). "An air pollution model for use in epidemiological studies: evaluation with measured levels of nitrogen dioxide and benzene." Journal of exposure analysis and environmental epidemiology **10**(1): 4-14.

Ragettli, M. S., H. C. Phuleria, M. Y. Tsai, C. Schindler, A. de Nazelle, R. E. Ducret-Stich, A. Ineichen, L. Perez, C. Braun-Fahrlander, N. Probst-Hensch and N. Kunzli (2015). "The relevance of commuter and work/school exposure in an epidemiological study on traffic-related air pollution." J Expo Sci Environ Epidemiol **25**(5): 474-481.

Roholm, K. (1937). "The fog disaster in the Meuse Valley, 1930: A fluorine intoxication." J. Industr. Hyg. and Toxicol **19**: 126-137.

Rosenlund, M., T. Bellander, T. Nordquist and L. Alfredsson (2009). "Traffic-generated air pollution and myocardial infarction." Epidemiology **20**(2): 265-271.

Rosenlund, M., S. Picciotto, F. Forastiere, M. Stafoggia and C. A. Perucci (2008). "Traffic-related air pollution in relation to incidence and prognosis of coronary heart disease." Epidemiology **19**(1): 121-128.

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

Santoni, G., S. Angleman, A. K. Welmer, F. Mangialasche, A. Marengoni and L. Fratiglioni (2015). "Age-related variation in health status after age 60." PLoS One **10**(3): e0120077.

Scheers, H., L. Jacobs, L. Casas, B. Nemery and T. S. Nawrot (2015). "Long-Term Exposure to Particulate Matter Air Pollution Is a Risk Factor for Stroke Meta-Analytical Evidence." Stroke **46**(11): 3058-3066.

Schwarze, P. E., J. Ovrevik, M. Lag, M. Refsnes, P. Nafstad, R. B. Hetland and E. Dybing (2006). "Particulate matter properties and health effects: consistency of epidemiological and toxicological studies." Hum Exp Toxicol **25**(10): 559-579.

Sellier, Y., J. Galineau, A. Hulin, F. Caini, N. Marquis, V. Navel, S. Bottagisi, L. Giorgis-Allemand, C. Jacquier, R. Slama, J. Lepeule and E. M.-C. C. S. Group (2014). "Health effects of ambient air pollution: do different methods for estimating exposure lead to different results?" Environ Int **66**: 165-173.

Shah, A. S., K. K. Lee, D. A. McAllister, A. Hunter, H. Nair, W. Whiteley, J. P. Langrish, D. E. Newby and N. L. Mills (2015). "Short term exposure to air pollution and stroke: systematic review and meta-analysis." BMJ **350**: h1295.

Stafoggia, M., G. Cesaroni, A. Peters, Z. J. Andersen, C. Badaloni, R. Beelen, B. Caracciolo, J. Cyrys, U. de Faire, K. de Hoogh, K. T. Eriksen, L. Fratiglioni, C. Galassi, B. Gigante, A. S. Havulinna, F. Hennig, A. Hilding, G. Hoek, B. Hoffmann, D. Houthuijs, M. Korek, T. Lanki, K. Leander, P. K. Magnusson, C. Meisinger, E. Migliore, K. Overvad, C. G. Ostenson, N. L. Pedersen, J. Pekkanen, J. Penell, G. Pershagen, N. Pundt, A. Pyko, O. Raaschou-Nielsen, A. Ranzi, F. Ricceri, C. Sacerdote, W. J. Swart, A. W. Turunen, P. Vineis, C. Weimar, G. Weinmayr, K. Wolf, B. Brunekreef and F. Forastiere (2014). "Long-Term Exposure to Ambient Air Pollution and Incidence of Cerebrovascular Events: Results from 11 European Cohorts within the ESCAPE Project." Environ Health Perspect.

Stafoggia, M., S. Picciotto, F. Forastiere, D. D'Ippoliti, G. Cattani, A. Marconi and C. A. Perucci (2005). "[Air pollution and fatal and non fatal coronary events in Rome]." Epidemiol Prev **29**(1): 40-47.

Su, J. G., M. Brauer, B. Ainslie, D. Steyn, T. Larson and M. Buzzelli (2008). "An innovative land use regression model incorporating meteorology for exposure analysis." Science of the Total Environment **390**(2-3): 520-529.

Täppefur, M. (2011). "Air pollution in Stockholm and Uppsala Counties, and in Gävle and Sandviken municipalities . Emission data for year 2011." SLB-ANALYS **APRIL 2013**.

Wandell, P. E., A. Wajngot, U. de Faire and M. L. Hellenius (2007). "Increased prevalence of diabetes among immigrants from non-European countries in 60-year-old men and women in Sweden." Diabetes Metab **33**(1): 30-36.

Wellenius, G. A., J. Schwartz and M. A. Mittleman (2005). "Air pollution and hospital admissions for ischemic and hemorrhagic stroke among medicare beneficiaries." Stroke **36**(12): 2549-2553.

WHO (2013). "Review of Evidence on Health Aspects of Air Pollution—REVIHAAP Project: Technical Report." WHO, Copenhagen.

WHO (2014a). "WHO Fact sheet available from (http://www.who.int/mediacentre/news/releases/2014/air-pollution/en/) ".

WHO (2014b). WHO "Indoor Air Quality Guidelines: Household Fuel Combustion" (http://www.ncbi.nlm.nih.gov/books/NBK264290/). Geneva.

Wichmann, J., T. Lind, M.-M. Nilsson and T. Bellander (2010). "PM 2.5, soot and NO 2 indoor–outdoor relationships at homes, pre-schools and schools in Stockholm, Sweden." Atmospheric Environment **44**(36): 4536-4544.

Vidale, S., A. Bonanomi, M. Guidotti, M. Arnaboldi and R. Sterzi (2010). "Air pollution positively correlates with daily stroke admission and in hospital mortality: a study in the urban area of Como, Italy." Neurological Sciences **31**(2): 179-182.

Wilkins, E. (1954). "Air pollution and the London fog of December 1952." JR Sanit. Inst.;(United States) **74**.

Wilton, D., A. Szpiro, T. Gould and T. Larson (2010). "Improving spatial concentration estimates for nitrogen oxides using a hybrid meteorological dispersion/land use regression model in Los Angeles, CA and Seattle, WA." Sci Total Environ **408**(5): 1120-1130.

Yamazaki, S., H. Nitta, M. Ono, J. Green and S. Fukuhara (2007). "Intracerebral haemorrhage associated with hourly concentration of ambient particulate matter: case-crossover analysis." Occup Environ Med **64**(1): 17-24.

Yorifuji, T., S. Kashima, T. Tsuda, K. Ishikawa-Takata, T. Ohta, K. Tsuruta and H. Doi (2013). "Long-term exposure to traffic-related air pollution and the risk of death from hemorrhagic stroke and lung cancer in Shizuoka, Japan." Sci Total Environ **443**: 397-402.

Yorifuji, T., S. Kashima, T. Tsuda, S. Takao, E. Suzuki, H. Doi, M. Sugiyama, K. Ishikawa-Takata and T. Ohta (2010). "Long-term exposure to traffic-related air pollution and mortality in Shizuoka, Japan." Occup Environ Med **67**(2): 111-117.

Karolinska Institutet

Institutet för Miljömedicin

Exposure to air pollution from road traffic and cardiovascular disease with a focus on exposure modeling

AKADEMISK AVHANDLING
som för avläggande av medicine doktorsexamen vid Karolinska Institutet offentligen försvaras i sal Farmakologi, Nanna Svartz väg 2

Fredagen den 11 December, 2015, kl 13.00

av
Michal Korek

Huvudhandledare:
Professor Göran Pershagen
Karolinska Institutet
Institutet för Miljömedicin

Bihandledare:
Professor Tom Bellander
Karolinska Institutet
Institutet för Miljömedicin

Dr. Johanna Penell
University of Surrey-Guildford
Faculty of Health & Medical Sciences

Fakultetsopponent:
Professor Ole Hertel
Aarhus University
Department for Environmental Science - Atmospheric chemistry and physics

Betygsnämnd:
Professor Bo Burström
Karolinska Institutet
Institutionen för Folkhälsovetenskap

Professor Johan Frostegård
Karolinska Institutet
Institutet för Miljömedicin

Professor Hans-Christen Hansson
Stockholms Universitet
Institutionen för Miljövetenskap och analytisk kemi

Stockholm 2015

ABSTRACT

Air pollution is an important environmental health factor contributing to the burden of disease. From a public health point of view cardiovascular effects of long-term exposure are predominant, primarily coronary events and stroke. However, sub-types of disease have not been well investigated and few studies have been conducted in areas with lower air pollution levels. The role of timing of exposure is also unclear.

In epidemiological studies different types of models are used to estimate exposure of study participants. It is therefore important to understand if modeled levels are similar for different model types. Furthermore, there is a need to develop better modeling techniques, and it has been proposed to combine models into so called hybrid models.

The aim of this thesis was to investigate the relation between individual long-term air pollution exposure from road traffic and the risk of coronary events and stroke in an area with comparatively low exposure levels, while considering timing of exposure. Furthermore a comparison of dispersion modeling (DM) and land use regression (LUR) was done in several study areas and a hybrid model based on DM and LUR was developed for Stockholm.

From four cohorts in Stockholm County, 20070 individuals were followed for an average of 12 years. Information on covariates was available from questionnaires and interviews from the time of recruitment. Air pollution exposure from traffic was assessed at residential addresses during follow-up using dispersion modeled levels of nitrogen oxides (NO_x), as a marker of exhaust emissions, and particles with an aerodynamic diameter of <10 µm (PM_{10}), as a marker of road dust. A suggestive association between road traffic exposure at the recruitment address and cardiovascular disease incidence was seen. For NO_x the hazard ratio for stroke and coronary events per 20µg/m^3 was 1.16 (0.83 -1.61) and 1.02 (0.82-1.27), respectively. Corresponding hazard ratios for PM_{10} were 1.14 (0.68-1.90) and 1.14 (0.87-1.49), respectively, per 10µg/m^3. Results did not appear to be modified by covariates, disease sub-types or exposure time windows.

LUR models and DMs were compared in 4 to13 European study areas depending on the pollutant. At study addresses, the median Pearson correlation (range) for annual mean concentrations of NO_2, PM_{10} and $PM_{2.5}$ were: 0.75 (0.19–0.89), 0.39 (0.23–0.66) and 0.29 (0.22–0.81). A hybrid model was developed for Stockholm for 93 bi-weekly NO_x observations using DM estimates, LUR variables, stationary monitoring and individual meteorological factors. The hybrid model explained NO_x levels at monitoring stations better (R^2 =89%) than the LUR and DM models (R^2 =58% and R^2 =68%, respectively).

In conclusion, our results suggest an elevated risk of coronary events and stroke related to traffic air pollution exposures in Stockholm County, however, no modification by time window of exposure could be detected. On average, estimates from LUR and DMs correlate well for NO_2 but less so for particulates. To combine DM and LUR seems promising for increasing the quality of the exposure assessment.

OPEN

Journal of Exposure Science and Environmental Epidemiology (2015) 25, 517–523

www.nature.com/jes

ORIGINAL ARTICLE

Traffic-related air pollution exposure and incidence of stroke in four cohorts from Stockholm

Michal J. Korek[1], Tom D. Bellander[1,2], Tomas Lind[2], Matteo Bottai[3], Kristina M. Eneroth[4], Barbara Caracciolo[5], Ulf H. de Faire[1], Laura Fratiglioni[5,6,7], Agneta Hilding[8], Karin Leander[1], Patrik K. E. Magnusson[9], Nancy L. Pedersen[9], Claes-Göran Östenson[8], Göran Pershagen[1] and Johanna C. Penell[1]

We investigated the risk of stroke related to long-term ambient air pollution exposure, in particular the role of various exposure time windows, using four cohorts from Stockholm County, Sweden. In total, 22,587 individuals were recruited from 1992 to 2004 and followed until 2011. Yearly air pollution levels resulting from local road traffic emissions were assessed at participant residences using dispersion models for particulate matter (PM_{10}) and nitrogen oxides (NO_X). Cohort-specific hazard ratios were estimated for time-weighted air pollution exposure during different time windows and the incidence of stroke, adjusted for common risk factors, and then meta-analysed. Overall, 868 subjects suffered a non-fatal or fatal stroke during 238,731 person-years of follow-up. An increment of 20 $\mu g/m^3$ in estimated annual mean of road-traffic related NO_X exposure at recruitment was associated with a hazard ratio of 1.16 (95% CI 0.83–1.61), with evidence of heterogeneity between the cohorts. For PM_{10}, an increment of 10 $\mu g/m^3$ corresponded to a hazard ratio of 1.14 (95% CI 0.68–1.90). Time-window analyses did not reveal any clear induction-latency pattern. In conclusion, we found suggestive evidence of an association between long-term exposure to NO_X and PM_{10} from local traffic and stroke at comparatively low levels of air pollution.

Journal of Exposure Science and Environmental Epidemiology (2015) **25,** 517–523; doi:10.1038/jes.2015.22; published online 1 April 2015

Keywords: epidemiology; exposure modeling; particulate matter; personal exposure

INTRODUCTION

Stroke is a leading cause of disease and death in the Western world although the incidence has decreased in recent decades.[1] The body of evidence regarding long-term air pollution exposure, especially to respirable particles, and various cardiovascular risk factors and diseases is growing.[2] For example, associations have been reported with the development of atherosclerosis, hyper tension, ischemic heart disease and cardiovascular mortality. Proposed pathophysiological mechanisms include systemic inflammation and oxidative stress, imbalance in the autonomic nervous system, endothelial dysfunction, vasoconstriction and thrombosis. Although some studies show associations with cerebrovascular incidence or mortality in relation to long-term air pollution exposure, evidence is limited and conflicting.[3–6] Data regarding the role of air pollution exposure for ischemic *vs* hemorrhagic stroke and the influence of potential effect modifiers are sparse and ambiguous.

In understanding the mechanisms behind effects of long-term exposure to ambient air pollution on cardiovascular disease, as well as to predict consequences of preventive measures, it is important to assess the role of timing of exposure in relation to occurrence of the adverse outcomes. It has been shown, that cardiovascular and other mortality decreased within a year after a coal ban in Dublin, Ireland, which drastically lowered the black smoke levels.[7] A follow-up of the Harvard Six City study and a study of myocardial infarction survivors, suggested that the relevant exposure period for mortality was the past few years[8,9] Studies of air pollution effects on cardiovascular disease or mortality have generally not detected marked associations with timing of exposure during follow-up periods of up to two decades.[3,10,11] Most of the studies on cerebrovascular effects of air pollution, however, did not address this issue.

The aim of the current study was to assess individual long-term exposure to air pollution from road traffic in relation to stroke incidence, in an area with relatively low air pollution levels. We took changes in residence and in annual exposure levels during follow-up into account and investigated different types of stroke, in a combined analysis of four cohorts from Stockholm County, Sweden.

METHODS

Study Population

The study included four cohorts based in Stockholm County, Sweden. The Stockholm Diabetes Preventive Program (SDPP),[12] a population-based prospective study, recruited 3128 men in 1992–1994 and 4821 women in 1996–1998 from five municipalities in Stockholm County. The study participants were 35–56 years old at recruitment. None had previously

[1]Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden; [2]Centre for Occupational and Environmental Medicine, Stockholm County Council, Stockholm, Sweden; [3]Division of Biostatistics, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden; [4]Environment and Health Administration, Stockholm, Sweden; [5]Department of Neurobiology, Care Sciences and Society, Aging Research Center, Karolinska Institutet, Stockholm, Sweden; [6]Stockholm Gerontology Research Center, Stockholm, Sweden; [7]Division of Clinical Geriatrics, Karolinska University Hospital, Stockholm, Sweden; [8]Department of Molecular Medicine and Surgery, Karolinska University Hospital, Stockholm, Sweden and [9]Department of Medical Epidemiology and Biostatistcs, Karolinska Institutet, Stockholm, Sweden. Correspondence: Michal Korek, Karolinska Institutet, Institute of Environmental Medicine, Nobels väg 13 Box 210, Stockholm, SE-171 77, Sweden.
Tel.: +46 8 524 87457. Fax: +46 8 304571.
E-mail: michal.korek@ki.se
Received 5 August 2014; revised 23 February 2015; accepted 23 February 2015; published online 1 April 2015

diagnosed diabetes and half of the cohort (53%) had a family history of diabetes (one first degree relative or two second degree relatives), while the other half was selected to match on age and sex of the first half. The cohort study of 60 year olds (SIXTY)[13] invited a random population sample consisting of one-third of all men and women who were living in Stockholm County and turned 60 years of age between August 1997 and March 1999. The SIXTY study included 4232 subjects. The Screening Across the Lifespan Twin Study (SALT)[14] screened all twins born in Sweden before 1958 for the most common complex diseases with a focus on cardiovascular diseases. Recruitment took place during 1998–2002. In the present study, the SALT participants residing in Stockholm County at recruitment were included, resulting in 7043 subjects with an age range of 42–100 years of age at recruitment. The Swedish National study of Aging and Care in Kungsholmen (SNAC-K)[15] included randomly sampled individuals $>=60$ years of age between March 2001 and June 2004 from a central area in Stockholm City. A total of 3363 subjects of 60–104 years of age were recruited.

In all four cohorts, individual data collected at enrolment on socio-demographic characteristics such as occupation status and education were obtained from questionnaires that also provided information on lifestyle factors, including smoking status, levels of physical activity and alcohol consumption. Information on diabetes and hypertension was obtained either from questionnaires or clinical data. Socio-economic variables aggregated at neighborhood level were retrieved from Statistics Sweden. Neighborhoods consisted of small geographical units with an average population of 1000–2000 subjects considered to be homogenous with regard to socio-economic characteristics. The study was approved by the Ethical Review Board Stockholm, Sweden.

Outcome Data

Data on individual stroke events were retrieved from the National Hospital Discharge Registry and the National Cause of Death Registry, including data from February 1964 to December 2011. A stroke event was defined based on the International Classification of Diseases (ICD) version 9 and 10: hospitalizations with principal diagnosis of ischemic stroke (ICD9: 433; 434; ICD10: I63), hemorrhagic stroke (ICD9: 431; ICD10: I61), unspecified stroke (ICD9: 436; ICD10: I64) and out-of-hospital deaths from cerebrovascular diseases (ICD9: 431–436; ICD10: I61-I64). If the person passed away within 28 days after a stroke event, the event was classified as fatal. Only stroke events after recruitment of the respective cohort were included in analyses, whereas earlier events were used to classify later events as non-incident.

Exposure Assessment

Long-term exposure to ambient air pollution from road traffic was estimated based on a methodology described in detail elsewhere.[16] Briefly, residential histories retrieved from the Swedish tax authorities were available for all cohort participants from 1991 until 2010, including data on all residential addresses and the date from which the person resided at each particular address. The residential address was known also earlier than 1991 for those moving in before this year. In case of emigration from Sweden or Stockholm County, this was recorded with a specific date used for censoring (see below). The residential addresses within Stockholm County were geocoded, 90% by automatic matching against the Swedish Mapping Cadastral and Land Registration Authority Databases, and an additional 9% manually. One percent of the addresses could not be geocoded, mainly because of insufficient address details.

Annual mean concentrations of nitrogen oxides (NO_x) and particulate matter with an aerodynamic diameter of less than 10 micrograms (PM_{10}) were calculated using a wind model and a Gaussian air quality dispersion model, both part of the Airviro Air Quality Management System (SMHI, Norrköping, Sweden; http://airviro.smhi.se). The emission inventory of the Stockholm and Uppsala County Air Quality Management Association supplied the input to the model. Information in this database has been updated yearly by the municipalities in the region since 1993. It is a geographic information system and contains detailed information about emissions from, for example, road and ferry traffic, petrol stations, industrial areas and households.[17] For the present study, only emissions from local road traffic were included as it is the dominating source of both NO_x and PM_{10}.[18,19] NO_x was used as a marker of road traffic-derived combustion pollutants, while PM_{10} was used as a marker for road wear. Emission factors for NO_x and exhaust particles from road traffic were obtained from the EVA model of the Swedish Transport Administration. Emission factors for non-exhaust PM (mainly road wear but including some contributions from brake and tyre wear) were obtained using NO_x as tracer for traffic emissions.[20] In Stockholm, road wear increases drastically because of the use of studded tyres and traction sand on streets during winter; up to 90% of the locally emitted PM_{10} may be due to road abrasion.[19,21] The contributions from road, tyre and brake wear in Stockholm is further analysed in a modelling study which clearly shows the dominance of road wear and that the surface moisture, and subsequent retention and suppression of suspension, also influences the PM_{10} levels.[22]

The model resolution for inner Stockholm and the urban parts of the municipalities of Solna, Järfälla and Södertälje was 25 m grid cells. The rest of Stockholm County had a resolution of 100 m or 500 m, respectively, in urban and rural areas. The model estimates air pollution concentrations 2 m above ground level and handles buildings by using a roughness parameter.[23] This results in underestimated concentrations in street canyons with heavy traffic. Therefore, a street canyon contribution was calculated using the SMHI-Airviro street canyon model (http://airviro.smhi.se). This contribution was added to all addresses with multistory houses on both sides within 30 m of the most polluted street segments in the inner city of Stockholm, and corresponded to air pollution levels at half the building height at those addresses.

The SMHI-Airviro Gaussian model has been validated in a number of previous studies, for example, Johansson *et al.*[24] and Eneroth *et al.*[25] The comparison between time series calculations of annual mean NO_x concentrations and urban background measurements from a continuous monitoring station in Stockholm during the period 1998–2005 provided R^2 of 0.74–0.80 for different years.[26] Within the European Study of Cohorts for Air Pollution Effects (ESCAPE), differences between dispersion model estimates and monitoring results were explored.[27] In Stockholm, SMHI-Airviro Gaussian model calculations were compared with measured concentrations of NO_2 and PM_{10}, respectively, at 39 and 19 monitoring sites, providing Spearman rank correlations of 0.755 for NO_2 and 0.580 for PM_{10}.

Levels of NO_x and PM_{10} resulting from local road traffic emissions were calculated for all geocoded addresses for every year from 1987 until the end of follow-up. Reduced emissions of exhaust particulates and NO_x due to stricter European vehicle regulations are included in the EVA emission model. The emission factors for non-exhaust PM were assumed to be constant as the proportion of cars with studded tyres in Stockholm has been relatively stable during the period of the present study. To compensate for trends in traffic volumes, the calculated levels of NO_x and PM_{10} were re-scaled based on measured traffic flow in and out of the regional centre of Stockholm on an annual basis. Concentrations of NO_2 were calculated based on modelled concentrations of NO_x. An empirical non-linear relationship was derived for each year based on measured concentrations of NO_x and NO_2 in Stockholm and its surroundings.The annual subject-specific exposure estimates were used to construct study entry and time-weighted exposure concentrations during different time windows.

Statistics

Cohort-specific Cox proportional hazard regression analysis was used to estimate hazard ratios (HRs) of stroke associated with long-term ambient air pollution exposure. NO_x and PM_{10} were analysed separately. Person-time at risk was calculated from enrolment into the study until stroke, death from another cause, emigration (i.e., to an address without information on air pollution exposure) or end of study (31 December 2011), whichever event occurred first. Age was used as the underlying time scale in all models. Risk estimates were calculated as HRs with 95% confidence intervals using increments of 20 $\mu g/m^3$ for NO_x and 10 $\mu g/m^3$ for PM_{10}. The data were divided into 6-month periods, allowing us to use exposure to PM_{10} and NO_x as time-varying covariates and to adjust for calendar year in 5-year periods. First, concentrations of NO_x and PM_{10} at the study entry address of each individual were used as exposure variables in cohort-specific analyses. Second, the NO_x and PM_{10} exposures were calculated for each subject for each subject-specific 6-month period during the follow-up. Exposure time windows were then created for 6–10-, 4–6-, 2–4- and 0–2-year intervals prior the end date of every 6-month interval during the follow-up.

Adjustment models were defined *a priori* and covariates were chosen based on the literature and available data from the cohorts. The fully adjusted models had a common set of individual-level covariates including gender, education level, smoking status, smoking intensity among current smokers and socio-economic index. This index was based on current or last (if retired) profession and categorized into low (blue collar worker),

medium (low and intermediate level white collar worker, and self-employed) and high (high-level white-collar worker). Additional variables were included if present for at least two cohorts.

The proportional-hazard assumption for all covariates was investigated. If any variable in the individual cohort models violated this assumption, effect estimates were compared with a stratified Cox analysis for that cohort.[3] In the analyses, we included only individuals with complete data on exposure estimates and confounders. We did not include confounders with missing data for more than 20% of the individuals; hence for SALT, alcohol consumption and occupational status and for SNAC-K physical activity was excluded.

To increase power, we combined the effect estimates of the four cohorts using a random effect meta-analysis model.[28] We investigated heterogeneity between the cohorts using the Higgins I^2 statistics.[29] Furthermore, we investigated a linear trend between subsequent exposure time windows in time and the risk of stroke. We used the time-window-specific effect estimates as a dependent variable and the time-window intervals as a categorical explanatory variable in a meta-regression model.

Potential effect modification by gender, smoking, hypertension (defined as $\geq$ 140 mm Hg systolic or $\geq$ 90 mm Hg diastolic BP, or intake of blood pressure-lowering medication, or in the SALT cohort, on self-reported data on prevalent hypertension) and diabetes were investigated by cohort-specific stratified analysis adjusted for the full set of covariates, and then combined into meta-analysis. When analysing effect modification by diabetes, calendar year was recoded into two periods instead of three periods in the model for SIXTY owing to the lack of model convergence when using 5-year intervals. Furthermore, the impact of family history of diabetes on the association between exposure and risk of stroke was investigated in SDPP and SIXTY by stratified analysis. Such data were lacking for the other two cohorts.

Sensitivity analyses were performed by restricting events to (i) only ischemic stroke, (ii) only non-fatal cases and (iii) only including incident cases after study enrolment. In a separate analysis, we explored the influence of contextual confounding by adding neighborhood mean income as an area-level socio-economic variable to the fully adjusted model

All analyses were performed using Stata version 11.0 (StataCorp LP, College Station, TX, USA).

RESULTS

A total of 22,587 subjects were recruited into the four cohorts (Table 1). After exclusion of the subjects recruited into more than one cohort or with missing data in any of the exposures or covariates, 20,070 subjects remained for the analysis. Overall, 6–13% of the subjects in each cohort were excluded owing to missing data.

At study entry, the mean age of all subjects was 60 years (range 35–104), and varying between cohorts (Table 2). Most participants had education up to secondary school or equivalent and were predominantly working or retired. Across cohorts, 15% to 25% of the participants were current smokers and about half reported regular alcohol consumption (daily/weekly). Diabetes prevalence was low, ranging from 1.6% to 8.6% across cohorts, while the proportion of hypertensive individuals was between 22.2% and 69.8% in the different cohorts. Most participants were either in the high or medium socio-economic category, and the average household income in the neighborhood was similar for three cohorts but higher for the cohort located in Stockholm city.

Air pollution linkage was successfully made for 99% of all individual addresses in Stockholm county ($n = 43{,}344$ addresses). The base-line exposure concentrations of NO_x and PM_{10} were similar for the two cohorts with recruitment in the whole of Stockholm County (SIXTY and SALT), whereas the SNAC-K cohort from Stockholm city had higher exposure levels. The SDPP cohort had the lowest levels and least variability for both PM_{10} and NO_x because of the recruitment of study participants from five suburban and semi-urban municipalities (Figure 1). Average concentrations varied somewhat across the exposure windows and followed the same between-cohort variability as for the study entry exposure data (Supplementary Table 1).

In general, NO_x concentrations were reduced during the observation period, whereas levels of PM_{10} were relatively constant. Modeled NO_x and PM_{10} were highly correlated for all cohorts, where SDPP, SIXTY and SALT had a high Pearson correlation ($r \sim 0.9$), whereas SNAC-K had a slightly lower correlation ($r = 0.75$).

A total of 868 subjects suffered a stroke during the 238,731 person-years at risk. Of the subjects, 775 (89%) were first ever cases of stroke after study entry, 755 (87%) were non-fatal and 737 (84%) were ischemic. The occurrence of stroke events were distributed over cohorts accordingly; 130 events in SDPP, 160 events in SIXTY, 314 events in SALT and 264 events in SNAC-K (Table 2).

The cohort-specific HR for total stroke per 20 $\mu g/m^3$ increment of road-traffic-related exposure to NO_x at enrolment address ranged between 0.84 and 1.78 (statistically significant only in the SIXTY cohort) when adjusting for all covariates (Figure 2, Supplementary Table 2). The combined HR was 1.16 (0.83–1.61). Similar risk estimates were seen per 10 $\mu g/m^3$ increase of traffic-related PM_{10}, where the cohort specific HR ranged between 0.59 and 2.21, also significant only in the SIXTY cohort. The combined analysis gave an overall HR of 1.14 (0.68–1.90). Moderate heterogeneity was suggested by the Higgin's I^2 statistic: 53.7% for NO_x and 66.9% for PM_{10}. There were no major or consistent differences in risk estimates between crude and adjusted models, indicating only limited confounding by the risk factors under study (Supplementary Table 2). Furthermore, the cohort specific and combined HRs were very similar using NO_2 and NO_X (Supplementary Figure 1).

In the time-window analysis for the fully adjusted model, no clear trend in the effect estimates could be detected after meta-analysis (Figure 3, Supplementary Table 3), but the confidence intervals were wide. Results were similar for NO_x and PM_{10} exposure, although a close to statistically significant excess risk was seen for NO_x exposure 6–10 years prior to the event. Moderate heterogeneity was also found in all separate time-window meta-analyses, ranging from $I^2 = 35.4\%$ to 67.0% for NO_x and 58.3% to 67.0% for PM_{10}. There was no significant trend between the effect estimates from the meta-analysed exposure time windows and the time interval they covered. The *P*-value for time-window category as an explanatory variable for the meta-analysed

Table 1. Number of individuals included in the analysis of four cohorts from Stockholm County.

Cohort[a]	*SDPP*	*SIXTY*	*SALT*	*SNAC-K*
Number recruited	7949	4232	7043	3363
Reasons for exclusion				
Participation in more than one cohort[a]	0	8	159	78
Missing exposure data for time window	111	47	168	17
Missing data on covariates	387	480	710	352
Total number in analysis (% of number recruited)	7451 (94%)	3697 (87%)	6006 (85%)	2916 (87%)

[a]Subjects are included in the first cohort into which they were selected.

Table 2. Characteristics of the study participants included in the analyses from the four cohorts in Stockholm County.

	SDPP[a]	*SIXTY*[b]	*SALT*[c]	*SNAC-K*[d]
N	7451	3697	6006	2916
Number of stroke events during follow-up	130	160	314	264
Years of enrolment	1992–1998	1997–1999	1998–2002	2001–2004
Age, years: median (minimum—maximum)	48 (35–56)	60 (59–61)	56 (42–97)	72 (60–104)
Male (%)	42.9	46.4	42.3	35.2
Education (%)				
Primary school or less	25.5	27.0	21.3	24.2
Up to secondary school or equivalent	45.5	44.3	42.8	42.5
University degree and more	29.0	28.7	35.9	33.3
Occupation status (%)				
Employed/self-employed	92.6	52.0	N/A[e]	26.5
Unemployed	7.4	9,9	N/A[e]	N/A[e]
Homemaker/housewife	N/A[e]	7.7	N/A[e]	N/A[e]
Retired	N/A[e]	30.3	N/A[e]	73.5
Smoking status (%)				
Current smoker	25.5	19.4	19.9	14.4
Former smoker	36.7	39.4	44.5	36.7
Never smoker	37.9	41.2	35.6	48.9
Number of cigarettes/day for current smokers mean ± SD	13.6 (7.4)	13.2 (7.2)	13.0 (7.5)	10.7 (8.0)
Alcohol consumption (%)				
Daily	7.4	4.4	N/A[e]	20.4
Weekly	37.0	20.4	N/A[e]	24.2
Seldom	51.6	45.1	N/A[e]	46.4
Never	4.1	30.1	N/A[e]	9.0
Diagnosed diabetes[f] *(%)*				
Yes (%)	1.6	3.9	4.1	8.6
No (%)	98.4	96.1	95.9	91.4
Diagnosed hypertension[g] *(%)*				
Yes	24.2	51.9	22.2	69.8
No	75.8	48.1	77.8	30.2
Socio-economic index (%)				
Low	28.3	22.3	29.3	18.2
Medium	26.8	55.2	52.5	30.2
High	44.9	22.6	18.2	51.7
Physical activity (%)				
Once a month or less / < 1 h/week	10.7	68.6	25.6	N/A[e]
About once a week / ~ 1 h/week	81.5	23.8	62.7	N/A[e]
3 times a week or more / > 2 h/week	7.8	7.6	11.7	N/A[e]
Average household income in neighborhood[h] (SEK) mean ± (SD)	288645 (50881)	300990 (85723)	305333 (81989)	351723 (27125)

[a]Stockholm Diabetes Prevention Program study. [b]60-year-old cohort study. [c]Screening Across the Lifespan Twin study. [d]Swedish National Study on Aging and Care in Kungsholmen. [e]Data not available for a sufficient number of individuals. [f]Information on diabetes was based on glucose tolerance test in SDPP and on questionnaire data in remaining cohorts. [g]Information on hypertension was defined by blood pressure measurements or intake of blood pressure lowering medication in three cohorts and through questionnaire in SALT. [h]Individual socio-economic status was based on current or last (if retired) profession and categorized into low (blue collar worker), medium (low and intermediate level white collar worker, and self-employed) and high (high-level white-collar worker).

time-window HRs was not significant ($P = 0.75$). There were rather high correlations in individual exposure between different time windows, that is, 0.7–1 for PM_{10} and 0.5–0.9 for NO_X, which contributes to the similar risk estimates. Only, about half of the study subjects moved during the follow-up period.

No strong effect modification of the association between NO_x or PM_{10} and stroke was observed by gender, smoking, hypertension or diabetes (Supplementary Figure 2). The HRs for NO_x and stroke for individuals with heredity for diabetes in SDPP and SIXTY were 2.07 (0.72–5.96) and 1.27 (0.38–4.29), respectively, compared with 1.36 (0.33–5.64) and 1.63 (1.10–2.40) for individuals without heredity for diabetes. Corresponding results for PM_{10} were 4.10 (0.98–17.11)) and 1.73 (0.39–7.71) for those with heredity for diabetes, respectively, compared with 0.8 (0.10–7.13) and 1.84 (1.10–3.09) for those without (data not shown). Restricting stroke events to non-fatal, ischemic or incident cases as well as adjusting the main model for mean income as area-level socio-economic indicator did not have any major impact on the risk estimates (Supplementary Figure 3). Further sensitivity analyses assessed the potential effect modification by age in the SALT and SNACK cohorts, which included a sizable fraction of older subjects. When individuals older than 75 or 85 years of age were excluded, there was no clear or consistent pattern of changes in the HRs (data not shown).

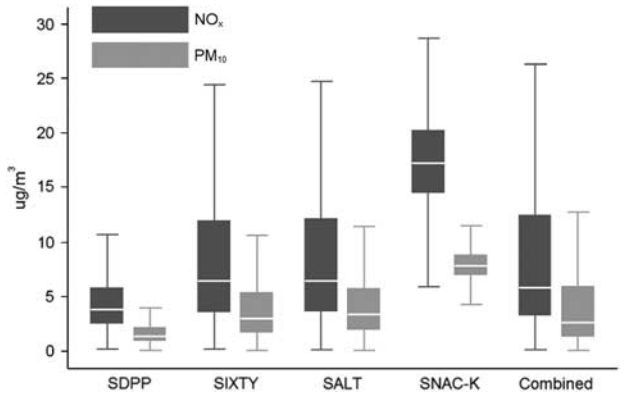


Figure 1. Modeled exposure levels of traffic-generated NO_x and PM_{10} ($\mu g/m^3$) at study entry addresses in four cohorts from Stockholm. Notes: Box layers describe the 75, 50, 25th percentile while outliers are not shown.

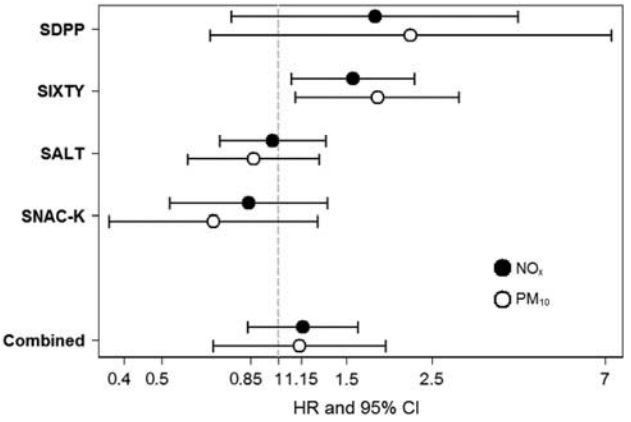


Figure 2. Exposure at recruitment from road traffic NO_x (per 20 $\mu g/m^3$) and PM_{10} (per 10 $\mu g/m^3$) and adjusted hazard ratio (HR) of stroke, in four cohorts in Stockholm County, separately and combined.

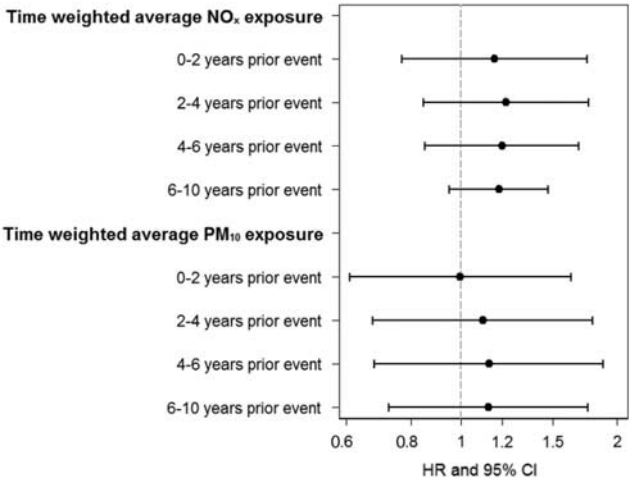


Figure 3. Adjusted hazard ratios (HR) of stroke, in relation to time-window exposure to NO_x (per 20 $\mu g/m^3$) and PM_{10} (per 10 $\mu g/m^3$) from road traffic in a meta-analysis of four cohorts from Stockholm County.

DISCUSSION

We found suggestive evidence of an association between air pollution from local road traffic, using NO_x and PM_{10} as indicators, and incidence of stroke in a region with comparatively low air pollution levels. No clear differences were indicated in effect estimates between various exposure time windows, but the power was limited in these analyses.

Some heterogeneity between cohorts in risk estimates for stroke related to long-term air pollution exposure was observed. Other studies have shown mixed results with borderline significant associations primarily for fatal stroke, with NO_2 (from dispersion modelling with high spatial detail) in Denmark,[5] and statistically significant associations for both ischemic and hemorrhagic stroke and NO_2 (exposure based on land-use regression models) in Japan.[30] The ESCAPE study on incidence of cerebrovascular events found suggestive evidence of an association for stroke and $PM_{2.5}$, PM $_{10}$ and Coarse PM but not for NO_2 or NO_x (exposure based on land-use regression models).[6] On the contrary, studies from England,[31] Oslo[3] and North America[4,10,32] did not see any elevated stroke risks associated with air pollution (mainly based on exposure assessment with less geographic detail). A study on women in USA, based on vicinity to urban background monitors, found larger effect estimates for $PM_{2.5}$ on cerebrovascular events (and death) for within-city exposure differences than for between-city differences, but no effects of NO_2.[33] These differences indicate that high spatial resolution is needed to describe the air pollution contrasts that may be associated with stroke risk.

One technical difference between our study and others' is that we have estimated the partial contribution to air pollution levels from road traffic only, whereas most other studies have estimated total levels. Within the Stockholm region, however, the local spatial differences in residential levels of both PM_{10} and NO_x are dominated by the emissions from road traffic.[34] Adding a regional background effect to all estimated values would not have changed our results, which were based on absolute rather than relative differences in exposure.

It is of particular interest to compare our results with those of the recently published study on cerebrovascular events from the ESCAPE project in which 20% of the stroke cases were in the four cohorts in the present study. Our exposure assessment was based on dispersion modeling, whereas ESCAPE used land use regression and our region constituted the lowest exposed area in ESCAPE, particularly for NO_x. We found suggestive evidence of associations for both NO_x and $PM_{10,}$ whereas no association was observed for NO_x in ESCAPE based on exposure at residential address at study enrolment. Furthermore, the magnitude of the effect was similar for PM_{10} in the two studies (14% and 11% per 10 $\mu g/m^3$). Notably, the cohort-specific effect estimates for long-term air pollution exposure and stroke in this study differed somewhat from the estimates presented for the same cohorts in the ESCAPE study (ESCAPE data retrieved through author correspondence). These variations could to some extent be explained by slightly diverse model adjustments, but a more probable explanation is the choice of exposure modeling technique. A major uncertainty with employment of both exposure assessment methodologies in most epidemiological studies is the failure to consider exposure contributions from occupational locations and commuting.[35] This generally would be expected to contribute to dilution of the associations.

Consistently with other European studies on long-term effects of air pollution on stroke, we did not find a difference between associating ischemic stroke and all stroke cases to PM_{10} and NO_x. Although short-term studies generally suggest a stronger association with ischemic stroke,[5,36–40] the literature on chronic effects of air pollution and types of stroke is sparse and results are mixed. A case–control study of ischemic stroke hospitalization in southern

Sweden and yearly mean NO_x averages prior to events showed no significant association.[41] In a Danish study using address-specific NO_2 concentrations weighted over 9.8 years, the strongest associations were found for non-specified and ischemic strokes whereas no association was found for hemorrhagic stroke.[5] A US study did not find a significant association between stroke type and the interquartile range (4 $\mu g/m^3$) change in average $PM_{2.5}$, $PM_{10-2.5}$ or PM_{10} exposure in the 12 months prior event in a cohort including only men.[10] Studies from Asia have linked air pollution to both ischemic and hemorrhagic stroke[30,42] where one long-term study found a significant association of yearly mean NO_2 levels from 1, 2 and 3 fiscal years prior a ischemic stroke and 2–3 years prior a hemorrhagic event.[30] In general, hemorrhagic stroke is less common than ischemic stroke, which leads to lower statistical precision and power in detecting risks for this type of stroke.

A recent review of epidemiological evidence on long-term exposure to air pollution and cardio-respiratory mortality found significant heterogeneity in $PM_{2.5}$ effect estimates across studies.[43] It was suggested that this was related to differences in particle composition, infiltration of particles indoors, population characteristics and methodological differences in exposure assessment and confounder control. In the ESCAPE study on cerebrovascular events based on 11 cohorts from 7 European countries,[6] heterogeneity was found for all exposure metrics but NO_x and the coarse PM fraction. Age was proposed as a major heterogeneity source but was suggested to correlate with other cohort characteristics.

On the other hand, studies within the ESCAPE project on acute coronary events[44] using the same cohorts, and on mortality adding 11 cohorts,[45] failed to detect such heterogeneity. In our study, we detected between-cohort heterogeneity in the effect estimates for long-term exposure to NO_X and PM_{10} on stroke incidence, even though the cohorts were based in only one region. The two cohorts in our study not showing associations differed from the other in certain aspects. One (SALT) lacked information on occupation status and alcohol consumption, available in all or most other cohorts, which probably led to poorer confounding control. The other (SNAC-K) was considerably older, with ages up to 105 years at recruitment, where less susceptible "survivors" may have been enriched

SNAC-K also differed considerably from the other cohorts in regard to prevalence of hypertension. On the other hand, the SDPP cohort (where associations were suggested) was selected so that diabetes heredity was more common, and those with such heredity appeared at higher risk of stroke associated with air pollution exposure. This cohort also had a longer observation period than the other because of earlier recruitment. However, taken together, we cannot find explanations that fully account for the observed heterogeneity.

A strength of our study consisted in the detailed assessment of air pollution for each subject, which took changes in residential address and calendar time into account. In addition, individual data on many cardiovascular risk factors were available along with information on potential contextual confounders. National health registries that are validated, for example regarding stroke,[46] were used to obtain the outcome data which minimized the risk of misclassification, although the quality of differentiation between ischemic and hemorrhagic stroke was lower during the earlier years of follow-up. Furthermore, the individual cohorts were rather small limiting the statistical power to detect associations, particularly in analyses of subgroups and interactions.

In conclusion, our findings indicated a possible association between local air pollution from road traffic and incidence of stroke in a combined analysis of four cohorts from Stockholm County. No clear differences in risk related to time windows of exposure were seen, but the interpretation was hampered by a limited statistical power in these analyses.

CONFLICT OF INTEREST

The authors declare that they have no competing interests.

ACKNOWLEDGMENTS

This project was funded by the Swedish Environmental Protection Agency, the Swedish Council for Working Life and Social Research, and the Swedish Heart-Lung Foundation. The SDPP cohort was additionally funded by the Stockholm County Council, the Swedish Research Council, the Swedish Diabetes Association and the Novo Nordisk Scandinavia. The SIXTY cohort was additionally funded by the Stockholm County Council and the Swedish Research Council (longitudinal research and 0593). The SALT cohort was additionally supported by NIH grant AG-08724. We acknowledge the Stockholm and Uppsala County Air Quality Management Association for enabling use of their emission inventory.

REFERENCES

1 Feigin VL, Forouzanfar MH, Krishnamurthi R, Mensah GA, Connor M, Bennett DA *et al.* Global and regional burden of stroke during 1990-2010: findings from the Global Burden of Disease Study 2010. *Lancet* 2014; **383**: 245–254.
2 Brook RD, Rajagopalan S, Pope CA 3rd, Brook JR, Bhatnagar A, Diez-Roux AV *et al.* Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the American Heart Association. *Circulation* 2010; **121**: 2331–2378.
3 Nafstad P, Haheim LL, Wisloff T, Gram F, Oftedal B, Holme I *et al.* Urban air pollution and mortality in a cohort of Norwegian men. *Environ Health Perspect* 2004; **112**: 610–615.
4 Pope CA 3rd, Burnett RT, Thurston GD, Thun MJ, Calle EE, Krewski D *et al.* Cardiovascular mortality and long-term exposure to particulate air pollution: epidemiological evidence of general pathophysiological pathways of disease. *Circulation* 2004; **109**: 71–77.
5 Andersen ZJ, Kristiansen LC, Andersen KK, Olsen TS, Hvidberg M, Jensen SS *et al.* Stroke and long-term exposure to outdoor air pollution from nitrogen dioxide: a cohort study. *Stroke* 2012; **43**: 320–325.
6 Stafoggia M, Cesaroni G, Peters A, Andersen ZJ, Badaloni C, Beelen R *et al.* Long-term exposure to ambient air pollution and incidence of cerebrovascular rvents: Results from 11 European cohorts within the ESCAPE Project. *Environ Health Perspect* 2014; **122**: 919–925.
7 Clancy L, Goodman P, Sinclair H, Dockery DW. Effect of air-pollution control on death rates in Dublin, Ireland: an intervention study. *Lancet* 2002; **360**: 1210–1214.
8 Laden F, Schwartz J, Speizer FE, Dockery DW. Reduction in fine particulate air pollution and mortality: Extended follow-up of the Harvard Six Cities study. *Am J Respir Crit Care Med* 2006; **173**: 667–672.
9 Zanobetti A, Schwartz J. Particulate air pollution, progression, and survival after myocardial infarction. *Environ Health Perspect* 2007; **115**: 769–775.
10 Puett RC, Hart JE, Suh H, Mittleman M, Laden F. Particulate matter exposures, mortality, and cardiovascular disease in the health professionals follow-up study. *Environ Health Perspect* 2011; **119**: 1130–1135.
11 Chen R, Zhang Y, Yang C, Zhao Z, Xu X, Kan H. Acute effect of ambient air pollution on stroke mortality in the China air pollution and health effects study. *Stroke* 2013; **44**: 954–960.
12 Eriksson AK, Ekbom A, Granath F, Hilding A, Efendic S, Ostenson CG. Psychological distress and risk of pre-diabetes and Type 2 diabetes in a prospective study of Swedish middle-aged men and women. *Diabet Med* 2008; **25**: 834–842.
13 Wandell PE, Wajngot A, de Faire U, Hellenius ML. Increased prevalence of diabetes among immigrants from non-European countries in 60-year-old men and women in Sweden. *Diabetes Metab* 2007; **33**: 30–36.
14 Lichtenstein P, Sullivan PF, Cnattingius S, Gatz M, Johansson S, Carlstrom E *et al.* The Swedish Twin Registry in the third millennium: an update. *Twin Res Hum Genet* 2006; **9**: 875–882.
15 Meinow B, Kareholt I, Lagergren M. According to need? Predicting the amount of municipal home help allocated to elderly recipients in an urban area of Sweden. *Health Soc Care Community* 2005; **13**: 366–377.
16 Bellander T, Berglind N, Gustavsson P, Jonson T, Nyberg F, Pershagen G *et al.* Using geographic information systems to assess individual historical exposure to air pollution from traffic and house heating in Stockholm. *Environ Health Perspect* 2001; **109**: 633–639.
17 Johansson C, Hadenius A, Johansson P-Å, Jonson T. SHAPE: The Stockholm study on health effects of air pollution and their economic consequences. Part 1: NO2 and particulate matter in Stockholm [Internet] 1999[cited 2015 Jan 23] 1999: 41. Available from http://slb.nu/slb/rapporter/pdf8/lvf1998_006.pdf.
18 Johansson C, Eneroth K. TESS–Traffic Emissions, Socioeconomic valuation and Socioeconomic measures, Part 1: Emissions and exposure of particles and NOx in

Greater Stockholm. Stockholm and Uppsala County Air Quality Management Association Report LVF [Internet]. 2007:2. [cited 2015 Jan 23] Available from http://slb.nu/slb/rapporter/pdf8/lvf2007_002.pdf.
19 Johansson C, Norman M, Gidhagen L. Spatial & temporal variations of PM10 and particle number concentrations in urban air. *Environ Monit Assess* 2007; **127**: 477–487.
20 Omstedt G, Bringfelt B, Johansson C. A model for vehicle-induced non-tailpipe emissions of particles along Swedish roads. *Atmospheric Environment* 2005; **39**: 6088–6097.
21 Sjödin Å, Ferm M, Björk A, Rahmberg M, Gudmundsson A, Swietlicki E *et al.* Wear particles from road traffic-a field, laboratory, and modelling study. Goteborg. IVL Swedish Environmental Research Institute Ltd.[Internet]: Sweden, 2010 [cited 2015 Jan 23] Available from http://www.ivl.se/webdav/files/B-rapporter/B1830.pdf.
22 Denby B, Sundvor I, Johansson C, Pirjola L, Ketzel M, Norman M *et al.* A coupled road dust and surface moisture model to predict non-exhaust road traffic induced particle emissions (NORTRIP). Part 1: Road dust loading and suspension modelling. *Atmospheric Environment* 2013; **77**: 283–300.
23 Gidhagen L, Johansson C, Langner J, Foltescu V. Urban scale modeling of particle number concentration in Stockholm. *Atmospheric Environment* 2005; **39**: 1711–1725.
24 Johansson C, Hadenius A, Johansson P-Å, Jonson T. SHAPE: The Stockholm study on health effects of air pollution and their economic consequences. Part 1: NO2 and particulate matter in Stockholm VAEGVERKET. PUBLIKATION [Internet]. 1999: 41[cited 2015 Jan 23] 1401-9612 Available from http://www.ivl.se/webdav/files/B-rapporter/B1830.pdf.
25 Eneroth K, Johansson C, Bellander T. Exposure comparison between measurements and calculations based on dispersion modelling (EXPOSE). Stockholm Environment and Health Protection Administration[Internet]. 2006. [cited 2015 Jan 23] Available from http://slb.nu/slb/rapporter/pdf8/lvf2006_012.pdf.
26 Johansson C, Andersson C, Bergström R, Krecl P. Exposure to particles due to local and non-local sources in Stockholm: estimates based on modelling and measurements 1997-2006. *ITM RAPPORT* 2008, Department of Applied Environmental Science, Stockholm University: Stockholm (Sweden). ITM report 175.
27 de Hoogh K, Korek M, Vienneau D, Keuken M, Kukkonen J, Nieuwenhuijsen MJ *et al.* Comparing land use regression and dispersion modelling to assess residential exposure to ambient air pollution for epidemiological studies. *Environ Int* 2014; **73**: 382–392.
28 DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986; **7**: 177–188.
29 Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002; **21**: 1539–1558.
30 Yorifuji T, Kashima S, Tsuda T, Ishikawa-Takata K, Ohta T, Tsuruta K *et al.* Long-term exposure to traffic-related air pollution and the risk of death from hemorrhagic stroke and lung cancer in Shizuoka, Japan. *Sci Total Environ* 2013; **443**: 397–402.
31 Atkinson RW, Carey IM, Kent AJ, van Staa TP, Anderson HR, Cook DG. Long-term exposure to outdoor air pollution and incidence of cardiovascular diseases. *Epidemiology* 2013; **24**: 44–53.
32 Chen H, Goldberg MS, Burnett RT, Jerrett M, Wheeler AJ, Villeneuve PJ. Long-term exposure to traffic-related air pollution and cardiovascular mortality. *Epidemiology* 2013; **24**: 35–43.
33 Miller KA, Siscovick DS, Sheppard L, Shepherd K, Sullivan JH, Anderson GL *et al.* Long-term exposure to air pollution and incidence of cardiovascular events in women. *N Engl J Med* 2007; **356**: 447–458.
34 Täppefur M. Air pollution in Stockholm and Uppsala Counties, and in Gävle and Sandviken municipalities. Emission data for year 2011. SLB-ANALYS. 2011; APRIL 2013.
35 Ozkaynak H, Baxter LK, Dionisio KL, Burke J. Air pollution exposure prediction approaches used in air pollution epidemiology studies. *J Expo Sci Environ Epidemiol* 2013; **23**: 566–572.
36 Vidale S, Bonanomi A, Guidotti M, Arnaboldi M, Sterzi R. Air pollution positively correlates with daily stroke admission and in hospital mortality: a study in the urban area of Como, Italy. *Neurol Sci* 2010; **31**: 179–182.
37 Wellenius GA, Schwartz J, Mittleman MA. Air pollution and hospital admissions for ischemic and hemorrhagic stroke among medicare beneficiaries. *Stroke* 2005; **36**: 2549–2553.
38 Chan CC, Chuang KJ, Chien LC, Chen WJ, Chang WT. Urban air pollution and emergency admissions for cerebrovascular diseases in Taipei, Taiwan. *Eur Heart J* 2006; **27**: 1238–1244.
39 Andersen ZJ, Olsen TS, Andersen KK, Loft S, Ketzel M, Raaschou-Nielsen O. Association between short-term exposure to ultrafine particles and hospital admissions for stroke in Copenhagen, Denmark. *Eur Heart J* 2010; **31**: 2034–2040.
40 Oudin A, Stromberg U, Jakobsson K, Stroh E, Bjork J. Estimation of short-term effects of air pollution on stroke hospital admissions in southern Sweden. *Neuroepidemiology* 2010; **34**: 131–142.
41 Oudin A, Stroh E, Stromberg U, Jakobsson K, Bjork J. Long-term exposure to air pollution and hospital admissions for ischemic stroke. A register-based case-control study using modelled NO(x) as exposure proxy. *BMC Public Health* 2009; **9**: 301.
42 Yamazaki S, Nitta H, Ono M, Green J, Fukuhara S. Intracerebral haemorrhage associated with hourly concentration of ambient particulate matter: case-crossover analysis. *Occup Environ Med* 2007; **64**: 17–24.
43 Hoek G, Krishnan RM, Beelen R, Peters A, Ostro B, Brunekreef B *et al.* Long-term air pollution exposure and cardio- respiratory mortality: a review. *Environ Health* 2013; **12**: 43.
44 Cesaroni G, Forastiere F, Stafoggia M, Andersen ZJ, Badaloni C, Beelen R *et al.* Long term exposure to ambient air pollution and incidence of acute coronary events: prospective cohort study and meta-analysis in 11 European cohorts from the ESCAPE Project. *BMJ* 2014; **348**: f7412.
45 Beelen R, Raaschou-Nielsen O, Stafoggia M, Andersen ZJ, Weinmayr G, Hoffmann B *et al.* Effects of long-term exposure to air pollution on natural-cause mortality: an analysis of 22 European cohorts within the multicentre ESCAPE project. *Lancet* 2014; **383**: 785–795.
46 Koster M, Asplund K, Johansson A, Stegmayr B. Refinement of Swedish administrative registers to monitor stroke events on the national level. *Neuroepidemiology* 2013; **40**: 240–246.

Supplementary Information accompanies the paper on the Journal of Exposure Science and Environmental Epidemiology website (http://www.nature.com/jes)

Long-term exposure to traffic-related air pollution and coronary events in four cohorts from Stockholm

Michal Korek[1], Tom Bellander[1, 2], Tomas Lind[2], Matteo Bottai[3], Kristina Eneroth PhD[10], Barbara Caracciolo[4], Ulf de Faire[1], Laura Fratiglioni[4, 5, 6], Agneta Hilding[7], Karin Leander[1] , Patrik KE Magnusson[8], Nancy L Pedersen[8], Claes-Göran Östenson[7], Göran Pershagen[1, 2], Johanna Penell[1, 9]

[1] Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden

[2] Centre for Occupational and Environmental Medicine, Stockholm County Council, Stockholm, Sweden

[3] Division of Biostatistics, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden

[4] Aging Research Center, Department of Neurobiology, Care Sciences and Society, , Karolinska Institutet & Stockholm, University, Stockholm, Sweden

[5] Stockholm Gerontology Research Center, Stockholm, Sweden;

[6] Division of Clinical Geriatrics, Karolinska University Hospital, Stockholm, Sweden

[7] Department of Molecular Medicine and Surgery, Karolinska Institutet, Karolinska University Hospital, Stockholm, Sweden

[8] Department of Medical Epidemiology and Biostatistcs, Karolinska Institutet, Stockholm Sweden

[9] Faculty of Health & Medical Sciences, School of Veterinary Medicine, Guildford Surrey, UK

[10] Environment and Health Administration, Stockholm City, Sweden

Corresponding author:

Michal Korek, MSc

Karolinska Institutet, Institute of Environmental Medicine,

Nobels väg 13 Box 210, SE-171 77 Stockholm, Sweden

E-mail: michal.korek@ki.se

Phone: +46-8-524 87457

Fax: +46-8-304571

Key words: air pollution, cohorts, exposure assessment, coronary events, time windows

ABSTRACT

Background

Few prospective studies have assessed the role of timing of exposure for the risk of developing coronary events (CE) following long-term exposure to ambient air pollution.

Methods

A total of 22,587 individuals aged 35-104 years were enrolled between 1992 and 2004 in four cohorts from Stockholm County and followed through 2011. Outdoor levels of NO_X and PM_{10} from local traffic were estimated annually for the residential addresses from 1987 to the end of follow-up, using dispersion models. Information on lifestyle and other environmental exposures was obtained from questionnaires and clinical tests, while national registries provided data on morbidity and mortality from CE. Cox proportional hazards regression analyses were performed for both pollution components at recruitment address and for different time windows in each cohort and combined in meta-analyses.

Results

A total of 913 coronary events were observed during the 237,723 person-years at risk. The meta-analyzed hazard risk ratios for CE related to traffic exposures at time of enrolment were 1.14 (0.78- 1.49) per 10 µg/m³ PM_{10} increase and 1.02 (0.82- 1.27) per 20 µg/m³ increase of NO_x concentration. No specific time window appeared to be of particular importance for the risk of CE. Sensitivity analysis did not reveal clear differences in risks between non-fatal CE, incident CE or myocardial infarction.

Conclusions

We observed suggestive evidence of elevated risk for coronary events from long-term exposure to comparatively low levels of air pollutants from road traffic, without any clear effect modification by timing of exposure.

INTRODUCTION

Epidemiological studies indicate that air pollution exposure is a risk factor for various cardiovascular outcomes and several pathophysiological pathways have been proposed.[1] Most studies of the association between long-term air pollution exposure and coronary events (CE) have focused on mortality and generally found increased risks.[2] However, coronary events are mostly non-fatal, which implies the need to also study incident coronary events. The few available reports on incident coronary events have mostly found no relation with NOx/NO_2 but more varying results for particulate matter exposure.[3-8] For instance, an intervention study in Ireland demonstrated decreased cardiovascular mortality within a year after drastically lowered levels of black smoke.[9] Similarly, exposure within a few years prior to the event has been associated with stronger effects compared to other time-periods for mortality from cardiovascular disease [10] and myocardial infarction.[11] Other studies on CE did not find time-windows of particular importance.[7, 12, 13] In the only study investigating time windows of exposure relevant to incident coronary events no effect modification was observed by timing of exposure.[14] Therefore, there is a need for further studies addressing the role of specific time-windows of exposure in relation to incident CE.

Previous studies on long-term air pollution exposure and CE have commonly included areas exceeding the WHO guidelines for particulate exposure, e.g. $20\mu g/m^3$ for annual mean exposure of particulate matter with an aerodynamic diameter of less than 10 micrometer (PM_{10}). However, the effect of air pollution on CE risk does not seem not to have a threshold level,[4] implying a need for risk assessment in areas with comparatively low air pollution levels. Furthermore, the effect of long-term air pollution exposure from specific sources on incident CE has not been investigated, since most studies have used exposure metrics capturing combinations of pollution sources. A further exploration of the effect of air

pollutants generated by road-traffic on incident CE could be of importance for prioritization of preventive measures.

The aim of this study was to investigate the relation between individual long-term exposure to air pollution from road traffic and incident coronary events in an area with comparatively low levels of exposure. We related the hazard of coronary events to subject specific exposure at study entry and in several time windows using time-space adjusted exposure data in a combined analysis of four cohorts from Stockholm County, Sweden.

MATERIAL AND METHODS

Study population

Individual data was collected on men and women who had been enrolled between 1992 and 2004 into four cohorts based in Stockholm County. The Stockholm Diabetes Preventive Program (SDPP) is a population-based prospective study which recruited 3128 men in 1992-1994 and 4821 women in 1996-1998 from five municipalities in Stockholm County. The study participants were 35-56 years old at recruitment and selected so that none had been previously diagnosed with diabetes and about half of the participants (53%) had a family history of diabetes (one first degree relative or two second degree relatives)[15]. The cohort study of 60 year olds (SIXTY) invited a random sample consisting of one third of all men and women who were living in Stockholm County and turned 60 years of age between August 1997 and March 1999, in total including 4232 subjects with a focus on investigating cardiovascular disease.[16] The Screening Across the Lifespan Twin Study (SALT) screened all twins born in Sweden before 1958 during 1998-2002 for the most common complex

diseases with a focus on cardiovascular diseases. In the present study, the SALT participants residing in Stockholm County at recruitment were included, resulting in 7043 subjects 42-100 years of age at recruitment.[17] The Swedish National study of Aging and Care in Kungsholmen (SNAC-K) included randomly sampled individuals >=60 years old between 2001 and 2004 from a central area in Stockholm City.[18] A total of 3363 subjects 60-104 years of age enrolled in SNACK. The individuals in the four cohorts were followed until the occurrence of a CE event, death, migration outside Stockholm County, or 31st December 2011, whichever came first.

At enrolment in the cohorts, individual data were collected from questionnaires on socio-demographic characteristics such as occupational status and education, and on life style factors including smoking status, levels of physical activity, alcohol consumption and antihypertensive treatment. For diabetes and hypertension, information was obtained either from questionnaires or clinical investigations. Socio-economic data on neighborhood level were retrieved from Statistics Sweden. The neighborhoods were small geographical units with an average population of 1000-2000 subjects considered to be homogenous with regard to socio-economic characteristics. Definitions of the covariates are provided in supplementary material table1, online supplements. The study was approved by the Ethical Review Board Stockholm, Sweden.

Outcome data

For all individuals, data on coronary events were retrieved from the National Hospital Discharge Registry and the National Cause of Death Registry, including data from February 1964 to December 2011. The following international classification of diseases (ICD) codes were used to define a coronary event: “Acute Myocardial Infarction” or “Other acute and sub-acute forms of ischemic heart disease” (ICD9: 410; 411; ICD10: I21, I23, 120.0, I24), in the

Hospital Discharge Registry and "Ischemic heart disease" (ICD9: 410-414; ICD10: I20-25) in the Cause of Death Registry. The event was classified as fatal if the person passed away within 28 days after the disease onset. Only coronary events after recruitment to the respective cohort were included in the main analyses. Events occurring before this time were used to define whether a coronary event during the study period was incident (i.e. the first-ever event) Hypertension was assessed for analyses of effect modification and defined as ≥ 140 mmHg systolic, or ≥ 90 mmHg diastolic blood pressure, or intake of blood pressure lowering medication or in the SALT cohort based on self-reported data on prevalent hypertension.

Exposure assessment

Individual long-term exposure to ambient air pollution from road traffic was estimated using NO_x as a marker of vehicle exhaust and particulate matter PM_{10} as a marker for road wear. Changes of residential address and year-specific and source-specific mean exposure levels were taken into account using a methodology described in detail elsewhere [19, 20]. In brief, residential histories retrieved from the Swedish tax authorities were available for all cohort participants from 1991 until 2010, including the date from which the person resided at each particular address. When transferring this system into electronic format in 1991, the address entered for each individual may have had a start date earlier than 1991 if this was still the valid residential address. Migration from Sweden or Stockholm County was also recorded by the Swedish tax authorities. Ninety percent of the residential addresses within Stockholm County were geocoded by matching against the databases of the Swedish Mapping Cadastral and Land Registration Authority, and additionally 9 % were coded manually. One percent of the addresses could not be geocoded, mainly due to insufficient address detail.

Outdoor concentrations of NO_x and PM_{10} from local road traffic emissions were calculated for all geocoded addresses for every year from 1987 until end of follow-up, using a Gaussian air

quality dispersion model and a wind model, both part of the Airviro Air Quality Management System (SMHI, Norrköping, Sweden; http://airviro.smhi.se). The emission inventory of the Stockholm and Uppsala County Air Quality Management Association provided input data to the model calculations. The inventory is updated yearly since 1993 and contains detailed emission information from several pollutant sources, including road traffic.[21] Tail pipe emissions of NOx and PM_{10} were described by emission factors for various vehicle and road types according to the EVA model of the Swedish Transport Administration. The EVA model includes scenarios for the composition of various types of vehicles and fuels, such as the share of diesel cars, as well as the composition of the vehicle fleet in terms of European emission standards (Euro classification) for different years. Non-exhaust PM, including road, break and tyre wear particles, were also included in the inventory.[22]

The spatial resolution of the dispersion calculations was 25 x 25 meters for central Stockholm and in the urban parts of the municipalities of Solna, Järfälla and Södertälje, 100 x 100 meters in other urban areas and 500 x 500 meters in rural areas. The model calculation height was 2 meter above ground level in open country and 2 meter above roof height in urban areas, handling buildings by using a roughness parameter .[23] The Gaussian model thereby underestimates the NO_x and PM_{10} concentrations in narrow street canyons with poor dispersion conditions Therefore, additional adjustment of the modelled concentrations was made to all addresses with multistory houses on both sides within 30 m of the most polluted street segments in the inner city of Stockholm, using the AirViro street canyon model (http://airviro.smhi.se). The calculated street canyon contribution corresponds to air pollution concentrations at half the building height.

The AirViro Gaussian model has been validated in several studies,[21, 24] most recently within the European ESCAPE study.[25] Modelled estimates of NO_2 and PM_{10} were compared with measured levels 39 respectively 19 monitoring sites within Stockholm County, providing Spearman rank correlations of 0.76 for NO_2 and 0.58 for PM_{10}. The Airviro Air Quality Management System was used to create emission data bases every five years from 1990 for NO_x and for the year 2004 for $PM_{10.}$ The residential NOx and PM_{10} concentrations for years for which no emission data-base had been constructed were estimated by re-scaling based on annual measurements of traffic flow across the urban area boundary of Stockholm. Time-space adjusted annual exposure estimates were used to construct both study entry exposure and time-weighted exposure concentrations in specific time-windows.

Statistical analysis

The analysis was performed for each exposure separately and in two steps. First, cohort-specific Cox proportional hazards regression models were fitted to estimate hazard ratios of coronary events associated with long-term ambient air-pollution exposure. In a second step cohort-specific effect estimates were combined in a meta-analysis. Person-time at risk was calculated from recruitment until CE diagnosis or death, with censoring at emigration or end of follow-up, whichever came first. We used age as the underlying time scale in all models. The data was divided into 6 month risk periods, allowing use of NO_x and PM_{10} as time-varying covariates and adjustment for calendar year in 5 year periods. First, subject specific residential exposure to NO_x and PM_{10} was related to CE at the time of study entry. Secondly, the exposure was calculated for each subject for each subject-specific 6-month risk period during the follow-up. From these data we created exposure time-windows as the mean averages of 0-2, 2-4, 4-6, and 6-10 year intervals prior to the end date of every 6-month risk

period during the follow-up. Each specific time-window was then evaluated in a separate model.

Adjustment models were defined a priori and covariates were chosen based on the literature and available data from the cohorts. The fully adjusted models had a common set of individual-level covariates including gender, education level, smoking status, smoking intensity among current smokers and socioeconomic index. This index was based on current or last (if retired) profession and categorized into low (blue collar worker), medium (low and intermediate level white collar worker, and self-employed) and high (high-level white-collar worker). In addition, other potential confounders were included if data was available for at least two cohorts with at least 80% non-missing observations per cohort (for SALT alcohol consumption and occupational status were not available and for SNAC-K physical activity). Diagnostic tools were used to check the proportional-hazard assumption for all categorical covariates. If any variable in the individual cohort models violated this assumption, effect estimates were compared with a stratified Cox analysis for that cohort and covariate.[12] The cohort-specific results were combined using random effects meta-analysis.[26] Presence of heterogeneity between the cohorts was addressed using the Higgins I^2 statistic.[27] Furthermore, we investigated a linear trend between subsequent exposure time windows and the risk of coronary events by using the time-window-specific effect estimates as a dependent variable and the time-window intervals as a categorical explanatory variable in a meta-regression model.

Potential effect modification by gender, smoking, diabetes, hypertension or having only one residential address during follow-up was investigated by cohort specific stratified analysis adjusted for the full set of covariates, and then combined into a meta-analysis. When

analyzing effect modification by diabetes, calendar year was recoded into two periods instead of three periods in the model for SIXTY due to lack of model convergence. In addition, effect modification by family history of diabetes was analyzed for SDPP and SIXTY (the two cohorts that had available data on diabetes heredity) by stratified analysis.

Sensitivity analyses were performed by restricting coronary events to only non-fatal cases, incident cases after study enrolment or myocardial infarction. In a separate analysis, we explored the influence of contextual confounding by adding neighborhood mean income as an area-level socio-economic variable to the fully adjusted model. All analyses were performed using Stata version 11.0 (StataCorp LP, College Station, Texas, USA).

RESULTS

A total of 22,587 subjects were recruited into the four cohorts with some overlap of recruitment into multiple cohorts. After exclusion of subjects from cohorts other than their first cohort entry, or with missing data in any of the exposures or covariates, 20,068 subjects remained under study, 7450 (94%) in SDPP, 3697 (87%) in SIXTY, 6004 (85%) in SALT and 2917(88%) in SNAC-K. The four cohorts are described in supplementary table S1. At study entry, the age of the participants was highly variable ranging from 35 to 104 years with a median age of 59 years. Most participants had education up to secondary school or equivalent and were either in the high or medium socio-economic category. Across cohorts 14 to 25% of the participants were current smokers and about 25 to 45 % reported regular alcohol consumption (daily/weekly). Diabetes prevalence was low, ranging from 1.5 to 8.7%, while the proportion of hypertensive individuals was between 22.2 and 51.7% in the different cohorts.

Air pollution was successfully linked to 99% of all individual addresses in Stockholm county (n=43,344 addresses, on average 2.2 per individual). The median exposure concentrations of NO_x and PM_{10} levels were stable over categories of covariates within cohorts but differed between cohorts (Table 1).

A similar situation held true for NO_x (data not shown). The exposure concentrations were most similar for the two cohorts with recruitment in the whole of Stockholm County (SIXTY and SALT), while the SNAC-K cohort from Stockholm city had comparatively higher exposure levels. The SDPP cohort had the lowest levels and least variability for both PM_{10} and NO_x due to the recruitment of study participants from five suburban and semi-urban municipalities. For the time windows, average concentrations followed the same between-cohort variability as for the study entry exposure data (data not shown). In general, NO_x levels became lower during the observation period while the levels of PM_{10} contributions were relatively constant. The Pearson's correlation between modeled NO_x and PM_{10} was high ($r \sim 0.9$) in SDPP, SIXTY and SALT, and lower ($r = 0.75$) in SNAC-K. Overall, about half of the study subjects changed address at least once during follow-up.

There were 913 coronary events during the 237,723 person-years of observation, 786 (86%) were first ever cases of CE after study entry, 831 (91%) were nonfatal and 638 (70%) were classified as myocardial infarction. The number of CE events was 206 each in SDPP, SIXTY and SNAC-K, and 295 in SALT (Table 2).

The cohort-specific hazard ratio (HR) for total CE per 20 $\mu g/m^3$ increase of road-traffic related exposure to NO_x at the enrolment address ranged between 0.72 and 1.21 when adjusting for all covariates (Figure 1, supplementary table S2). The corresponding meta-

analyzed HR was 1.02 (0.82- 1.27). For PM_{10} the cohort specific HR per 10 µg/m³ increase ranged between 0.97 and 1.49 with a combined HR of 1.14 (0.78- 1.49). No heterogeneity was indicated for either exposure. In general, the confounding by risk factors was limited in all cohorts as indicated by similar hazard ratios in the crude and fully adjusted models (supplementary table S2).

In the time-window analysis for the fully adjusted model, meta-analyzed effect estimates did not indicate exposure periods of particular importance. (Figure 2, supplementary table S3), but the confidence intervals were wide. Results were similar for NO_x and PM_{10} exposure. Furthermore, we did not find a significant trend in the relation between hazard ratios from the meta-analyzed exposure time-windows. There was no clear effect modification due to hypertension, gender, diabetes status, and smoking status or if individuals changed home address during follow-up (Figure 3). Restricting CE to non-fatal cases, incident cases or myocardial infarction cases only, or adjusting the main model for mean income as an area-level socioeconomic indicator, did not have a major impact on hazard ratios (supplementary figure S1).

DISCUSSION

In this study, we found suggestive evidence of association between long-term exposure to traffic-related air pollution and the incidence of coronary events, particularly for PM_{10}. Our study region had comparatively low air pollution concentrations, with PM_{10} levels well below the WHO guideline value for yearly mean exposure (supplementary figure S2). We found no clear effect modification by sex, diabetes, blood pressure or change of address, or when restricting analyses to incident CE, non-fatal CE or myocardial infarction. Furthermore, the time-window analyses did not indicate exposure periods of particular importance.

The hazard ratios for incident coronary events in this study, 1.02 (0.82- 1.27) per 20 µg/m³ NO_x, and 1.14 (0.78- 1.49) per 10 µg/m³ PM_{10} are similar to the respective point estimates in a European-wide multi cohort study (ESCAPE) 1.01(0.98-1.05) per 20 µg/m^3 of NO_X and 1.12 (1.01 – 1.25) per 10 µg/m^3 of PM_{10}, although air-pollution exposure was estimated using a different technique often referred to as "land use regression".[4] The larger sample size in the ESCAPE study is reflected by the narrower confidence intervals compared with our study which contributed with 15% of the CE events in the ESCAPE study. Our findings are consistent with those of other studies relating incident CE to NOx or NO_2, reporting no association [4-6], although some studies found increased risks.[3, 5, 12] For PM_{10}, the hazard risk ratio in our study appeared larger compared with reported risk ratios in three studies from the US where the reported excess risk related to PM_{10} or coarse PM for any kind of CE was 4 % per 7-µg/m^3 or lower.[6, 7, 14] Partly, the difference in the effect estimates may be due to differences in particle composition in the US and Europe where Europe has more diesel emissions. In a study from the UK, no association was found for PM_{10} and myocardial infarction (MI), HR: 1.01(0.98–1.05) although an excess risk of 1.09 (1.05–1.14) per 3µg/m^3 increase in PM_{10} was seen for heart failure.[5] A higher HR was also suggested in a Greek cohort study where the risk ratio for IHD related to PM_{10} was 1.41 (0.91–2.17) per10µg/m^3 .[3] The PM_{10} levels in the UK and in the Greek studies were considerably higher than in our study, which may have contributed to the higher risk estimates.

Long-term air pollution exposure has more often been related to CE mortality than to incident or non-fatal CE, both for NO_2/NO_x and particulate exposure.[6-8, 12, 28-35] An earlier case-control study from Stockholm reported associations for both PM_{10} and NO_2 and fatal MI, especially for out of hospital deaths, but not for non-fatal MI.[36] The same was reported in a case-control study on NO_2 and incident MI in Rome.[37] The evidence suggests that air

pollution exposure affects etiological pathways leading to mortality to a greater extent than those contributing to development of non-fatal events.

We did not find a clear difference in the effect estimates between CE related to traffic derived exposure at the enrollment address versus in different time-windows, although the statistical power was limited for detecting heterogeneity. Still, the result is consistent with most of the few studies investigating the role of timing of exposure on CE mortality,[12, 13] and incidence.[7, 14] A US study based on MI survivors found stronger associations for both MI and congestive heart failure related to PM_{10} exposure during 1-2 years prior to the event but not for exposure during the same year or three years earlier.[11] Overall, the evidence is not clear regarding a role of specific time windows of exposure to air pollution for development of cardiovascular disease. Very large populations are necessary to address these questions, particularly in view of the substantial correlations in exposure of individuals between different time windows.

Our study had several strengths. Subject specific exposure was assessed with high spatial resolution and the exposure estimates considered all addresses during the observation period i.e. from prior study entry until the end of follow-up. We included detailed information on several individual cardiovascular risk factors and investigated the effect of potential effect modifiers and contextual confounders. Information on outcome was retrieved from validated National health registries minimizing the risk of bias. On the other hand, our study material was rather small, which resulted in a limited statistical power, particularly in sub-group analyses.

To conclude, we found suggestive evidence of association between local traffic derived air pollution and CE in a combined analysis of four cohorts from Stockholm County. There were no clear differences in risk for types of CE or for various time-windows of exposure but the statistical power in the subgroup analysis was limited.

Conflicts of Interest and Source of Funding

This project was funded by, the Swedish Environmental Protection Agency, the Swedish Council for Working Life and Social Research and the Swedish Heart-Lung Foundation. Additional sources of funding included the Swedish Research Council, the Swedish Diabetes Association, the Novo Nordisk Scandinavia (SDPP cohort) and the NIH grant AG-08724 (SALT cohort). The authors declare that they have no competing interests.

Running title: **Traffic-related air pollution exposure and coronary events**

References

1. Brook RD, Rajagopalan S, Pope CA, 3rd, et al. Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the American Heart Association. Circulation 2010;121:2331-78 doi: 10.1161/CIR.0b013e3181dbece1published Online First: Epub Date]|.
2. Hoek G, Krishnan RM, Beelen R, et al. Long-term air pollution exposure and cardio- respiratory mortality: a review. Environ Health 2013;12:43 doi: 1476-069X-12-43 [pii]

10.1186/1476-069X-12-43published Online First: Epub Date]|.
3. Katsoulis M, Dimakopoulou K, Pedeli X, et al. Long-term exposure to traffic-related air pollution and cardiovascular health in a Greek cohort study. Sci Total Environ 2014;490:934-40 doi: 10.1016/j.scitotenv.2014.05.058published Online First: Epub Date]|.
4. Cesaroni G, Forastiere F, Stafoggia M, et al. Long term exposure to ambient air pollution and incidence of acute coronary events: prospective cohort study and meta-analysis in 11 European cohorts from the ESCAPE Project. BMJ 2014;348:f7412 doi: 10.1136/bmj.f7412

bmj.f7412 [pii]published Online First: Epub Date]|.
5. Atkinson RW, Carey IM, Kent AJ, van Staa TP, Anderson HR, Cook DG. Long-term exposure to outdoor air pollution and incidence of cardiovascular diseases. Epidemiology 2013;24:44-53 doi: 10.1097/EDE.0b013e318276ccb8published Online First: Epub Date]|.
6. Lipsett MJ, Ostro BD, Reynolds P, et al. Long-term exposure to air pollution and cardiorespiratory disease in the California teachers study cohort. Am J Respir Crit Care Med 2011;184:828-35 doi: 10.1164/rccm.201012-2082OC

201012-2082OC [pii]published Online First: Epub Date]|.
7. Puett RC, Hart JE, Suh H, Mittleman M, Laden F. Particulate matter exposures, mortality, and cardiovascular disease in the health professionals follow-up study. Environ Health Perspect 2011;119:1130-5 doi: 10.1289/ehp.1002921published Online First: Epub Date]|.
8. Miller KA, Siscovick DS, Sheppard L, et al. Long-term exposure to air pollution and incidence of cardiovascular events in women. N Engl J Med 2007;356:447-58 doi: 356/5/447 [pii]

10.1056/NEJMoa054409published Online First: Epub Date]|.
9. Clancy L, Goodman P, Sinclair H, Dockery DW. Effect of air-pollution control on death rates in Dublin, Ireland: an intervention study. Lancet 2002;360:1210-4 doi: 10.1016/S0140-6736(02)11281-5published Online First: Epub Date]|.
10. Laden F, Schwartz J, Speizer FE, Dockery DW. Reduction in fine particulate air pollution and mortality: Extended follow-up of the Harvard Six Cities study. Am J Respir Crit Care Med 2006;173:667-72 doi: 10.1164/rccm.200503-443OCpublished Online First: Epub Date]|.
11. Zanobetti A, Schwartz J. Particulate air pollution, progression, and survival after myocardial infarction. Environ Health Perspect 2007;115:769-75 doi: 10.1289/ehp.9201published Online First: Epub Date]|.
12. Nafstad P, Haheim LL, Wisloff T, et al. Urban air pollution and mortality in a cohort of Norwegian men. Environ Health Perspect 2004;112:610-5
13. Chen H, Goldberg MS, Burnett RT, Jerrett M, Wheeler AJ, Villeneuve PJ. Long-term exposure to traffic-related air pollution and cardiovascular mortality. Epidemiology 2013;24:35-43 doi: 10.1097/EDE.0b013e318276c005published Online First: Epub Date]|.
14. Puett RC, Hart JE, Yanosky JD, et al. Chronic fine and coarse particulate exposure, mortality, and coronary heart disease in the Nurses' Health Study. Environ Health Perspect 2009;117:1697-701 doi: 10.1289/ehp.0900572published Online First: Epub Date]|.
15. Eriksson AK, Ekbom A, Granath F, Hilding A, Efendic S, Ostenson CG. Psychological distress and risk of pre-diabetes and Type 2 diabetes in a prospective study of Swedish middle-aged men

and women. Diabet Med 2008;25:834-42 doi: 10.1111/j.1464-5491.2008.02463.xpublished Online First: Epub Date]|.
16. Wandell PE, Wajngot A, de Faire U, Hellenius ML. Increased prevalence of diabetes among immigrants from non-European countries in 60-year-old men and women in Sweden. Diabetes Metab 2007;33:30-6 doi: 10.1016/j.diabet.2006.07.001published Online First: Epub Date]|.
17. Lichtenstein P, Sullivan PF, Cnattingius S, et al. The Swedish Twin Registry in the third millennium: an update. Twin research and human genetics : the official journal of the International Society for Twin Studies 2006;9:875-82 doi: 10.1375/183242706779462444published Online First: Epub Date]|.
18. Angleman SB, Santoni G, Von Strauss E, Fratiglioni L. Temporal Trends of Functional Dependence and Survival Among Older Adults From 1991 to 2010 in Sweden: Toward a Healthier Aging. The Journals of Gerontology Series A: Biological Sciences and Medical Sciences 2014:glu206
19. Bellander T, Berglind N, Gustavsson P, et al. Using geographic information systems to assess individual historical exposure to air pollution from traffic and house heating in Stockholm. Environ Health Perspect 2001;109:633-9 doi: sc271_5_1835 [pii]published Online First: Epub Date]|.
20. Gruzieva O, Bellander T, Eneroth K, et al. Traffic-related air pollution and development of allergic sensitization in children during the first 8 years of life. J Allergy Clin Immunol 2012;129:240-6 doi: 10.1016/j.jaci.2011.11.001

S0091-6749(11)01739-8 [pii]published Online First: Epub Date]|.
21. Johansson C, Hadenius A, Johansson P-Å, Jonson T. SHAPE: The Stockholm study on health effects of air pollution and their economic consequences. Part 1: NO2 and particulate matter in Stockholm VAEGVERKET. PUBLIKATION.1999: 41Available at: http://www.ivl.se/webdav/files/B-rapporter/B1830.pdf 1999.
22. Omstedt G, Bringfelt B, Johansson C. A model for vehicle-induced non-tailpipe emissions of particles along Swedish roads. Atmospheric Environment 2005;39:6088-97
23. Gidhagen L, Johansson C, Langner J, Foltescu V. Urban scale modeling of particle number concentration in Stockholm. Atmospheric Environment 2005;39:1711-25
24. Eneroth K, Johansson C, Bellander T. Exposure comparison between measurements and calculations based on dispersion modelling (EXPOSE). Stockholm Environment and Health Protection Administration 2006
25. de Hoogh K, Korek M, Vienneau D, et al. Comparing land use regression and dispersion modelling to assess residential exposure to ambient air pollution for epidemiological studies. Environ Int 2014;73:382-92 doi: 10.1016/j.envint.2014.08.011published Online First: Epub Date]|.
26. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7:177-88 doi: 0197-2456(86)90046-2 [pii]published Online First: Epub Date]|.
27. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Statistics in medicine 2002;21:1539-58 doi: 10.1002/sim.1186published Online First: Epub Date]|.
28. Beelen R, Hoek G, Houthuijs D, et al. The joint association of air pollution and noise from road traffic with cardiovascular mortality in a cohort study. Occup Environ Med 2009;66:243-50 doi: 10.1136/oem.2008.042358

oem.2008.042358 [pii]published Online First: Epub Date]|.
29. Raaschou-Nielsen O, Andersen ZJ, Jensen SS, et al. Traffic air pollution and mortality from cardiovascular disease and all causes: a Danish cohort study. Environ Health 2012;11:60 doi: 10.1186/1476-069X-11-60published Online First: Epub Date]|.
30. Pope CA, 3rd, Burnett RT, Thurston GD, et al. Cardiovascular mortality and long-term exposure to particulate air pollution: epidemiological evidence of general pathophysiological pathways of disease. Circulation 2004;109:71-7 doi: 10.1161/01.CIR.0000108927.80044.7F

01.CIR.0000108927.80044.7F [pii]published Online First: Epub Date]|.
31. Yorifuji T, Kashima S, Tsuda T, et al. Long-term exposure to traffic-related air pollution and mortality in Shizuoka, Japan. Occup Environ Med 2010;67:111-7 doi: 10.1136/oem.2008.045542published Online First: Epub Date]|.
32. Puett RC, Schwartz J, Hart JE, et al. Chronic particulate exposure, mortality, and coronary heart disease in the nurses' health study. Am J Epidemiol 2008;168:1161-8 doi: 10.1093/aje/kwn232

kwn232 [pii]published Online First: Epub Date]|.
33. Crouse DL, Peters PA, van Donkelaar A, et al. Risk of nonaccidental and cardiovascular mortality in relation to long-term exposure to low concentrations of fine particulate matter: a Canadian national-level cohort study. Environ Health Perspect 2012;120:708-14 doi: 10.1289/ehp.1104049published Online First: Epub Date]|.
34. Beelen R, Raaschou-Nielsen O, Stafoggia M, et al. Effects of long-term exposure to air pollution on natural-cause mortality: an analysis of 22 European cohorts within the multicentre ESCAPE project. Lancet 2014;383:785-95 doi: 10.1016/S0140-6736(13)62158-3

S0140-6736(13)62158-3 [pii]published Online First: Epub Date]|.
35. Cesaroni G, Badaloni C, Gariazzo C, et al. Long-term exposure to urban air pollution and mortality in a cohort of more than a million adults in Rome. Environ Health Perspect 2013;121:324-31 doi: 10.1289/ehp.1205862published Online First: Epub Date]|.
36. Rosenlund M, Berglind N, Pershagen G, Hallqvist J, Jonson T, Bellander T. Long-term exposure to urban air pollution and myocardial infarction. Epidemiology 2006;17:383-90 doi: 10.1097/01.ede.0000219722.25569.0fpublished Online First: Epub Date]|.
37. Rosenlund M, Picciotto S, Forastiere F, Stafoggia M, Perucci CA. Traffic-related air pollution in relation to incidence and prognosis of coronary heart disease. Epidemiology 2008;19:121-8 doi: 10.1097/EDE.0b013e31815c1921

00001648-200801000-00019 [pii]published Online First: Epub Date]|.

TABLE 1. Median exposure levels of PM_{10} in different categories of covariates in four cohorts in Stockholm County.

	SDPP[a]	SIXTY[b]	SALT[c]	SNAC-K[d]
N	7450	3697	6004	2917
Years of enrolment	1992-1998	1997-1999	1998-2002	2001-2004
Age				
<65	1.3	3.0	3.2	7.8
>=65	no observations	no observations	3.8	7.9
Gender				
Male	1.2	2.9	3.5	7.7
Female	1.5	3.1	3.2	7.9
Education (%)				
Primary school or less	1.3	2.9	3.4	8.0
Up to secondary school or equivalent	1.3	3.0	3.3	7.8
University degree and more	1.3	3.3	3.5	7.7
Occupation status (%)				
Employed/self-employed	1.3	3.0	N/A[h]	7.9
Unemployed	1.3	3.1	N/A[h]	N/A[h]
Homemaker/housewife	N/A[h]	3.1	N/A[h]	N/A[h]
Retired	N/A[h]	3.2	N/A[h]	7.8
Smoking status (%)				
Current smoker	1.3	3.4	3.5	8.0
Former smoker	1.3	3.0	3.2	7.7
Never smoker	1.3	3.0	3.4	7.9
Alcohol consumption (%)				
Daily	1.5	3.1	N/A[h]	7.6
Weekly	1.4	3.1	N/A[h]	7.8
Seldom	1.3	2.9	N/A[h]	7.9
Never	1.2	3.2	N/A[h]	8.0
Diagnosed diabetes[e] (%)				
Yes	1.3	3.2	3.6	7.9
No	1.3	3.0	3.4	7.8
Diagnosed hypertension[f] (%)				
Yes	1.4	2.9	3.4	7.9
No	1.3	3.2	3.4	7.7

Socioeconomic index (%)[g]				
Low	1.3	2.8	3.3	8.0
Medium	1.4	3.1	3.4	7.9
High	1.3	3.3	3.4	7.8
Physical activity (%)				
Once a month or less /<1h/week	1.4	3.1	3.6	N/A[h]
About once a week /~1h/week	1.3	2.9	3.3	N/A[h]
3 times a week or more />2h/week	1.3	2.8	3.4	N/A[h]
Average household income in neighborhood (SEK)				
< p50[i]	1.3	3.0	3.3	8.5
>=p50[i]	1.4	3.0	3.4	7.8

[a]Stockholm Diabetes Prevention Program study

[b]60-year-old cohort study

[c] Screening Across the Lifespan Twin study

[d] Swedish National Study on Aging and Care in Kungsholmen

[e] Information on diabetes was based on glucose tolerance test in SDPP and on questionnaire data in remaining cohorts.

[f] Information on hypertension was defined by blood pressure measurements or intake of blood pressure lowering medication in three cohorts and through questionnaire in SALT.

[g] Individual socioeconomic status was based on current or last (if retired) profession and categorized into low (blue collar worker), medium (low and intermediate level white collar worker, and self-employed) and high (high-level white-collar worker).

[h] Data not available for a sufficient number of individuals.

[i] p = percentile

TABLE 2. Type of specific coronary events in the analysis of four cohorts in Stockholm County

	SDPP		SIXTY		SALT		SNACK	
	n	% [a]	n	% [a]	n	% [a]	n	% [a]
Total amount of coronary events after recruitment	206	100	206	100	295	100	206	100
Non-fatal coronary events after recruitment	197	96	196	95	265	90	173	84
Incident[b] coronary events after recruitment	195	95	180	87	246	83	165	80
Myocardial infarction[c]	152	74	132	64	199	67	155	75

[a] Percentage of CE events relative to the total amount of CE events.

[b] Excluding individuals with CE events prior to the respective follow-up.

[c] First events during follow-up not restricted to incident cases only .

The average follow-up time was 12 years.

FIGURE 1. Exposure at recruitment from road traffic NO_x (per 20 µg/m^3) and PM_{10} (per10 µg/m^3) and adjusted hazard ratio (HR) of coronary events, in four cohorts in Stockholm County, separately and combined

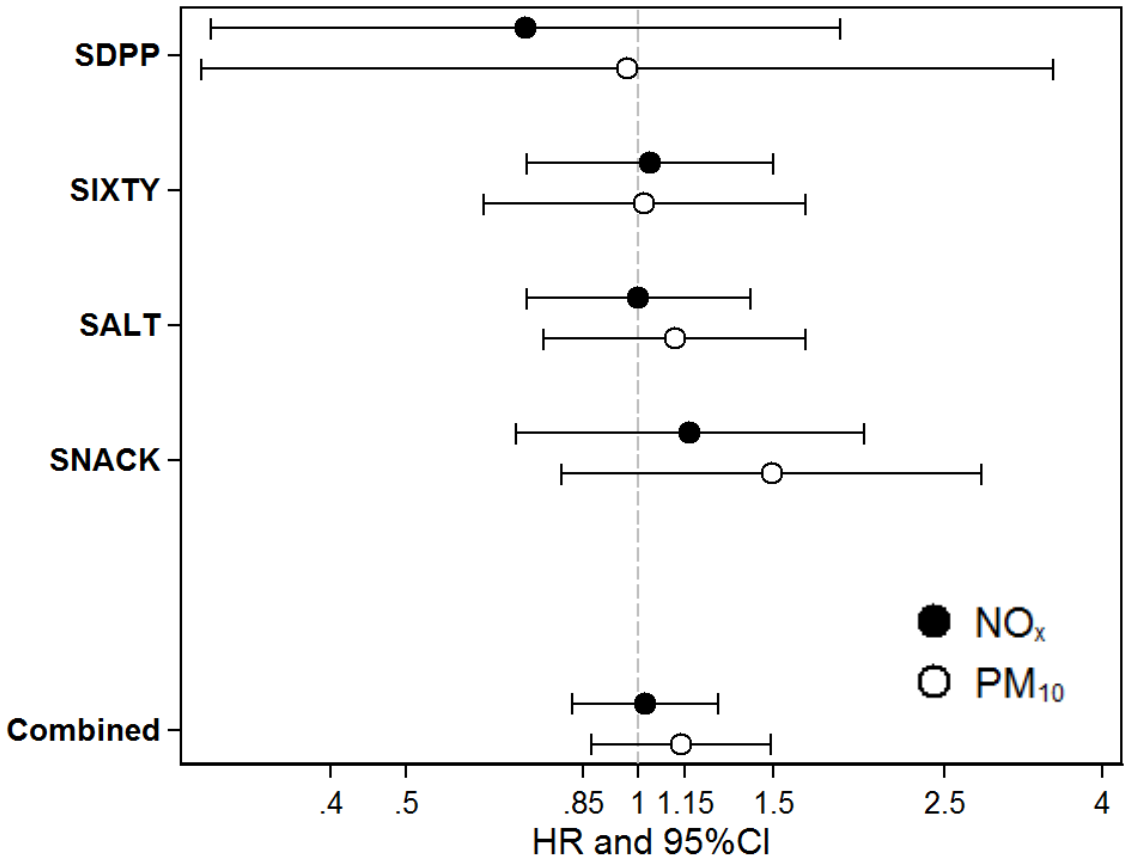


FIGURE 2. Adjusted Hazard ratios (HR) of coronary events, in relation to time-window exposure to NO_x (per 20 µg/m^3) and PM_{10} (per 10 µg/m^3) from road traffic in a meta-analysis of four cohorts from Stockholm County

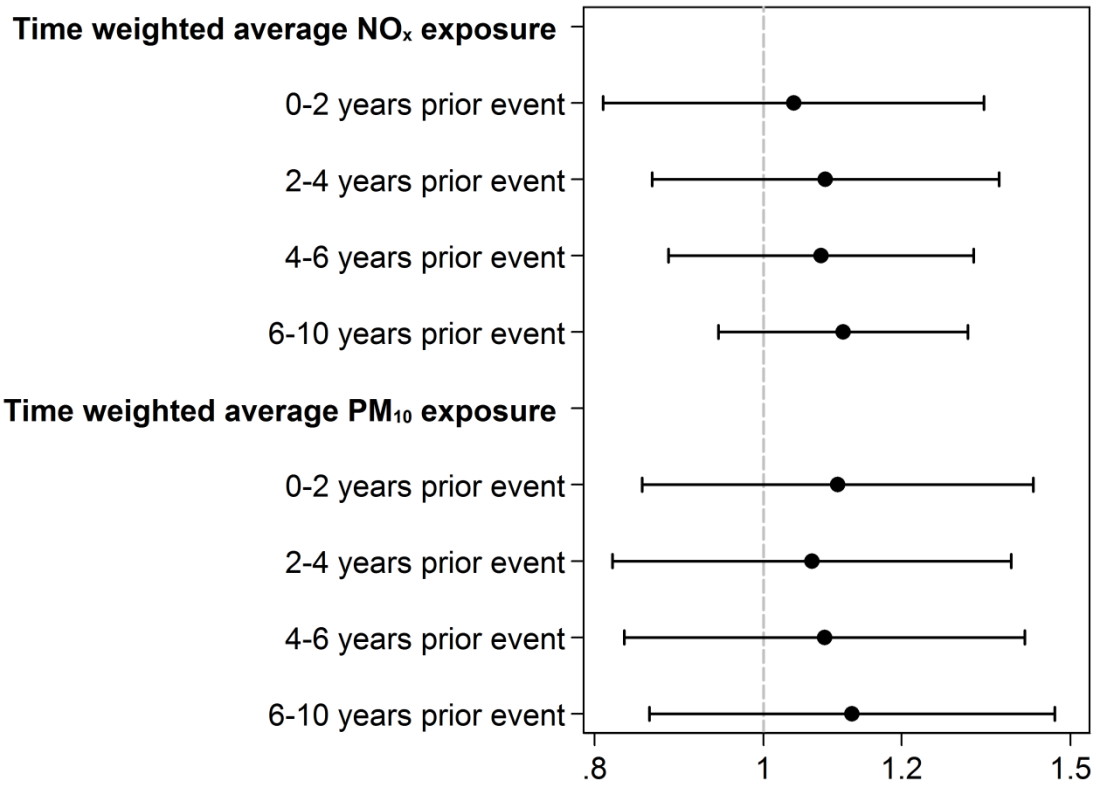
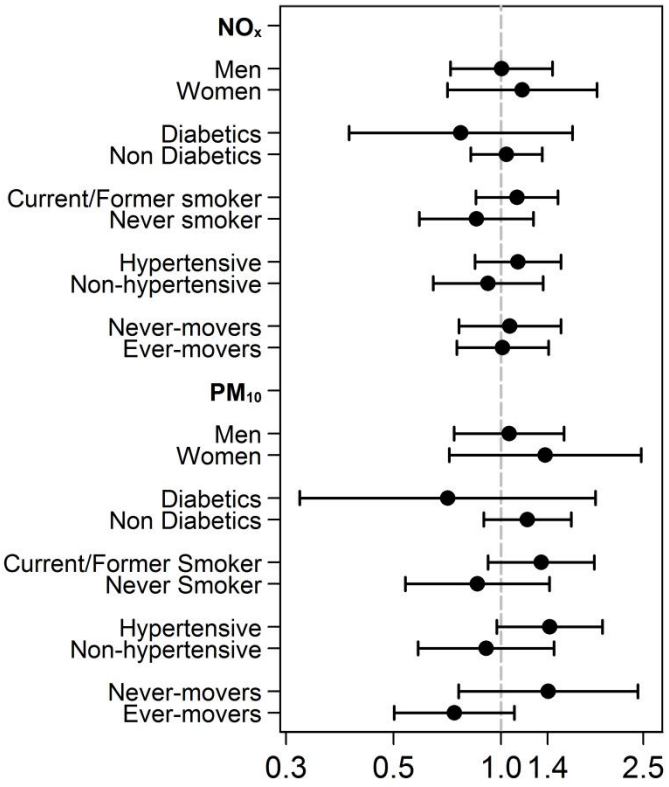


FIGURE 3. Effect modification by gender, smoking, diabetes, hypertension and relocation during follow-up of the association between NO_x or PM_{10} at recruitment and coronary events in a meta-analysis of four cohorts from Stockholm County



Supplementary Materials

Long-term exposure to traffic-related air pollution and coronary events in four cohorts from Stockholm

Content:

Supplemental Table S1: Characteristics of the study participants included in the analyses of four cohorts in Stockholm County

	SDPP[a]	SIXTY[b]	SALT[c]	SNACK[d]
N	7450	3697	6004	2917
Years of enrolment	1992-1998	1997-1999	1998-2002	2001-2004
Age, years: median (minimum - maximum)	48 (35- 56)	60 (59 - 61)	56 (42- 97)	72 (60- 104)
Male (%)	42.6	45.9	42.0	35.1
Education (%)				
Primary school or less	25.5	27.1	21.3	24.5
Up to secondary school or equivalent	45.4	44.3	42.8	42.2
University degree and more	29.1	28.7	35.9	33.3
Occupation status (%)				
Employed/self-employed	92.5	52	N/A[h]	26.6
Unemployed	7.5	9.9	N/A[h]	N/A[h]
Homemaker/housewife	N/A[h]	7.7	N/A[h]	N/A[h]
Retired	N/A[h]	30.3	N/A[h]	73.4
Smoking status (%)				
Current smoker	25.4	19.3	19.8	14.4
Former smoker	36.7	39.5	44.4	36.7
Never smoker	38.0	41.3	35.9	48.9
Number of cigarettes/day for current smokers mean ± SD	13.4 (7.3)	13.0 (7.1)	13.0 (7.6)	10.7 (7.9)
Alcohol consumption (%)				
Daily	7.4	4.3	N/A[h]	20.1
Weekly	37.1	20.3	N/A[h]	24.3
Seldom	51.5	45.2	N/A[h]	46.6
Never	4.0	30.2	N/A[h]	9.1
Diagnosed diabetes[e] (%)				
Yes	1.5	3.9	4.1	8.7
No	98.5	96.2	96.0	91.3
Diagnosed hypertension[f] (%)				
Yes	24.1	51.7	22.2	30.1
No	75.9	48.3	77.8	69.9
Socioeconomic index[g] (%)				
Low	28.3	22.1	29.3	18.4
Medium	26.7	55.4	52.5	30
High	45.0	22.6	18.2	51.6
Physical activity (%)				
Once a month or less /<1h/week	10.7	68.6	25.6	N/A[h]
About once a week /~1h/week	81.5	23.8	62.7	N/A[h]
3 times a week or more />2h/week	7.8	7.6	11.9	N/A[h]
Average household income in neighborhood (SEK) mean ± (SD)	288638 (50935)	301303 (84981)	305654 (82263)	351752 (27057)

[a]Stockholm Diabetes Prevention Program study

[b]60-year-old cohort study

[c] Screening Across the Lifespan Twin study[d] Swedish National Study on Aging and Care in Kungsholmen

[e] Information on diabetes was based on glucose tolerance test in SDPP and on questionnaire data in remaining cohorts.

[f] Information on hypertension was defined by blood pressure measurements or intake of blood pressure lowering medication in three cohorts and through questionnaire in SALT.

[g] Individual socioeconomic status was based on current or last (if retired) profession and categorized into low (blue collar worker), medium (low and intermediate level white collar worker, and self-employed) and high (high-level white-collar worker).

[h] Data not available for a sufficient number of individuals

Supplemental Table S2. Associations of modeled annual mean road traffic NO_x (per 20 µg/m^3) and PM_{10} (per10 µg/m^3) exposure at study entry address or in different exposure time windows and the incidence of CE in four cohorts in Stockholm County.

NOx

Cohort	**SDPP**						**SIXTY**						**SALT**						**SNACK**					
Time of exposure	Adjusted for sex HR (95% CI)			Fully adjusted HR (95% CI)[a]			Adjusted for sex HR (95% CI)			Fully adjusted HR (95% CI)[a]			Adjusted for sex HR (95% CI)			Fully adjusted HR (95% CI)[a]			Adjusted for sex HR (95% CI)			Fully adjusted HR (95% CI)[a]		
Study entry	**0.74**	0.29	1.89	**0.72**	0.28	1.83	**1.03**	0.72	1.46	**1.04**	0.72	1.50	**0.98**	0.71	1.36	**1.00**	0.72	1.40	**1.21**	0.72	2.03	**1.17**	0.69	1.97
0-2 years	**0.95**	0.41	2.18	**0.97**	0.42	2.25	**1.14**	0.74	1.77	**1.16**	0.74	1.82	**1.07**	0.74	1.55	**1.03**	0.70	1.52	**1.20**	0.68	2.13	**0.92**	0.51	1.66
2-4 years	**0.89**	0.38	2.08	**0.91**	0.39	2.16	**1.02**	0.68	1.52	**1.02**	0.68	1.54	**1.16**	0.84	1.61	**1.11**	0.78	1.56	**1.65**	0.99	2.74	**1.23**	0.71	2.12
4-6 years	**1.13**	0.54	2.36	**1.13**	0.53	2.39	**1.09**	0.77	1.54	**1.09**	0.76	1.55	**1.08**	0.81	1.44	**0.99**	0.73	1.36	**1.56**	0.99	2.46	**1.26**	0.79	2.03
6-10 years	**0.87**	0.42	1.82	**0.86**	0.41	1.84	**1.11**	0.83	1.47	**1.10**	0.82	1.48	**1.13**	0.90	1.43	**1.08**	0.84	1.38	**1.00**	1.00	1.00	**1.28**	0.88	1.86

PM_{10}

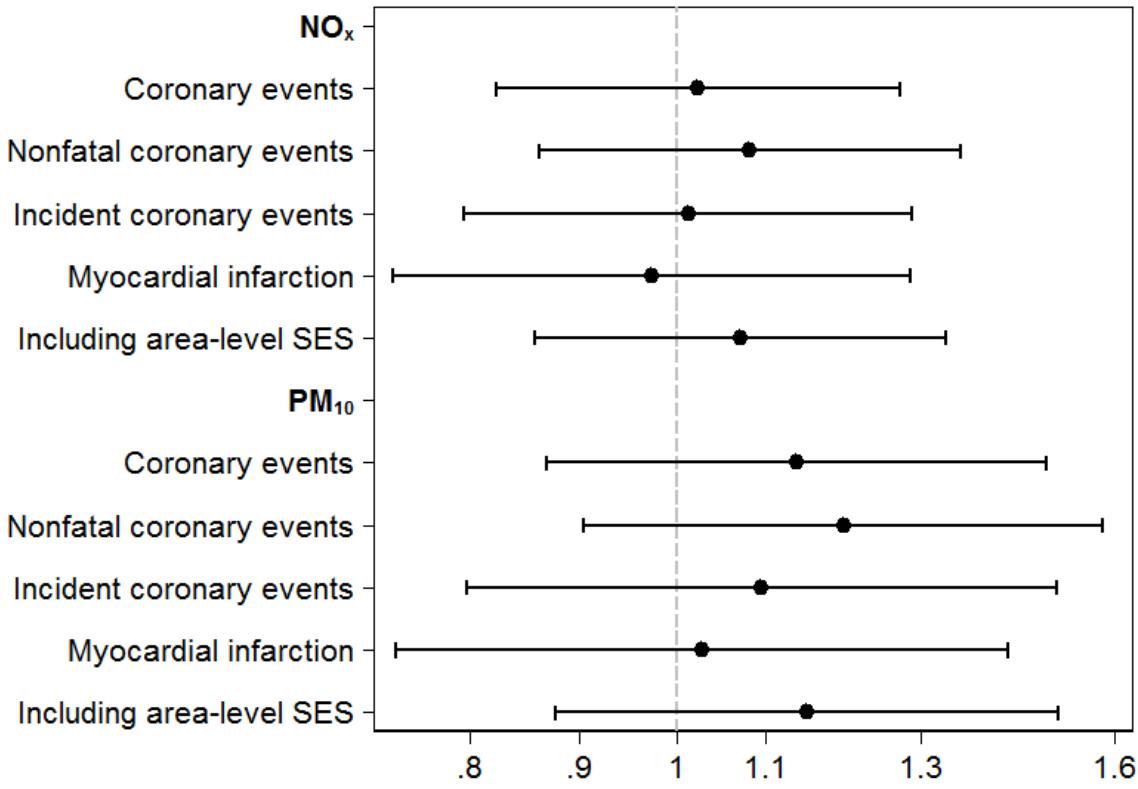
Cohort	SDPP						SIXTY						SALT						SNACK					
Time of exposure	Adjusted for sex HR (95% CI)			Fully adjusted HR (95% CI)[a]			Adjusted for sex HR (95% CI)			Fully adjusted HR (95% CI)[a]			Adjusted for sex HR (95% CI)			Fully adjusted HR (95% CI)[a]			Adjusted for sex HR (95% CI)			Fully adjusted HR (95% CI)[a]		
Study entry	**1.01**	0.29	3.60	**0.97**	0.27	3.45	**1.02**	0.64	1.65	**1.02**	0.63	1.65	**1.13**	0.77	1.68	**1.12**	0.75	1.65	**1.58**	0.85	2.94	**1.49**	0.80	2.79
0-2 years	**1.13**	0.52	2.47	**1.18**	0.54	2.55	**1.09**	0.68	1.74	**1.09**	0.68	1.74	**1.03**	0.69	1.54	**1.03**	0.69	1.54	**1.38**	0.74	2.54	**1.28**	0.69	2.40
2-4 years	**1.15**	0.51	2.61	**1.19**	0.53	2.70	**0.98**	0.61	1.59	**0.98**	0.60	1.60	**1.00**	0.67	1.49	**0.99**	0.66	1.49	**1.45**	0.78	2.67	**1.35**	0.73	2.52
4-6 years	**1.21**	0.51	2.86	**1.23**	0.51	2.92	**1.04**	0.64	1.68	**1.03**	0.63	1.67	**1.02**	0.68	1.52	**1.01**	0.68	1.52	**1.37**	0.74	2.54	**1.30**	0.70	2.42
6-10 years	**0.88**	0.31	2.46	**0.88**	0.31	2.47	**1.07**	0.66	1.74	**1.05**	0.64	1.72	**1.09**	0.73	1.62	**1.08**	0.72	1.62	**1.46**	0.80	2.68	**1.47**	0.80	2.68

[a]The fully adjusted models included; gender, education, smoking status, smoking frequency, alcohol consumption, socioeconomic index and physical activity where occupation status and alcohol consumption were missing in the SALT cohort and physical activity in the SNAC-K cohort

Supplemental Table S 3. Adjusted Hazard ratios (HR) of CE in relation to exposure per NO_x (20 µg/m^3) and PM_{10} (10 µg/m^3) from road traffic at study entry and during different time-windows in a meta-analysis of four cohorts from Stockholm County

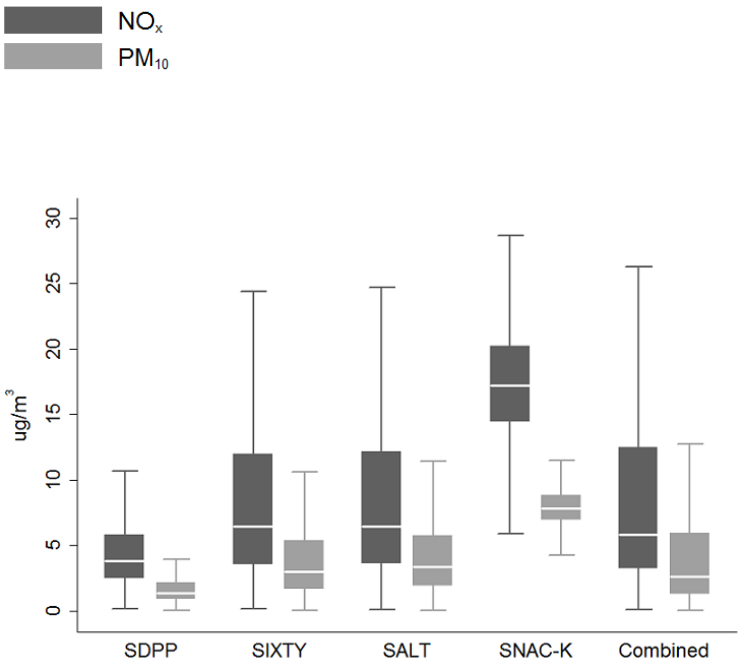
Exposure	Time of exposure	HR (95% CI)	
NO_x	At entry	1.02	(0.82-1.27)
	0-2 years	1.04	(0.81-1.34)
	2-4 years	1.09	(0.86-1.37)
	4-6 years	1.08	(0.88-1.32)
	6-10 years	1.11	(0.94-1.31)
PM_{10}	At entry	1.14	(0.87-1.49)
	0-2 years	1.10	(0.85-1.43)
	2-4 years	1.07	(0.82-1.39)
	4-6 years	1.08	(0.83-1.41)
	6-10 years	1.12	(0.86-1.47)

Supplemental Figure S1. Adjusted Hazard ratios for total number of coronary events or restricted to nonfatal, incident or myocardial infarction, or including area-level SES, in relation to exposure to NO_x (per 20 µg/m^3) and PM_{10} (per 10 µg/m^3) from road traffic in a meta-analysis of four cohorts from Stockholm County



Supplementary figure S2

Modeled exposure levels of traffic generated NO_x and PM_{10} ($\mu g/m^3$) at study entry addresses in four cohorts from Stockholm



Notes: Box plots are defined by the 75, 50, 25th percentile, whiskers indicating the 1.5 IQR of the nearer quartile

Environment International 73 (2014) 382–392

Contents lists available at ScienceDirect

Environment International

journal homepage: www.elsevier.com/locate/envint

Comparing land use regression and dispersion modelling to assess residential exposure to ambient air pollution for epidemiological studies

Kees de Hoogh [a,b,c,*], Michal Korek [c], Danielle Vienneau [a,b], Menno Keuken [e], Jaakko Kukkonen [f], Mark J. Nieuwenhuijsen [g,h,i], Chiara Badaloni [j], Rob Beelen [k], Andrea Bolignano [l], Giulia Cesaroni [j], Marta Cirach Pradas [g,h,i], Josef Cyrys [m,n], John Douros [o], Marloes Eeftens [a,b,k], Francesco Forastiere [j], Bertil Forsberg [p], Kateryna Fuks [q], Ulrike Gehring [k], Alexandros Gryparis [r], John Gulliver [c], Anna L Hansell [c,s], Barbara Hoffmann [q,t], Christer Johansson [u], Sander Jonkers [e], Leena Kangas [f], Klea Katsouyanni [r,v], Nino Künzli [a,b], Timo Lanki [w], Michael Memmesheimer [x], Nicolas Moussiopoulos [o], Lars Modig [p], Göran Pershagen [d], Nicole Probst-Hensch [a,b], Christian Schindler [a,b], Tamara Schikowski [a,b,q], Dorothee Sugiri [q], Oriol Teixidó [y], Ming-Yi Tsai [a,b,z], Tarja Yli-Tuomi [w], Bert Brunekreef [k,aa], Gerard Hoek [k], Tom Bellander [d,ab]

[a] Swiss Tropical and Public Health Institute, Basel, Switzerland
[b] University of Basel, Basel, Switzerland
[c] MRC-PHE Centre for Environment and Health, Department of Epidemiology and Biostatistics, Imperial College London, London, United Kingdom
[d] Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden
[e] Netherlands Organization for Applied Research, Utrecht, The Netherlands
[f] Finnish Meteorological Institute, Helsinki, Finland
[g] Center for Research in Environmental Epidemiology (CREAL), Barcelona, Spain
[h] IMIM (Hospital del Mar Research Institute), Barcelona, Spain
[i] CIBER Epidemiología y Salud Pública (CIBERESP), Spain
[j] Epidemiology Department, Lazio Regional Health Service, Rome, Italy
[k] Institute for Risk Assessment Sciences, Utrecht University, P.O. Box 80178, 3508 TD Utrecht, The Netherlands
[l] Environmental Protection Agency, Lazio Region, Italy
[m] Helmholtz Zentrum München, German Research Center for Environmental Health, Institutes of Epidemiology I and II, Neuherberg, Germany
[n] University of Augsburg, Environmental Science Center, Augsburg, Germany
[o] Laboratory of Heat Transfer and Environmental Engineering, Aristotle University of Thessaloniki, Aristotle University, Thessaloniki, Greece
[p] Department of Public Health and Clinical Medicine, Occupational and Environmental Medicine, Umeå University, Sweden
[q] IUF Leibniz Research Institute for Environmental Medicine, University of Düsseldorf, Düsseldorf, Germany
[r] Department of Hygiene, Epidemiology and Medical Statistics University of Athens, Medical School, Athens, Greece
[s] Directorate of Public Health and Primary Care, Imperial College Healthcare NHS Trust, London, UK
[t] Medical Faculty, Heinrich-Heine University of Düsseldorf, Düsseldorf, Germany
[u] Department of Applied Environmental Science, Stockholm University, Stockholm, Sweden
[v] Department of Primary Care & Public Health Sciences and Environmental Research Group, King's College London, United Kingdom
[w] Department of Environmental Health, National Institute for Health and Welfare (THL), Kuopio, Finland
[x] Rhenish Institute for Environmental Research (RIU), Köln, Germany
[y] Energy and Air quality Department, Barcelona Regional, Barcelona, Spain
[z] Department of Environmental and Occupational Health Sciences, University of Washington, Seattle, WA, United States
[aa] Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands
[ab] Centre for Occupational and Environmental Medicine, Stockholm County Council, Stockholm, Sweden

* Corresponding author at: Department of Epidemiology and Public Health, Swiss Tropical and Public Health Institute, Socinstrasse 57, CH-4051, Basel, Switzerland. Tel.: +41 61 284 8749; fax: +41 61 284 8105.
E-mail address: c.dehoogh@unibas.ch (K. de Hoogh).

http://dx.doi.org/10.1016/j.envint.2014.08.011

ARTICLE INFO

Article history:
Received 11 July 2014
Accepted 19 August 2014
Available online 16 September 2014

Keywords:
Land use regression
Dispersion modelling
Air pollution
Exposure
Cohort

ABSTRACT

Background: Land-use regression (LUR) and dispersion models (DM) are commonly used for estimating individual air pollution exposure in population studies. Few comparisons have however been made of the performance of these methods.
Objectives: Within the European Study of Cohorts for Air Pollution Effects (ESCAPE) we explored the differences between LUR and DM estimates for NO_2, PM_{10} and $PM_{2.5}$.
Methods: The ESCAPE study developed LUR models for outdoor air pollution levels based on a harmonised monitoring campaign. In thirteen ESCAPE study areas we further applied dispersion models. We compared LUR and DM estimates at the residential addresses of participants in 13 cohorts for NO_2; 7 for PM_{10} and 4 for $PM_{2.5}$. Additionally, we compared the DM estimates with measured concentrations at the 20–40 ESCAPE monitoring sites in each area.
Results: The median Pearson *R* (range) correlation coefficients between LUR and DM estimates for the annual average concentrations of NO_2, PM_{10} and $PM_{2.5}$ were 0.75 (0.19–0.89), 0.39 (0.23–0.66) and 0.29 (0.22–0.81) for 112,971 (13 study areas), 69,591 (7) and 28,519 (4) addresses respectively. The median Pearson *R* correlation coefficients (range) between DM estimates and ESCAPE measurements were of 0.74 (0.09–0.86) for NO_2; 0.58 (0.36–0.88) for PM_{10} and 0.58 (0.39–0.66) for $PM_{2.5}$.
Conclusions: LUR and dispersion model estimates correlated on average well for NO_2 but only moderately for PM_{10} and $PM_{2.5}$, with large variability across areas. DM predicted a moderate to large proportion of the measured variation for NO_2 but less for PM_{10} and $PM_{2.5}$.

1. Introduction

A large number of epidemiological studies have shown a clear association between long-term ambient air pollution exposure and adverse health effects (WHO, 2013). Several of these studies estimated individual air pollution exposures from stationary monitoring data, e.g. by using the nearest air pollution monitor to represent the pollution in entire cities (Dockery et al., 1993) to more complex approaches including spatial interpolation and kriging (Brauer et al., 2008; Künzli et al., 2005). Such methods provide estimates of large-scale spatial differences in air pollution concentrations, but are less effective in assessing intra-urban variation particularly when the number of monitoring sites is small. Recent studies have focused on intra-urban variation of air pollution, using indicators or proxies such as distance to the nearest road as well as pollutant levels estimated by land use regression (LUR), dispersion modelling (DM) including Chemical transport models (CTM) and hybrid models (HEI, 2010).

The LUR method, first developed by Briggs et al. (1997), uses least squares regression to combine monitored data with Geographic Information System (GIS)-based predictor data reflecting pollutant sources, to build a prediction model applicable to non-measured locations, e.g. residential addresses of cohort members. LUR modelling has been increasingly used in epidemiological studies because it is relatively cheap and can be easily implemented on the basis of purpose-designed monitoring campaigns or routinely measured concentrations and appropriate geographic predictors of air pollution sources (Hoek et al., 2008).

DMs are based on detailed knowledge of the physical, chemical, and fluid dynamical processes in the atmosphere. DMs use information on emissions, source characteristics, chemical and physical properties of the pollutants, topography, and meteorology to model the transport and transformation of gaseous or particulate pollutants through the atmosphere to predict, e.g., ground level concentrations (Holmes and Morawska, 2006; Kukkonen et al., 2012). Gaussian based DMs were originally developed as air quality management tools but have also been used in environmental epidemiology to model long-term exposures (Bellander et al., 2001; Wu et al., 2011). Chemical Transport Models have also been used to model short- and long-term exposure periods (Hennig et al., 2014). Few studies to date have conducted comparisons between LUR and DMs for their performance in estimating exposures (Beelen et al., 2010; Cyrys et al., 2005; Dijkema et al., 2011; Gulliver et al., 2011; Marshall et al., 2008; Sellier et al., 2014). These studies included different models, spatial resolution, pollutants and study areas, factors likely to have contributed to inconsistent findings within individual studies. As both LUR and DM are applied in epidemiology, there is a need for more comparison studies of these methods, addressing their respective advantages and strengths depending on the specific air pollution and health-related questions which are sought to be answered.

We compare LUR and DM to assess spatial variation of annual average ambient air pollution estimates at residential addresses within the framework of the European Study of Cohorts for Air Pollution Effects (ESCAPE), not taking into account population activity patterns or indoor air pollution. The ESCAPE study developed LUR models to estimate exposure at the residential addresses of cohort participants based on uniform monitoring campaigns and uniform modelling approaches in 36 study areas (Beelen et al., 2013; Cyrys et al., 2012; de Hoogh et al., 2013; Eeftens et al., 2012a,b). To several of these study areas we apply DM or use existing DM output, allowing for an in depth comparison to better understand the differences and/or agreements between LUR and DM estimates for use in epidemiological studies with long-term exposures. We include a range of exposure environments and populations across Europe, and focus, in particular, on the differences in estimated exposure at the individual participant level which is most relevant for interpretation of epidemiological studies.

2. Materials and methods

We estimated annual average outdoor air pollution concentrations for NO_2 in 13, PM_{10} in 7 and $PM_{2.5}$ in 4 of the 36 European cities/areas included in the ESCAPE study using both LUR and DM (Umeå region, Sweden; Stockholm County, Sweden (PM_{10}); Helsinki—Vantaa region, Finland ($PM_{2.5}$); Bradford, UK; London, UK (PM_{10}); Netherlands (PM_{10} & $PM_{2.5}$); Ruhr Area (PM_{10} & $PM_{2.5}$), Germany; Basel, Switzerland; Geneva, Switzerland; Lugano, Switzerland (PM_{10}); Rome, Italy ($PM_{2.5}$); Barcelona, Spain (PM_{10}); Athens, Greece (PM_{10})). The selection of study areas was based on the availability of existing dispersion models. A general discussion of these two modelling approaches is reported elsewhere (Hoek et al., 2008; Özkaynak et al., 2013).

We conducted several comparisons, depending on the comparability of the model outputs. The main comparison between the methods was made at the residential address of cohorts participants (referred to as LUR-DM). We also compared the DM estimates with measured concentrations at the ESCAPE monitoring sites. This was an independent validation, as monitoring data from the ESCAPE sites were not used as input data in the DM models. Recent studies have documented that the model R^2 and the leave-one out cross-validation R^2 overestimate the predictive ability of LUR models at independent sites (Basagaña

et al., 2012; M. Wang et al., 2013). Therefore we cannot directly compare the explained variance of the LUR models with the explained variance of the dispersion models. Furthermore, we did not have a sufficiently large set of independent monitoring data available within the study areas to serve as an independent test set for both LUR and DM.

2.1. Description of cohorts

We used address locations of cohort participants as the basis for the LUR-DM comparison by study area. The majority of cohorts in this analysis were also used in the ESCAPE health studies: EPIC in Umeå (SE), SDPP, 60 years cohort, SALT and SNAC in Stockholm (SE), FINRISK in Helsinki (FI), Born in Bradford (UK), EPIC-Oxford in London (UK), PIAMA in the Netherlands, Heinz Nixdorf Recall (HNR) study in the Ruhr area (DE), SAPALDIA in Basel, Geneva and Lugano (CH) and SIDRIA in Rome (IT). For Barcelona (ES), we chose the larger population of the ARIBA cohort ($n = 8{,}402$), rather than the ECRHS cohort ($n = 297$) used in ESCAPE. Due to confidentiality, address locations of the EPIC cohort in Athens (GR) were not available; instead we used 1500 randomly selected addresses across the study area to act as a cohort surrogate. Most of the study areas were large cities and the surrounding suburban or rural communities; however, some of the cohorts covered larger regions, such as PIAMA in the Netherlands. In total, we used 112,971 address locations over 13 cohorts.

2.2. Land use regression modelling

The ESCAPE study involved harmonised monitoring campaigns for NO_2 in 36 study areas and PM_{10}/$PM_{2.5}$ in 20 study areas, as described in Cyrys et al. (2012) and Eeftens et al. (2012a). In brief, in each study area a measurement campaign was carried out during three 2-week periods within one year. The complete monitoring period across all study areas was between 2008 and 2011. Ogawa badges were used for monitoring of NO_2 and Harvard Impactors were used for monitoring of PM. Care was taken to select site locations to incorporate relevant intra-urban spatial variation in traffic and land use characteristics. Adjusted annual mean concentrations for each site were then estimated with the aid of measurement data from an all-year running reference site in an urban background location in each study area.

Based on these measurements, LUR models were developed in each study area following a standardised approach (Beelen et al., 2013; Eeftens et al., 2012b). Geographical Information Systems (GIS) predictor variables were collected for all study areas centrally (EU-wide datasets including CORINE land cover, EuroStreets road network, altitude and population density) and locally (traffic data and, where available, more detailed land cover data). Circular buffers with radii of 25, 50, 100, 300, 500, and 1000 m were used to calculate traffic and road variables for each monitoring location. For land use and population, buffers of 100, 300, 500, 1000, and 5000 m were calculated. LUR models were developed combining the adjusted annual means and the GIS predictor data in each study area following a stringent set of rules. Linear regression was performed in a stepwise logical standardised approach, detailed by Eeftens et al. (2012b). Predictors giving the highest adjusted R^2 were subsequently added to the model if they conformed to the direction of effect defined a priori and added more than 1% to the adjusted R^2. Final models were checked for p-value (removed when p-value >0.10), co-linearity (variables with Variance Inflation Factor (VIF) >3 were removed and model rerun) and influential observations (models with Cook's D >1 were further examined). The final models were evaluated by leave-one-out cross validation (LOOCV).

Model structure, model R^2 and LOOCV R^2 of the LUR models in the included 13 study areas are shown in Table A.1. LUR model predictions at the cohort address were based on predictor values restricted to the range of observed values at the monitoring sites, in order to prevent extrapolation beyond the range for which the model was developed.

2.3. Dispersion modelling

DM was applied in the 13 study areas by third parties using input data including traffic flow, road geometry, other non-traffic pollution sources (e.g. industrial and agricultural sources), meteorological parameters and concentrations measured at regional and urban background sites. In ten of the 13 study areas a Gaussian plume DM was used: Airviro in Stockholm and Umeå Region; CAR-FMI in Helsinki; ADMS-Urban in Bradford, London and Barcelona; CAR and Pluim Snelweg (motorway) in the Netherlands; Pollumap DM 2010 in Basel, Geneva, and Lugano. Two areas used Eulerian or chemical transport models: EURAD-CTM in Ruhr area; Flexible Air quality Regional Model (FARM) in Rome and one used a Computational Fluid Dynamic (CFD) model; MEMO/MARS-aero in Athens. Information about the DM by study area is shown in Table 1. Models differed in the sources included (all models including traffic sources but some additionally including industry and agricultural sources), the scale of assessment and the representation of regional background (most used routine monitoring data). The effective spatial scale of the receptor-oriented methods depends on several factors, e.g. the precision of the spatial description of sources and topography, and could not be estimated. DM estimates were extracted to the addresses of the cohorts involved.

2.4. Statistical analyses

Exposure estimates from LUR and DM were compared at the address level. We calculated Pearson (R) and Spearman (Rho) correlation coefficients and show scatterplots of the relationship. The LUR and DM exposure estimates were also categorised into quintiles as epidemiological studies often use categorical analyses to relax the assumption of a linear association. Kappa coefficients were calculated to assess the level of agreement beyond chance. Bland–Altman plots were produced to further investigate the agreement between the two methods, specifically to test whether the difference between LUR and DM depends on the absolute concentrations. In addition, the correlation between the DM estimates and monitored concentrations at the ESCAPE monitoring sites was calculated (R and Rho) and visualised in scatterplots (DM-MON).

Statistical analysis was carried out in STATA version 11.0 (StataCorp LP, College Station, Texas, USA).

3. Results

3.1. Comparison of LUR and DM at address level

Distributions of LUR and DM predictions at the cohort addresses, the correlation and Kappa coefficients are shown in Table 2. Fig. 1 shows scatterplots of LUR and DM predictions.

3.1.1. NO_2

LUR and DM estimates of NO_2 levels for cohort members were available for the 13 study areas at a total of 112,971 residential addresses. The correlation (Pearson R) between LUR and DM estimates of NO_2 levels at cohort addresses varied from 0.19 (Athens) to 0.89 (The Netherlands; Fig. 1, Table 2). The Spearman rank correlation (R) ranged from 0.21 to 0.90. The median Pearson and Spearman correlation coefficients were 0.75 and 0.77 respectively, indicating overall good agreement. The agreement by quintiles ranged from 24% to 62%. Kappa statistics ranged from 0.005 to 0.52 (Table 2).

The overall median of estimated NO_2 concentrations was slightly higher for LUR (21.4 μg/m^3) than for DM predictions (17.3 μg/m^3). The difference between LUR and DM median estimates was up to 11.9 μg/m^3 (Rome; Table 2). In the areas with the largest differences between LUR and DM estimates, the DM/CTM modelled an average concentration over an area of 0.25–1 km^2, in contrast to LUR which modelled concentrations at individual address (receptor) points. The

Table 1
Details of atmospheric dispersion models used to predict air pollution concentrations in each study area.

Study area	Name of dispersion model	Type	Pollutants	Geographical resolution output	Year output	Regional background	Sources	Street canyon	Reference(s)
Umeå region, SE	Airviro Gauss dispersion model	Gaussian plume	NO_2	50 × 50 m	2010	Monitoring[a]	T, P, R[b]	No	SMHI (1993)
Stockholm County, SE	Airviro Gauss dispersion model	Gaussian plume	NO_2, PM_{10}	25 × 25 m in urban, 500 × 500 m in rural area	2009	Monitoring[a]	T	Yes	SMHI (1993)
Helsinki-Vantaa region. FI	CAR-FMI (Contaminants in the Air from a Road – Finnish Meteorological Institute)	Gaussian plume	NO_2, $PM_{2.5}$	At unique receptor points	2010	Monitoring[a]	T	No	Kukkonen et al. (2001), Karppinen et al. (2000)
Bradford, UK	ADMS-Urban	Gaussian plume	NO_2	At unique receptor points	2009	Monitoring[a]	T, A	No	Carruthers et al. (2000)
London, Oxford, UK	ADMS-Urban	Gaussian plume	NO_2, PM_{10}	10 × 10 m	2011	Monitoring[a]	T, A	Yes	Carruthers et al. (2000)
Netherlands	GCN (Generic Concentrations in the Netherlands), for the regional/urban background, Pluim Snelweg for the motorways and provincial roads, CAR model for the urban roads	Gaussian plume	NO_2, PM_{10}, $PM_{2.5}$	25 × 25 m	2009	Model[c]	T, P, A	Yes, included in CAR	Velders et al. (2013) Wesseling and Visser (2003) Wesseling and Sauter (2007)
Ruhr Area, DE	EURopean Air Pollution Dispersion (EURAD) model system	Dispersion and chemical transport model	NO_2, PM_{10}	1 × 1 km	2006-2008	Monitoring	T, P, R, A	No	Memmesheimer et al. (2004)
Basel, Geneva and Lugano, CH	Pollumap dispersion model 2010	Gaussian plume	NO_2 (All), PM_{10} (Lugano only)	100 × 100 m	2010		T, A	No	SAEFL (2003) Gariazzo et al. (2007)
Rome, IT	Flexible Air quality Regional Model (FARM)	Eulerian chemical transport model	NO_2, $PM_{2.5}$	1 × 1 km	2007		T, A	No	Gariazzo et al. (2007) Finardi et al. (2009)
Barcelona, ES	ADMS-Urban	Gaussian plume	NO_2, PM_{10}	5 × 5 m for NO_2, 100 × 100 m for PM_{10}	2008	Monitoring[a]	T, P, R, A	Yes	Carruthers et al. (2000)
Athens, GR	MEMO/MARS-aero	Eulerian chemical transport model	NO_2, PM_{10}	500 × 500 m	2008	Model		No	Moussiopoulos et al. (2012)

[a] Monitoring data from regional background station.
[b] T = traffic; P = point sources; R = residential heating; A = area source for all non-traffic sources.
[c] Combination of monitoring and modelling at 1 × 1 km scale.

relative difference between the median NO_2 LUR and DM predictions was however not large in these areas (<~30%).

The estimated ranges of NO_2 concentrations differed for the two methods, with some study areas showing a distinctly narrower range for LUR estimates compared to DM estimates (Bradford and the Netherlands) and other study areas showing a larger range for LUR estimates than for DM estimates (Ruhr Area, Athens, Lugano, Barcelona, and London).

3.1.2. PM_{10}

PM_{10} concentrations were modelled with LUR and DM for 69,591 residential addresses in 7 study areas. The correlation between LUR and DM was generally lower and the differences in levels larger than for NO_2 (Table 2, Fig. 1). A large difference of 20 μg/m^3, for instance, was found between median PM_{10} concentrations for LUR and DM in Athens, whereas the differences in the Netherlands and Lugano were small (0.3 and 1.2 μg/m^3 respectively). The median Pearson and Spearman correlation coefficients between LUR and DM estimates were 0.39 and 0.49 respectively. Lugano, the Netherlands and London showed the highest correlations (Pearson) between the 2 methods (R = 0.66, 0.56 and 0.52 respectively). In several of the LUR predictions the impact of truncation to the highest value of predictor variables at the monitoring sites is visible, e.g. in the Netherlands (Fig. 1). In Stockholm, the dispersion model had a lower bound, defined by the measured regional background used as input in the model. The percentage of agreement by quintiles ranged from 25 to 55%.

3.1.3. $PM_{2.5}$

Estimated $PM_{2.5}$ concentrations were modelled in four study areas (Helsinki—Vantaa region, the Netherlands, the Ruhr Area, and Rome) for a total of 28,159 residential addresses. In the Netherlands there was a high correlation (Pearson R = 0.81), with similar median $PM_{2.5}$ concentrations for both methods, but with a larger range for DM estimates (14.5 μg/m^3) compared to LUR estimates (6.2 μg/m^3). The other three study areas showed low correlations between the LUR and DM estimates.

The Bland–Altman plots (Fig. A.1) were inspected to assess the agreement over the concentration range between the two methods. The majority of points were located within +/− 2 times the standard deviation; however, there were quite different patterns for the different study areas and pollutants. Fig. A.1 shows that bias rarely is zero (only Basel (NO_2), Netherland (NO_2, PM_{10}, and $PM_{2.5}$), Lugano (PM_{10}) and Helsinki, and Rome ($PM_{2.5}$) have an absolute mean difference of less than 1 μg/m^3). Secondly the upper- and lower-limits of the 95% range

Table 2
Descriptive and comparison statistics of LUR and dispersion estimates (μg/m^3) at cohort address for NO_2, PM_{10} and $PM_{2.5}$.

		LUR predictions (μg/m^3)				DM predictions (μg/m^3)				Comparison of LUR with DM						
										Continuous: DM = Constant + Slope × LUR					Quintiles	
Study area	N[a]	Median	P_{05}	P_{95}	P_{95}–P_{05}	Median	P_{05}	P_{95}	P_{95}–P_{05}	Spearman's Rho	Pearson *R*	Constant	Slope	RMSE	Agreement (%)[b]	Kappa
NO_2																
Umeå region, SE[c]	4575	6.8	4.1	16.4	12.2	12.5	5.6	20.6	15.0	0.782	0.792	5.17	0.93	2.63	48.3	0.352
Stockholm County, SE[c]	39409	9.6	6.4	20.9	14.5	6.5	3.3	18.1	14.9	0.791	0.856	−1.98	0.93	2.46	48.9	0.361
Helsinki—Vantaa region, FI[c]	5871	16.0	9.0	25.5	16.5	9.0	7.0	17.0	10.0	0.762	0.745	2.01	0.52	2.34	43.7	0.297
Bradford, UK[c]	20919	24.0	18.9	29.0	10.1	18.3	14.0	26.5	12.5	0.820	0.667	−1.62	0.86	3.06	49.2	0.365
London, UK[c]	7089	33.3	21.7	45.5	23.8	32.0	21.1	42.6	21.4	0.836	0.798	8.55	0.70	4.05	55.2	0.441
Netherlands[c]	7295	22.7	12.7	33.9	21.2	24.0	11.4	38.2	26.8	0.901	0.891	−2.37	1.13	3.70	61.8	0.523
Ruhr Area, DE[d]	4809	29.6	23.4	38.6	15.2	37.5	30.8	44.1	13.3	0.428	0.389	28.45	0.30	3.51	31.0	0.138
Basel, SU[c]	1118	29.0	18.3	34.3	16.0	30.5	21.4	34.4	13.1	0.771	0.768	11.11	0.65	2.71	48.9	0.362
Geneva, SU[c]	737	26.4	16.2	38.7	22.6	31.7	24.4	36.0	11.7	0.708	0.657	21.73	0.36	2.84	41.4	0.267
Lugano, SU[c]	1090	26.6	11.8	39.2	27.3	30.9	22.9	34.8	12.0	0.773	0.819	20.43	0.37	1.97	50.2	0.377
Rome, IT[d]	10157	38.1	25.5	56.1	30.5	50.0	31.5	59.4	27.8	0.406	0.386	33.35	0.36	7.65	29.4	0.120
Barcelona, ES[c]	8402	57.1	38.5	85.1	46.6	54.0	39.7	78.4	38.7	0.687	0.688	21.41	0.59	8.84	43.3	0.292
Athens, GR[d]	1500	36.0	23.4	59.5	36.0	47.0	36.5	56.4	19.8	0.207	0.188	42.86	0.10	6.35	23.9	0.005
All	112971	21.4				17.3										
PM_{10}																
Stockholm County, SE[c]	39409	15.1	6.2	20.4	14.2	10.0	7.8	16.6	8.8	0.378	0.367	6.83	0.29	2.82	31.2	0.140
London, UK[c]	7089	16.9	14.9	20.9	6.1	21.7	20.7	23.0	2.4	0.554	0.517	17.94	0.22	0.65	55.2	0.441
Netherlands[c]	7295	24.6	23.8	27.1	3.3	24.9	20.4	27.2	6.7	0.625	0.556	−4.88	1.16	1.91	42.0	0.275
Ruhr Area, DE[d]	4809	27.5	25.3	31.6	6.3	18.0	15.1	22.5	7.4	0.328	0.346	5.97	0.43	2.18	24.8	0.060
Lugano, SU[c]	1087	23.3	18.0	27.4	9.4	24.5	20.4	25.9	5.5	0.575	0.659	13.87	0.43	1.25	39.8	0.248
Barcelona, ES[c]	8402	39.0	37.0	47.5	10.6	37.4	35.7	44.2	8.5	0.495	0.393	24.14	0.35	2.62	33.1	0.163
Athens, GR[d]	1500	47.0	24.7	64.1	39.4	27.0	23.4	30.3	7.0	0.272	0.233	24.70	0.046	2.36	26.5	0.080
All	69591	16.6				15.1										
$PM_{2.5}$																
Helsinki—Vantaa region, FI[c]	5871	8.0	5.6	9.1	3.5	8.5	8.2	9.3	1.1	0.215	0.252	7.85	0.093	0.37	25.8	0.073
Netherlands[c]	7295	16.5	15.4	17.3	1.9	16.8	13.1	18.6	5.6	0.879	0.812	−20.40	2.23	0.41	50.4	0.380
Ruhr Area, DE[d]	4809	18.3	16.9	20.4	3.5	14.7	13.1	16.7	3.6	0.391	0.327	8.21	0.35	1.12	28.0	0.100
Rome, IT[d]	10544	18.9	17.3	23.3	6.0	20.1	16.5	21.6	5.0	0.252	0.223	16.03	0.19	1.53	26.5	0.081
All	28159	17.4				16.8										

[a] Number of residential addresses in the participating cohorts.
[b] Percentage of residential addresses falling in the same quintile.
[c] Spatial resolution of DM estimates ≤100 × 100 m.
[d] Spatial resolution of DM estimates ≥500 × 500 m.

differ widely between the study areas. Fig. A1 also shows which of the two methods tends to provide higher or lower concentration estimates. For example NO_2 estimates in Bradford are mostly higher with LUR (95% range = −12.7 to 0.7 μg/m^3) while the opposite is true in Umeå (95% range = −0.6 to 9.8 μg/m^3). For the coarse-scale models and the three Swiss models, the DM model predictions were lower than the LUR predictions for the highest concentrations, mostly traffic locations.

3.2. Comparison DM with ESCAPE monitoring results

Correlations between dispersion modelled annual average concentrations and adjusted annual average concentrations based on measurements at the ESCAPE monitoring sites are shown in Table 3 (scatterplots in Fig. 2, Table A.2). Pearson *R*'s correlation coefficients ranged from 0.09 (Athens) to 0.86 (Umeå) for NO_2, with a median of 0.74. Dispersion models that aimed to predict at specific receptor points or predict with a very small resolution of <100 × 100 m predicted NO_2 concentrations better than the coarser Eulerain/CFD models. The median correlation for PM_{10} (0.58, ranging from 0.36 (Barcelona) to 0.88 (London)) was lower than for NO_2, which again was mainly driven by the difference in scale. Among the four study areas with a DM for $PM_{2.5}$, the two models that estimated at unique receptor points or on a small spatial scale (Helsinki—Vantaa region, Netherlands) predicted measured concentrations with correlations of 0.66 and 0.54 (Pearson), respectively. Correlations for the larger scale models were 0.39 (the Ruhr Area) to 0.61 (Rome). For most of the study areas Spearman correlations were moderate to high (ranges: NO_2 0.15 to 0.88; PM_{10} 0.47 to 0.70 and $PM_{2.5}$ 0.49 to 0.70). For the majority of the study areas DM thus tend to predict a fairly large proportion ($R > 0.6$) of the variation across the measurement sites. Scatter plots of the DM-MON comparison are shown in Fig. 2. The regression lines for NO_2 generally follow the 1:1 line, whereas regression lines for PM_{10} and $PM_{2.5}$ show departures from the 1:1 line. Relatively large differences in NO_2 concentrations were found only in Umeå (DM > measured) and Helsinki—Vantaa region (DM < measured), though in both areas the correlation was higher than 0.6. PM_{10} concentrations were higher than the model predictions in Athens, though the correlation was reasonable. Fig. 3 illustrates that the agreement between LUR and DM at the cohort addresses increases with increasing correlation between the DM and measured concentrations at the monitoring sites. The agreement between LUR and DM did not depend on the LOOCV of the LUR model.

4. Discussion

To our knowledge, this is the first study to compare LUR and DM for assigning air pollution exposures to a large number of residential addresses in different geographical areas. In general, a distinction between two types of DM can be made: one estimates receptor-specific concentrations (Gaussian) and the other estimates average concentrations for an area (Eulerian/CFD). This has potential implications on the comparability of air pollution estimates at the address level and for the downstream epidemiology.

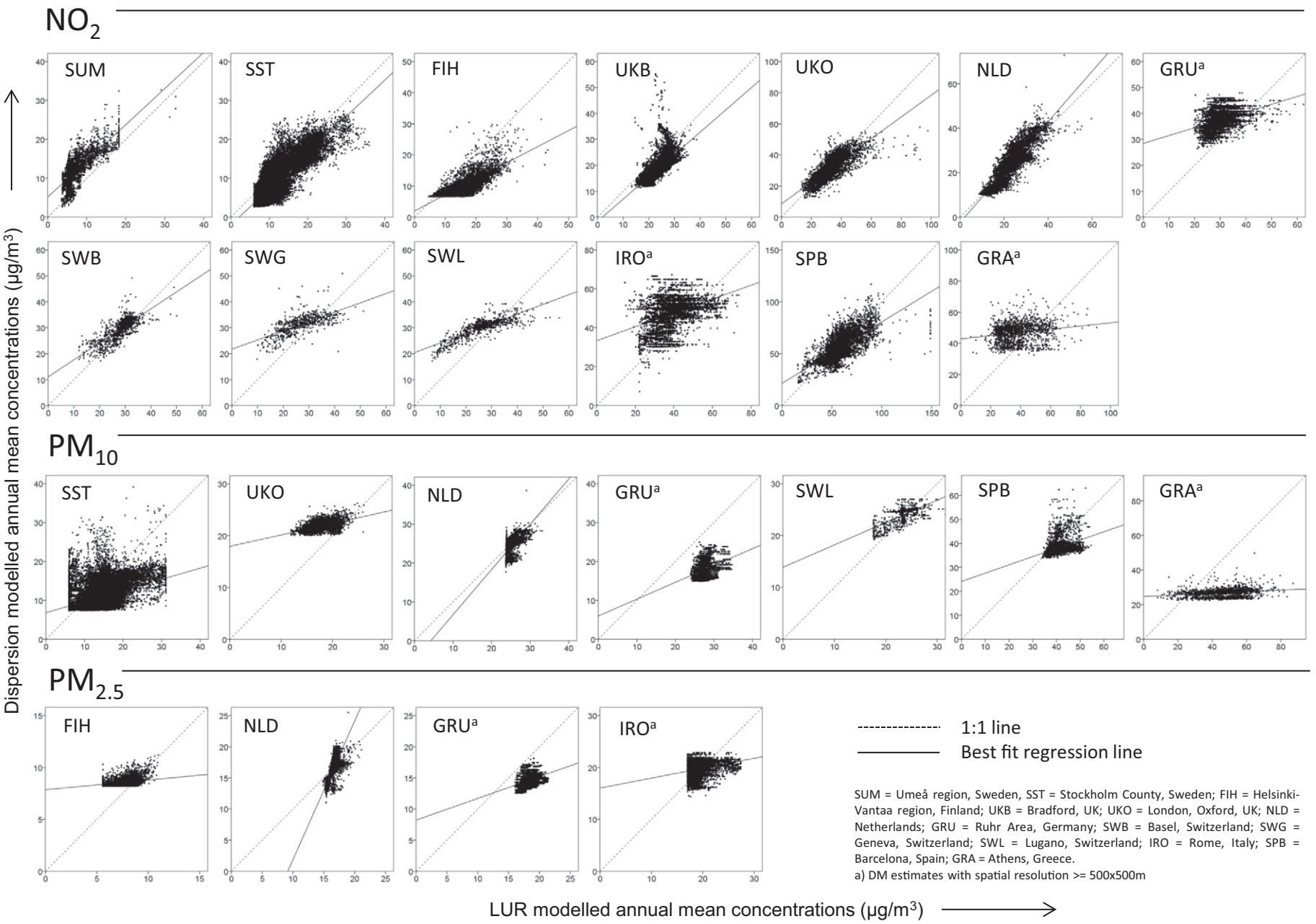


Fig. 1. Scatter plots of DM (y-axis) against LUR (x-axis) estimates (µg/m^3) at residential addresses for NO_2, PM_{10} and $PM_{2.5}$.

Table 3
Descriptive and comparison statistics of DM estimates and measurements (μg/m^3) at ESCAPE monitoring sites for NO_2, PM_{10} and $PM_{2.5}$.

		Monitored concentrations at ESCAPE sites (μg/m^3)			DM predictions at ESCAPE sites (μg/m^3)			Comparison of DM predictions with measured concentrations at ESCAPE sites				
Study area	N[a]	Median	Min	Max	Median	Min	Max	Spearman's Rho	Pearson R	Constant	Slope	RMSE
NO_2												
Umeå region, SE[b]	20	9.3	5.3	35.8	15.5	7.4	31.0	0.878	0.858	-5.36	1.02	3.88
Stockholm County, SE[b]	39	14.8	2.1	33.0	13.0	2.9	25.3	0.775	0.755	4.41	0.84	4.94
Helsinki—Vantaa region, FI[b]	25	19.7	12.2	28.5	10.6	6.6	26.7	0.753	0.658	12.64	0.63	3.78
Bradford, UK[b]	40	25.2	16.7	36.7	19.8	13.0	38.0	0.806	0.743	11.99	0.62	3.59
London, UK[b]	27	39.7	29.2	102.7	37.7	23.0	79.9	0.681	0.849	-10.83	1.39	9.06
Netherlands[b]	68	28.0	12.8	57.1	27.7	11.1	47.1	0.897	0.852	1.05	0.99	5.45
Ruhr Area, DE[c]	29	31.2	22.2	58.4	39.2	28.5	50.2	0.459	0.391	5.47	0.72	8.98
Basel, SU[b]	40	31.4	16.1	47.8	31.8	21.4	35.2	0.492	0.598	-14.10	1.46	5.98
Geneva, SU[b]	41	30.1	16.1	51.3	31.2	20.4	40.9	0.642	0.540	-5.90	1.17	7.66
Lugano, SU[b]	42	27.1	12.2	59.2	31.7	23.4	39.3	0.749	0.764	-33.72	2.00	5.37
Rome, IT[c]	40	41.7	13.6	72.6	50.0	26.6	62.1	0.568	0.614	-8.76	1.08	10.96
Barcelona, ES[b]	40	54.7	13.8	109.0	51.2	28.5	78.5	0.805	0.754	-4.61	1.15	13.40
Athens, GR[c]	40	35.9	13.3	71.0	40.4	34.4	52.2	0.154	0.089	27.86	0.20	12.04
PM_{10}												
Stockholm County, SE[b]	19	18.5	5.7	35.6	15.2	7.5	19.3	0.472	0.580	11.33	1.09	5.65
London, UK[b]	13	18.4	16.1	31.2	22.3	21.5	30.7	0.484	0.877	-13.47	1.46	2.03
Netherlands[b]	34	26.2	21.9	33.0	25.7	20.8	30.5	0.671	0.696	3.74	0.88	2.18
Ruhr Area, DE[c]	15	27.4	22.5	33.3	18.2	15.3	32.3	0.521	0.392	22.95	0.25	2.90
Lugano, SU[b]	18	23.9	18.5	32.5	24.1	20.1	25.3	0.552	0.668	-8.77	1.38	2.67
Barcelona, ES[b]	20	38.6	17.8	48.5	36.7	34.5	51.3	0.699	0.356	10.71	0.71	6.82
Athens, GR[c]	20	42.9	27.3	58.0	24.5	22.8	30.3	0.522	0.397	9.42	1.32	6.81
$PM_{2.5}$												
Helsinki—Vantaa region, FI[b]	13	8.9	7.9	10.4	8.8	8.2	10.1	0.703	0.657	1.39	0.86	0.64
Netherlands[b]	34	17.4	12.7	21.0	17.5	13.4	21.0	0.485	0.540	8.4	0.52	1.54
Ruhr Area, DE[c]	15	18.5	15.5	21.1	14.9	13.0	25.1	0.492	0.387	15.8	0.17	1.48
Rome, IT[c]	18	18.5	14.2	27.0	20.5	16.6	21.9	0.598	0.612	−11.0	1.53	2.74

[a] Number of ESCAPE monitoring sites.
[b] Spatial resolution of DM estimates ≤100 × 100 m.
[c] Spatial resolution of DM estimates ≥500 × 500 m.

Overall, agreement between DM and LUR was quite strong for NO_2 in 7 out of 13 study areas, (Pearson $R > 0.70$). Lower agreement was found for PM_{10} and $PM_{2.5}$. Agreement between LUR and DM at the address level was higher for areas where the DM correlated more strongly with the measurements.

4.1. Prediction of measured concentrations at monitoring sites

Gaussian DMs generally predicted the spatial variation of NO_2 at monitoring sites well, reflecting the small-scale variation of this pollutant. On the other hand Eulerian/CFD DMs that modelled average NO_2 concentrations on a coarser spatial scale reflected larger scale variations of urban background within cities. Most models also predicted the concentration levels well (within about 30%), partly due to the incorporation of measured regional background concentrations. Prediction of PM was less effective, similar to LUR models (Beelen et al., 2013; Eeftens et al., 2012b).

As we did not have independent data available for a sufficiently large number of locations in our cities, we cannot make a solid comparison between the two models' predictive ability for the study areas. The correlations between DM and measured concentrations were however lower than for the LUR models (median LOOCV R^2 was 0.80 (0.55), 0.77 (0.34) and 0.61 (0.33) for LUR (DM) NO_2, PM_{10} and $PM_{2.5}$ respectively (Table A.2)). This does not necessarily imply better performance at unmeasured locations. The model R^2 only represents the predictive ability at the monitoring sites and recent studies have documented that the LOOCV R^2 used in LUR studies only partly compensates for the over-fitting. Hold-out validation R^2 has been shown to be potentially 20–40% lower than the model R^2, with larger differences observed for LUR models based on a smaller number of sites (Basagaña et al., 2012; M. Wang et al., 2013).

The RMSE of the comparison between DM and measurements (Table 3) was larger than the RMSE of the comparison between DM and LUR (Table 2). Although based on different locations, this might indicate that both models may have similar errors in explaining measurements.

Several previous studies have compared LUR and DM at monitoring sites. Beelen et al. (2010) found moderate agreement ($R = 0.55$) between LUR and DM estimates for annual average NO_2 concentrations at a 100 × 100 m grid in the Rijnmond area of the Netherlands with the URBIS performing better than the LUR model ($R = 0.77$ vs 0.47) at 18 independent sites. This is likely because the LUR model was developed for the whole of the Netherlands and lacking specific local information for the Rijnmond area. A study in Amsterdam (NL) by Dijkema et al. (2011) compared NO_2 concentrations estimated by 2 LUR models (regional and city specific) against the Dutch CAR dispersion model. All models explained between 50 and 60% of the variance, although CAR overestimated at background and underestimated at traffic monitoring sites. In Vancouver, Canada, Marshall et al. (2008) compared LUR and a 4 × 4 km chemical transport DM (CMAQ) to estimate NO, NO_2, CO and ozone. They found that LUR was better in predicting the small spatial variations at the neighbourhood scale, whereas DM tended to be better in predicting the urban scale variations. Cyrys et al. (2005) also compared LUR and dispersion modelling for NO_2 and $PM_{2.5}$ in Munich, Germany, at 40 monitoring sites and at 1669 addresses. The model estimates correlated well at the 40 monitoring sites and addresses ($R > 0.79$). Gulliver et al. (2011) compared LUR and DM at 52 routine monitoring stations in London (UK) using a grouped jack-knife approach Results showed that LUR ($R^2 = 0.47$) outperformed DM ($R^2 = 0.28$). Most recently Sellier et al. (2014) compared LUR and DM estimates for NO_2 at cohort addresses in Nancy and Poitiers (France) finding a good correlation between the two methods ($R = 0.87$).

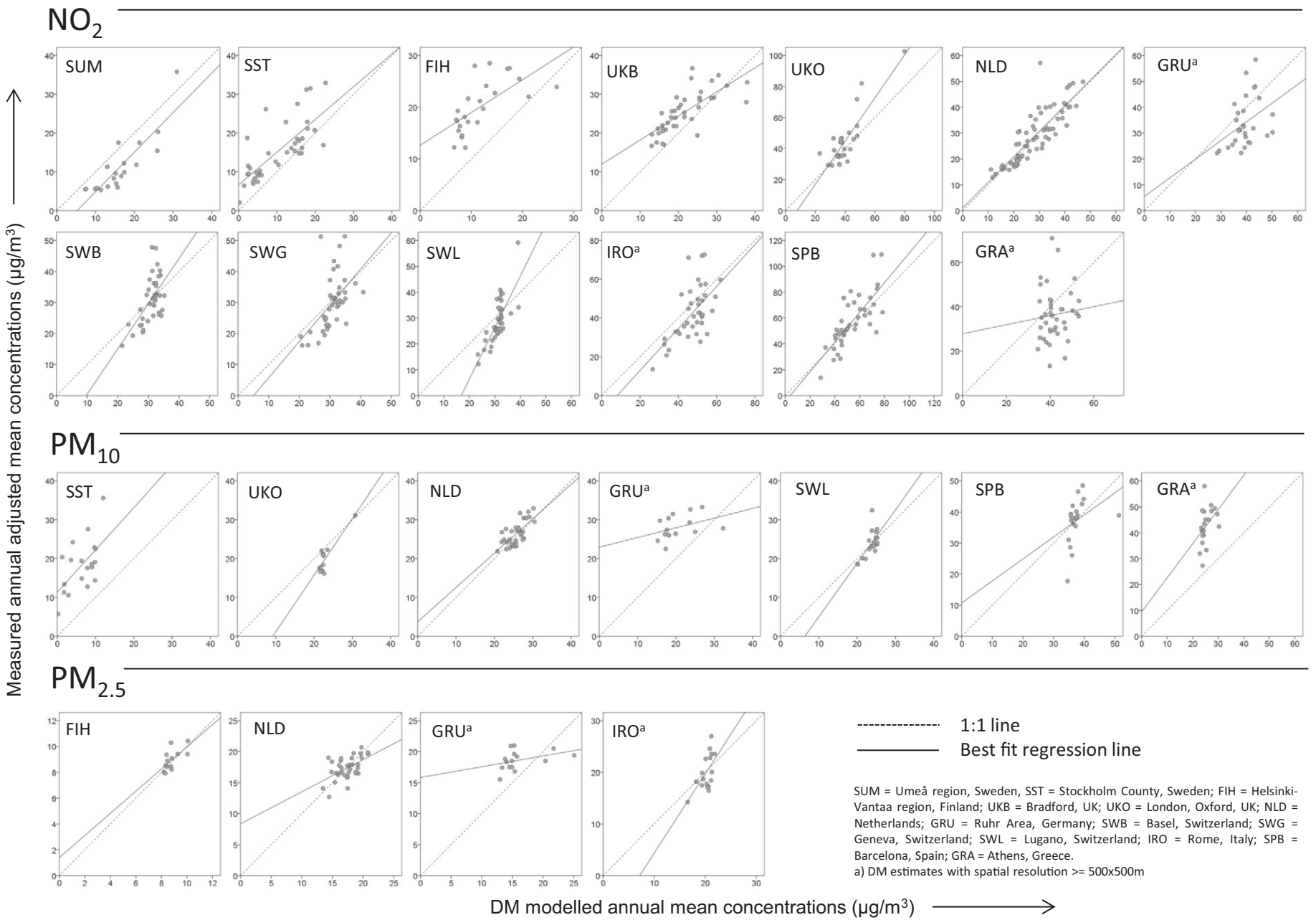


Fig. 2. Comparison of measured annual adjusted concentrations (*x*-axis) against DM (*y*-axis) estimates (µg/m^3) at ESCAPE monitoring sites.

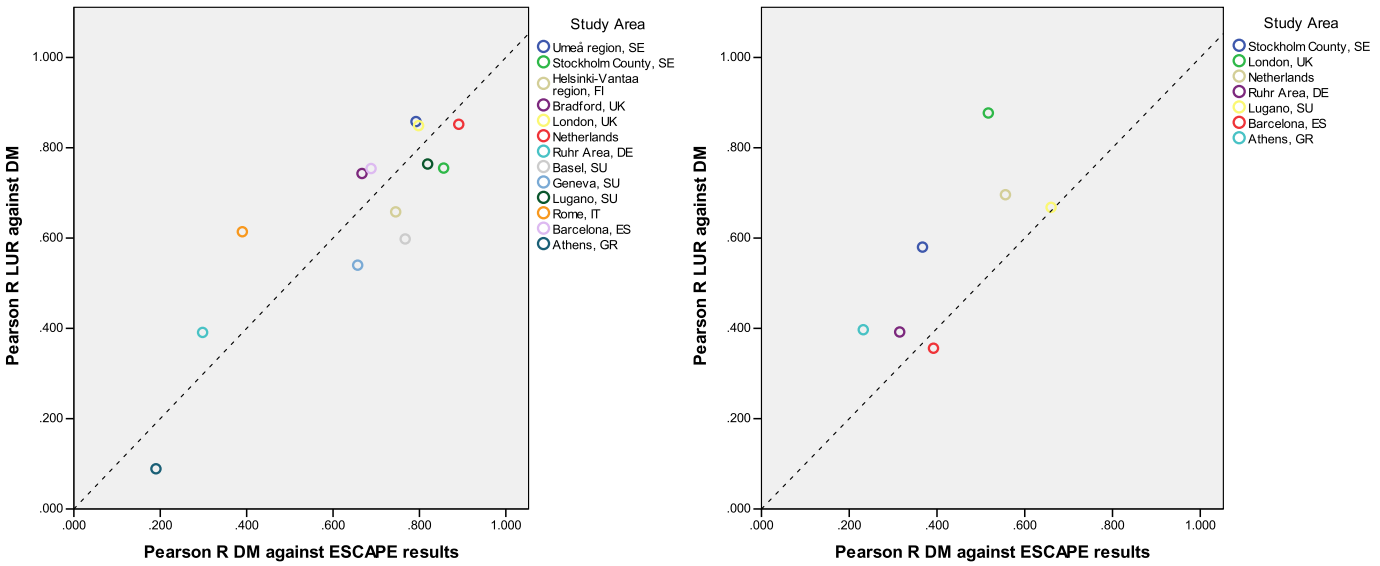


Fig. 3. Scatterplots of Pearson *R*'s between the LUR–DM and DM–ESCAPE comparisons for both NO_2 and PM_{10}.

4.2. Predictions of address level exposure

Despite the different modelling approaches of LUR and DM, the agreement in predicting NO_2 concentrations at cohort addresses was relatively good in most study areas. This is probably due to the importance of traffic affecting small-scale spatial variation of NO_2 in the predominantly urban areas. DMs have been developed extensively for modelling NO_2 traffic sources, and LUR models are most effective for modelling traffic because of the availability of predictor variables such as traffic intensity and distance to major roads. In a recent paper of PM composition, LUR models predicted traffic-related components (Black carbon, Cu, Fe) much better than elements for which non-traffic sources were dominant e.g. Ni, V and S (de Hoogh et al, 2013).

Compared to NO_2, the lower agreement between DM and LUR predictions for PM_{10} is likely due to a combination of random error related to the smaller spatial variation of PM_{10}, the lower predictive power of both models to predict concentrations and the smaller number of monitoring sites available to develop LUR models (20 PM versus 40 NO_2 in most areas). In general, the spatial variation of the measured PM and the predictions by both models was smaller than for NO_2, consistent with observations of a high regional background contribution to fine particle concentrations and a smaller influence of local sources (Eeftens et al., 2012b). In several areas, for both models, the spatial variation of PM was relatively small compared to the prediction errors as reflected by the root mean squared error.

Some of the differences in agreement between the two models at the cohort addresses were caused by the different model types. The Eulerian/CFD models used in the Ruhr, Rome and Athens areas correlated less strongly with LUR estimates than the Gaussian models for both NO_2 and PM_{10}. This is probably in part caused by the coarser resolution used by the Eulerian/CFD models compared to the Gaussian models which therefore better predicting receptor-specific concentrations as modelled in LUR. In epidemiological studies using the Ruhr Area model, the coarser resolution dispersion model was therefore supplemented with distance to major roads to account for the small-scale variation (Hoffmann et al., 2009). Fig. 3 illustrates that the agreement at the cohort addresses depended on how well the DM predicted the measurements at the ESCAPE monitoring sites. In addition to scale of the model, the complexity and size of the urban environment likely affect how well DM and LUR can predict spatial patterns. DMs for Mediterranean cities have some additional challenges such as describing local flows in coastal areas with complex terrain, as well as accounting for the intricacies of boundary layer development. In the case of Athens, emissions have exhibited large variability (inter-annual as well as spatial) over the last couple of years due to the effects of the economic crisis. Therefore, the amount of emission uncertainty involved in the Athens calculations has conceivably played a key role in the DM calculations. Interestingly, the LOOCV for the LUR models was also relatively low in Athens.

DM and LUR models generally explained a lower fraction of measured spatial variation of PM_{10} compared to NO_2 (Table 3 and Tables A.1 and A.2). The continental and regional scale chemical transport models commonly underestimated both the measured $PM_{2.5}$ and PM_{10} concentrations at the ESCAPE monitoring sites, which were designed to capture specifically the variation in traffic-related pollutants and therefore oversampled high traffic sites. Other reasons might include missing or under-estimated source categories (such as wild-land fires, desert dust, biogenic sources, non-exhaust emissions from traffic, shipping, fugitive dust, and sea salt), and by missing or inadequately treated processes in the models (such as the formation of secondary organic aerosols). Because of the urban character of the study areas, all the Gaussian models used measured concentration values at regional background stations; the above mentioned PM modelling deficit for chemical transport models does not therefore influence the predicted results in those cases. However, some dispersion models clearly underpredicted PM_{10} concentrations at the ESCAPE monitoring sites, in case of Stockholm, Ruhr Area and Athens, as can be seen based on the results presented in Fig. 2. For those models predicting average concentrations on a larger scale (i.e. Ruhr Area, Athens) this is a logical consequence of the fact that these models are not designed to predict concentrations at traffic sites. Consistently, the Ruhr Area Eulerian DM model predicted NO_2, PM_{10} and $PM_{2.5}$ better ($R = 0.53, 0.69, 0.68$ respectively) when the traffic sites were excluded.

As previously mentioned, LUR models are less effective for sources other than traffic (de Hoogh et al., 2013). The simple dispersion assumptions in LUR models apply better to traffic emissions than industrial point emissions, emitted at potentially hundreds of metres above ground. In Bradford, our NO_2 LUR model under-predicted at a number of residential addresses which were located in one residential area with a high activity of chemical processes. While this emission source was included in the ADMS-Urban model emission inventory, the LUR model for Bradford did not include an industry variable, because no ESCAPE monitoring sites were located near industrial sources.

A discussion about the Bland–Altman plots and Kappa-coefficients can be found in the Appendix (p. 4).

4.3. Implications for epidemiological studies

The overall high correlation between LUR and (fine scale) DM for NO_2 suggests that similar effects may be obtained if both approaches are applied in epidemiological studies to assess associations with health. However, if predicted concentration ranges differ, the size of the effect estimates may be different. The lower correlation for PM suggests that health effect estimates could be more different when applied in epidemiological studies. It remains important, however, to test directly in epidemiological studies differences in effect estimates related to exposure models. A recent study from Sellier et al. (2014) which applied four different exposure methods, including LUR and DM, to a cohort in Nancy and Poitiers (France), showed some differences in estimated health effect despite moderate to high correlations between NO_2 exposure estimates at the cohort level

The ESCAPE study was specifically designed to investigate health effects of long term air pollution exposure, using standardised LUR as the method of choice. Both LUR and DM are equally equipped to predict long term exposures, but an advantage of DM is that it can more easily deal with different time periods (e.g. hours, days, weeks, years and decades, also in retrospect) by using diagnostic or real-time emission and meteorological data. LUR models estimating daily concentrations have been developed and applied (Gryparis et al., 2014) but their evaluation and use are still limited. LUR model application is further restricted to the time period and geographical area of the monitoring campaign, although some recent studies suggest that LUR models in some circumstances can be transferred both back in time as well as geographically (Gulliver et al., 2013; R. Wang et al., 2013). An advantage of LUR models, however is that exposure estimates can be generated for absorbance, UFP, elemental composition (de Hoogh et al., 2013; Eeftens et al., 2012b) for which few dispersion models are available.

DMs can also be used for evaluating the contributions originating from various sources or source categories at selected locations. A specific strength of DM is its use for retrospective evaluations as well as for scenarios for the future. DM, however is also inherently source specific and as such requires several accurate input datasets like emission inventories, and ideally, pre-processed representative meteorological data, a thorough discussion of which has been presented by Kukkonen et al. (2012). Although the initial development of a LUR model takes some time, the subsequent application to residential addresses is fairly light in terms of computing power and time. DM on the other hand needs a lot more expertise to run and is relatively heavier in data demand and running time.

5. Conclusions

Dispersion model estimates for outdoor NO_2 with high spatial resolution showed, in most countries, high correlation with measured values and with the corresponding land-use regression estimates for cohort addresses. This implies that both methods may be useful for epidemiological studies of small-scale variations of outdoor combustion-related air pollution, typically from road traffic. The agreement for PM levels was considerably lower than for NO_2, probably reflecting smaller spatial variation, less precise source characterization and/or lack of related land use descriptors. The agreement between LUR and dispersion models with lower spatial resolution was reduced. These Eulerian/CFD DMs provide average concentrations in a small area, thus modelling a different aspect of person-specific exposure. The influence of data requirements and whether the methods tend to give different results in epidemiological studies need to be further explored.

Conflict of interest

All authors declare no actual or potential conflict of financial or other interests.

Acknowledgments

The research leading to these results was funded by the European Community's Seventh Framework Program (FP7/2007–2011) projects ESCAPE (grant agreement number: 211250) and TRANSPHORM (ENV.2009.1.2.2.1). We also like to thank José Lao from the Energy & Air Quality Department, Barcelona Regional, Barcelona, Spain for his help in the Barcelona dispersion modelling; Pekka Taimisto and Arto Pennanen for their field work in Helsinki Vantaa region: and Christine McHugh from CERC, UK for her help in dispersion modelling for London. We thank all study participants and the dedicated personnel of the Heinz Nixdorf Recall Study. We gratefully acknowledge the collaboration with K.-H. Jöckel, D. Grönemeyer, R. Seibel, K. Mann, L. Vollbracht, and K. Lauterbach. We thank the North Rhine-Westphalia State Agency for Nature, Environment and Consumer Protection for providing road maps with traffic data and emission data from the reference sites for back-extrapolation. The study was supported by the Heinz Nixdorf Foundation [chairman: M. Nixdorf; former chairman: G. Schmidt (deceased)], the German Ministry of Education and Science, the German Research Foundation (DFG; projects JO-170/8-1, HO 3314/2-1, SI 236/8-1, and SI236/9-1). SAPALDIA: we thank Study directorate: NM Probst Hensch, T Rochat, N Künzli, C Schindler, JM Gaspoz, the Swiss National Science Foundation (grant nos 33CSCO-134276/1, 33CSCO-108796, 3247BO-104283, 3247BO-104288, 3247BO-104284, 3247-065896, 3100-059302, 3200-052720, 3200-042532, 4026-028099, PMPDP3_129021/1, PMPDP3_141671/1), and the Federal Office for Forest, Environment and Landscape (10.0022.PJ/J112-0392). For full study team: see Appendix A (p. 2).

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.envint.2014.08.011.

References

Basagaña X, Rivera M, Aguilera I, Agis D, Bouso L, Elosua R, et al. Effect of the number of measurement sites on land use regression models in estimating local air pollution. Atmos Environ 2012;54:634–42.

Beelen R, Voogt M, Duyzer J, Zandveld P, Hoek G. Comparison of the performances of land use regression modelling and dispersion modelling in estimating small-scale variations in long-term air pollution concentrations in a Dutch urban area. Atmos Environ 2010;44:4614–21.

Beelen R, Hoek G, Vienneau D, Eeftens M, Dimakopoulou K, Pedeli X, et al. Development of NO_2 and NOx land use regression models for estimating air pollution exposure in 36 study areas in Europe—the ESCAPE project. Atmos Environ 2013;72:10–23.

Bellander T, Berglind N, Gustavsson P, Jonson T, Nyberg F, Pershagen G, et al. Using geographic information systems to assess individual historical exposure to air pollution from traffic and house heating in Stockholm. Environ Health Perspect 2001;109:633–9.

Brauer M, Lencar C, Tamburic L, Koehoorn M, Demers P, Karr C. A cohort study of traffic-related air pollution impacts on birth outcomes. Environ Health Perspect 2008;116(5):680–6.

Briggs D, Collins S, Elliot P, Fischer P, Kingham S, Lebret E, et al. Mapping urban air pollution using GIS: a regression-based approach. Int J Geogr Inf Sci 1997;11:699–718.

Carruthers DJ, Edmunds HA, Lester AE, McHugh CA, Singles RA. Use and validation of ADMS-Urban in contrasting urban and industrial environments. Int J Environ Pollut 2000;14:1–6.

Cyrys J, Hochadel M, Gehring U, Hoek G, Diegmann V, Brunekreef B, et al. GIS-based estimation of exposure to particulate matter and NO_2 in an urban area: stochastic versus dispersion modeling. Environ Health Perspect 2005;113:987–92.

Cyrys J, Eeftens M, Heinrich J, Ampe C, Armengaud A, Beelen R, et al. Variation of NO_2 and NOx concentrations between and within 36 European study areas: results from the ESCAPE study. Atmos Environ 2012;62:374–90.

de Hoogh K, Wang M, Adam M, Badaloni C, Beelen R, Birk M, et al. Development of land use regression models for particle composition in 20 study areas in Europe. Environ Sci Technol 2013;47:5778–86.

Dijkema MBA, Gehring U, van Strien RT, van der Zee SC, Fischer P, Hoek G, et al. A comparison of different approaches to estimate small scale spatial variation in outdoor NO_2 concentrations. Environ Health Perspect 2011;119(5):670–5.

Dockery DW, Pope 3rd CA, Xu X, Spengler JD, Ware JH, Fay ME, et al. An association between air pollution and mortality in six US cities. N Engl J Med 1993;329:1753–9.

Eeftens M, Tsai MY, Ampe C, Anwander B, Beelen R, Bellander T, et al. Spatial variation of PM2.5, PM10, PM2.5 absorbance and PM coarse concentrations between and within

20 European study areas and the relationship with NO_2—results of the ESCAPE project. Atmos Environ 2012a;62:303–17.
Eeftens M, Beelen R, de Hoogh K, Bellander T, Cesaroni G, Cirach M, et al. Development of Land Use Regression models for PM(2.5), PM(2.5) absorbance, PM(10) and PM(coarse) in 20 European study areas; results of the ESCAPE project. Environ Sci Technol 2012b; 46:1119511205.
Finardi S, D'Allura A, Maddalena M, Silibello C, Radice P, Morelli M, et al. An air quality forecast system for Rome metropolitan area: First evaluation and identification of critical issues. In: Proceedings of the 7th International Conference on Air Quality Science and Application, 24-27 March 2009, Istanbul, Turkey; 2009.
Gariazzo C, Silibello C, Finardi S, Radice P, Piersanti A, Calori G, et al. A gas/aerosol air pollutants study over the urban area of Rome using a comprehensive chemical transport model. Atmos Environ 2007;41:7286–303.
Gryparis A, Dimakopoulou K, Pedeli X, Katsouyanni K. Spatio-temporal semiparametric models for NO_2 and PM10 concentration levels in Athens, Greece. Sci Total Environ 2014;479–480:21–30.
Gulliver J, de Hoogh K, Fecht D, Vienneau D, Briggs D. Comparative assessment of GIS-based methods and metrics for estimating long-term exposures to air pollution. Atmos Environ 2011;45(39):7072–80.
Gulliver J, de Hoogh K, Hansell A, Vienneau D. Development and back-extrapolation of NO_2 land use regression models for historic exposure assessment in Great Britain. Environ Sci Technol 2013;47:7804–11.
HEI Panel on the Health Effects of Traffic-Related Air Pollution. Traffic-related air pollution: a critical review of the literature on emissions, exposure, and health effects. HEI Special Report 17. Boston: Health Effects Institute; 2010 [Available: http://pubs.healtheffects.org/getfile.php?u=553 [accessed 3 January 2014]].
Hennig F, Fuks K, Moebus S, Weinmayr G, Memmesheimer M, Jakobs H, et al. Association between source-specific particulate matter air pollution and hs-CRP: local traffic and industrial emissions. Environ Health Perspect 2014. http://dx.doi.org/10.1289/ehp.1307081.
Hoek G, Beelen R, de Hoogh K, Vienneau D, Gulliver J, Fischer P, et al. A review of land-use regression models to assess spatial variation of outdoor air pollution. Atmos Environ 2008;42:7561–78.
Hoffmann B, Moebus S, Dragano N, Stang A, Möhlenkamp S, Schmermund A, et al. Chronic residential exposure to particulate matter air pollution and systemic inflammatory markers. Environ Health Perspect 2009;117:1302–8. http://dx.doi.org/10.1289/ehp.0800362.
Holmes NS, Morawska L. A review of dispersion modelling and its application to the dispersion of particles: an overview of different dispersion models available. Atmos Environ 2006;40:5902–28.
Karppinen A, Kukkonen J, Elolähde T, Konttinen M, Koskentalo T, Rantakrans E. A modelling system for predicting urban air pollution, Model description and applications in the Helsinki metropolitan area. Atmos Environ 2000;34(22):3723–33.
Kukkonen J, Härkönen J, Walden J, Karppinen A, Lusa K. Validation of the dispersion model CAR-FMI against measurements near a major road. Int J Environ Pollut 2001; 16:1–6.
Kukkonen J, Olsson T, Schultz DM, Baklanov A, Klein T, Miranda AI, et al. A review of operational, regional-scale, chemical weather forecasting models in Europe. Atmos Chem Phys 2012;12:1–87. http://dx.doi.org/10.5194/acp-12-1-2012.
Künzli N, Jerrett M, Mack WJ, Beckerman B, LaBree L, Gilliland F, et al. Ambient air pollution and atherosclerosis in Los Angeles. Environ Health Perspect 2005;113(2):201–6.
Marshall JD, Nethery E, Brauer M. Within-urban variability in ambient air pollution: comparison of estimation methods. Atmos Environ 2008;42:1359–69.
Memmesheimer M, Friese E, Ebel A, Jakobs HJ, Feldmann H, Kessler C, et al. Long-term simulations of particulate matter in Europe on different scales using sequential nesting of a regional model. Int J EnvirOn Pollut 2004;22:108–32.
Moussiopoulos N, Douros I, Tsegas G, Kleanthous S, Chourdakis E. An Air Quality Management System for Policy Support in Cyprus. Advances in Meteorology 2012;Vol 2012. http://dx.doi.org/10.1155/2012/959280. [Article ID 959280].
Özkaynak H, Baxter LK, Dionisio KL, Burke J. Air pollution exposure prediction approaches used in air pollution epidemiology studies. J Expo Sci Environ Epidemiol 2013;23: 566–72.
SAEFL. Modelling of PM10 and PM2.5 ambient concentrations in Switzerland 2000 and 2010. Technical report. Swiss agency for the environment, forests and landscape, environmental documentation No. 169; 2003.
Sellier Y, Galineau J, Hulin A, Caini F, Marquis N, Navel V, et al. Health effects of ambient air pollution: do different methods for estimating exposure lead to different results? Environ Int 2014;66:165–73.
SMHI. Indic Airviro. Swedish Meteorological and Hydrological Institute. Norrkoping, Sweden: Technical Description of the Dispersion Models Swedish Meteorological and Hydrological Institute; 1993.
Velders GJM, Aben JMM, Geilenkirchen GP, den Hollander HA, Jimmink BA, van der Swaluw E, et al. Grootschalige concentratie- en depositiekaarten Nederland: Rapportage. RIVM Report 680362003; 2013 [in Dutch].
Wang M, Beelen R, Basagana X, Becker T, Cesaroni G, de Hoogh K, et al. Evaluation of land use regression models for NO_2 and particulate matter in 20 European study areas: the ESCAPE project. Environ Sci Technol 2013a;47(9):4357–64.
Wang R, Henderson SB, Sbihi H, Allen RW, Brauer M. Temporal stability of land use regression models for traffic-related air pollution. Atmos Environ 2013b;64:312–9.
Wesseling JP, Sauter FJ. Calibration of the CAR II Program with Measurements of RIVM. Bilthoven, The Netherlands: RIVM; 2007 [Report 680705004 www.rivm.nl (in Dutch)].
Wesseling JP, Visser GT. An Intercomparison of the TNO Traffic Model, Field Data and Wind Tunnel Measurements TNO, Utrecht, The Netherlands. Report 2003/207CAR model; 2003.
WHO. Review of evidence on health aspects of air pollution—REVIHAAP Project: final technical report. Bonn: The WHO European Centre for Environment and Health; 2013.
Wu J, Wilhelm M, Chung J, Ritz B. Comparing exposure assessment methods for traffic-related air pollution in an adverse pregnancy outcome study. Environ Res 2011; 111:685–92.

Can Dispersion modeling of air pollution be improved by land use regression?

Michal Korek [a], Christer Johansson [b], Nina Svensson [b], Tomas Lind [a, c], Rob Beelen [d, e], Gerard Hoek [e], Göran Pershagen [a], Tom Bellander [a, c]

[a] Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden

[b] Environment and Health Administration, Stockholm City, Sweden and Department of Environmental Science and Analytical Chemistry, Stockholm University, Sweden

[c] Centre for Occupational and Environmental Medicine, Stockholm County Council, Stockholm, Sweden

[d] National Institute for Public Health and the Environment (RIVM),Utrecht, the Netherlands

[e] Institute for Risk Assessment Sciences, Utrecht University, Utrecht, the Netherlands;

Corresponding author:

Michal Korek

Karolinska Institutet, Institute of Environmental Medicine

Nobels väg 13 Box 210, SE-171 77 Stockholm, Sweden

E-mail: michal.korek@ki.se

Phone: +46-8-524 87457

Fax: +46-8-304571

Running title: **Combining land use regression and dispersion modeling**

Conflicts of Interest and Source of Funding

This project was supported financially the Swedish Environmental Protection Agency, the European Commission and the Swedish Council for Healthy, Working Life and Welfare (FORTE). The authors declare that they have no competing interests.

Abstract

Both dispersion modeling (DM) and land use regression modeling (LUR) are often used for assessment of long-term air pollution exposure in epidemiological studies, but seldom in combination. We developed a hybrid DM-LUR model using 93 bi-weekly observations of NO_x at 31 sites in greater Stockholm (Sweden). The DM was based on spatially resolved topographic, physiographic and emission data, and hourly meteorological data from a diagnostic wind model. Other data were from land-use, meteorology and routine monitoring of NO_x. We built a linear regression model for NO_x, using stepwise forward selection of covariates. The resulting model predicted observed NO_x (R^2 = 0.89) better than the DM without covariates (R^2= 0.68, p-interaction <0.001), and with minimal apparent bias. The model included (in descending order of importance) DM, traffic intensity on the nearest street, population density within 100 m radius, global radiation (direct sunlight plus diffuse or scattered light), and urban contribution to NO_x levels (routine urban NO_x less routine rural NO_x). Our results indicate that there is a potential for improving estimates of air pollutant concentrations based on DM, by incorporating further spatial characteristics of the immediate surroundings, possibly accounting for imperfections in the emission data.

Introduction

The number of studies on the relation between both long and short-term air pollution exposure from road traffic and various adverse health effects continues to grow (1, 2). Over time, developments in computer technology and exposure modeling have advanced modern exposure assessment from the use of large spatial scale exposure estimates based on a few continuous monitoring sites (3), to exposure estimates on much finer "local" scales, describing intra-urban, cross-sectional (4) as well as temporally resolved exposure variation (5, 6). The need of better modeling techniques is underlined by documented large intra-urban variations found in monitoring studies (7) and adequate exposure assessment is a crucial part of environmental epidemiology with direct influence on study validity (8).

Meteorological dispersion modeling (DM) and Land use regression modeling (LUR) are alternative methods describing small scale variations in air pollution levels and both have been documented to estimate urban outdoor concentrations of NO_2 and NO_x well (9). DM's calculate the geographic distribution of air pollutants by combining data on emission (point, line and area sources), the geophysical properties of the study area, and meteorological conditions. In contrast to LUR, the DM can calculate concentrations at assigned locations at any time scale while adjusting for the interacting spatio-temporal effects of sources and meteorology.

Land Use Regression is a multiple linear regression technique used to establish the relation between spatially dispersed monitored air pollutant levels in a study area, and the area specific land use data (predictors) that best explain the monitored exposure variability. In a second step, exposure at any given site is calculated by applying the established regression formula on predictor data specific for that site (10, 11). LUR is mostly used to calculate spatial exposure relations for longer time periods i.e. months or years but in contrast to DM, the LUR design does not require an emission inventory.

The exposure concentrations based on DM and LUR models generally correlate well at the same address sites but some disparity in between modeled levels can be found (9). Model selection has been demonstrated to somewhat impact the health effects estimates in the same epidemiological study (12). A proposed way to further minimize errors in exposure estimates is to combine existing models into so called “hybrid models” (10).

It has been shown that LUR modeling may be enhanced by the use of time varying traffic data and meteorological data such as temperature, relative humidity and wind speed (13, 14). Meteorological data has also been used to define the form of buffer areas as to capture the effect of an average wind direction (15) as well as wind speed (16, 17). These studies did however not explore the effect of other meteorological factors such as mixing height. Furthermore, unlike the estimates from DM, the constructed LUR predictors did not describe the effect of spatio-temporally related variations in both meteorology and traffic. Such effects were captured in a LUR- dispersion hybrid model developed for NO_2 and NO_x (18). In this hybrid model, DM estimates were included in LUR modeling as one predictor variable together with data on road, land use, population and distance to coast. This hybrid model

was created to capture near road exposure gradients in urban and sub-urban areas and added further promise to a hybrid LUR-DM model by explaining more observed variability compared to using only DM or LUR. However, the monitors were positioned as to exclude infrastructural influences on meteorology by high-rise buildings and street canyons, therefore omitting exposure scenarios for living conditions common to many people in city centers and similar areas. At present, the benefits of a hybrid DM-LUR model compared to DM and LUR separately have been little investigated in a metropolitan setting. Additionally, the potential to use LUR to investigate areas of improvement in DM has not been explored.

The aim of this study is to develop a hybrid spatio-temporal model for outdoor NO_x levels in an large urban area, using NO_x estimates from dispersion modelling as well as land use variables, meteorology and fixed monitoring data while adjusting for street canyon effects. We evaluated whether the hybrid model predicted better than either dispersion modelling or land-use regression modelling developed separately, and could be used to identify potential improvement of both dispersion and land-use regression modelling.

Material and Methods

Dispersion models

Two different dispersion models have been applied to the greater Stockholm area (35 by 35 km). Both used a detailed emission inventory for traffic sources, with regularly updated traffic flows reported by the municipalities. This is a local road network database covering 90% of all the roads in Stockholm County including information on traffic intensities for every road segment above 500 vehicles/24hrs. The amount of heavy traffic per road segment was approximated to be 4-10 % of the total traffic depending on road type although up to 90% at some bus routes. The road network was digitalized by each municipality separately for 1993-2008. Emission factors were calculated for street segments as the emission per vehicle and distance (NO_x/vkm), based on the HBEFA model (http://www.hbefa.net/e/index.html). The number of vehicles at a street segment was adjusted for average differences according to month of year, type of day, hour of day, speed limit. The inventory also included other sources (19) although the dominant source of NO_x in Stockholm County was traffic (20) for which modeling was done in this study. Both types of DM estimates were complemented with the corresponding 2-week average or rural NO_x from a routine monitoring station. Both models were used to calculate one hour average NO_x concentrations for the EXCAPE monitoring sites, averaged to correspond to the actual 2-w samplings at each site.

A multisource Gaussian dispersion model was used to calculate urban background concentrations and non- canyon traffic sites at a 500 m spatial resolution. The model is part

of the Airviro Air Quality Management System (SMHI, Norrköping, Sweden http//airviro.smhi.se). Meteorological data was obtained from a 50 m high mast in a suburban district (Högdalen) in southern Stockholm, and these data were input to a diagnostic wind model in order to calculate a wind field over the whole model domain. It should be noted that the modelled values represents the contributions to the concentrations from all traffic sources to a 500 m by 500 m area, whereas the monitored values are single points within these areas (individual streets or building effects are not resolved by this model). Buildings, park trees etc. are parameterized as rough surfaces, which increases turbulent dilution.

The Gaussian model has been used extensively in epidemiological studies describing both short-term and long-term exposure concentrations on address level in Stockholm County (21-23). Furthermore, the model has presented a high correlation with monitored annual exposure R^2 = 0.74-0.80 over several years (1998-2005) (24).

To describe the NO_x concentrations at street canyon traffic sites, the SIMAIR-road model was used(25) The domain of the SIMAIR-road model covered greater Stockholm although not the municipalities of Sundbyberg and Solna. This model calculates concentrations along individual streets with buildings on both sides of the street. Meteorological data were supplied to the SIMAIR-road model from a system called MESAN (MESoscale Analysis), which makes use of all available measurement stations and radar and satellites combined with a background field forecast (25).

Spatially distributed measurements

The spatially distributed observations of NO_x were from the multinational project "ESCAPE" (European Study of Cohorts for Air Pollution Effects). Within the ESCAPE project a coordinated campaign for monitoring of study area specific levels of NO_x and other pollutants was organized in several European countries. Based on these measurements area-specific LUR models were later developed. The details of the measurements have been described elsewhere (7). Briefly, the monitoring campaign in Stockholm County was conducted from 01-12-2008 until 11-07-2009. The spatial variation of NO_x was measured at 40 monitoring sites distributed to capture traffic-related exposure scenarios at home addresses in Stockholm County. Site specific measurements were obtained for three bi-weekly periods. The choice of periods aimed to cover seasonal variations, and up to 10 sites were monitored simultaneously. NO_x was measured using Ogawa diffusion badges (7). Geographical coordinates were attributed to each monitoring site by the Swedish mapping, cadastral and land registration authority.

For the purpose of this study, the following categories and inclusion criteria for the ESCAPE monitoring sites were successively applied:

1. The site should be within the spatial domains of our DM models. Six sites were situated outside the area of the models. One of these is a rural background site (at Norr Malma, ca 70 km north of Stockholm) and was used as an estimate of the non-urban source contribution to the modelled concentrations.

2. To avoid a strong influence from single sites with very high traffic volumes, only sites with < 100.000 vehicles/day on the nearest street were included (one site was excluded).
3. Traffic sites should be located close to the road (≤ 15m) and if situated on a building, ≤ 10m above street height. The façade should also face a street with ≥ 10000 veh/day
    a. For traffic sites in street canyons the SIMAIR model was used. These site had buildings on both sites of the street (one street canyon site was outside the domain of the SIMAIR model and therefor excluded).
    b. For traffic sites in open street configurations the Gauss urban background model was used.
4. Urban background sites were estimated using the Gaussian model. These sites were ≥15 meters from the nearest street with traffic intensities of ≥5000 veh/day. One urban site did not meet the criteria 3 or 4 and was therefore excluded.

The final dataset thus included 31 sites: 11 traffic, 16 urban and 4 rural sites. Each monitored 2 week average was considered as one observation, yielding (3*31) 93 observations.

We additionally collected continuous NO_x data from three stationary routine monitoring stations (STAT) representing the regional background (Norr Malma), urban background (Torkel Knutssonsgatan) and traffic (Hornsgatan, located in a street canyon). The measurements were provided by the Environment and Health Administration of Stockholm (www.slb.nu) and covered the same dates as the monitor-specific 2-week periods.

The NO_x concentrations observed at Norr Malma (regional background) were added to the calculated concentrations from DMs since the models only considered the contributions from the urban traffic sources. A 14-day period "delta urban NO_x" predictor was calculated as the difference between Norr Malma (regional) and Torkel Knutsson (urban) and a "delta traffic NO_x" predictor was calculated as the difference between Torkel Knutsson (urban) and Hornsgatan (traffic). The last two variables were offered as predictors in LUR modeling. Descriptive data for the stationary monitoring and STAT predictors can be found in the (Supplemental Table 1).

Additional spatial and temporal data

The extraction and definition of land use data has been described elsewhere (26). Briefly, based on coordinates for the study-specific monitoring sites, predictor data were collected in a geographic information system (ArcMap 9). Predictors based on land use and population data were created in the form of buffer zones around monitoring sites while predictors based on traffic data were also based on distance from the site to the road.

The traffic variables for Stockholm County were primarily based on the road network provided by the Eastern Sweden Air Quality Management Association (www.slb.nu/lvf). i.e. the same database used for the DM. For LUR, predictors were calculated as the inverse distance and the inverse distance squared to nearest road and nearest major road (m^{-1},m^{-2}). The total length of roads (m) based on all roads and major roads only, were calculated in buffers of 25, 50, 100, 300, 500 and 1000 meter radii. A major road was defined as a road with >5000 veh/24h. The buffer sizes were selected as to describe near sources and sources

of urban background levels (27). For the same buffer sizes, the "traffic load" on the nearest roads was calculated as the sum of the lengths of road segments multiplied by the traffic intensity attributed to each segment. The same calculations were then done using heavy traffic intensities only.

To adjust for missing roads, particularly for rural locations we complemented the road network with the Euro streets digital road network version 3.1. This road network is based on the TeleAtlas MultiNet TM from the year 2008. The MultiNet TM road network covers roads in Stockholm County with traffic intensities of less than 500 veh/day but lacks information on traffic intensity. The additional road information allowed us to better estimate the distance from all monitors to the nearest road. Furthermore, a proxy of 500 vehicles/24h and 0 heavy vehicles/24h was attributed to these roads. This information was only used for predictors based on distance to road.

Land use data were extracted from the CORINE (Coordination and Information on the Environmental programme) land cover data 2000 (CLC2000), governed by the European Environmental Agency. The data was originally based on images from the Landsat-7 satellite although used in vector form in this study. The minimum mapping unit (size of vector) was 25ha corresponding e.g. to a 500x500m square. Final predictor variables covered urban green, semi-natural areas, forest areas, high density residential land, low density residential land, industry and ports. Each predictor was based on the amount of surface area in buffer zones with radii of 100, 300, 500, 1000 and 5000 meter. Population density was modeled as the number of individuals within the buffer zones using a 100x 100m grid map with counts of citizens attached to each grid for the year 2005. The amount of surface water within buffer

zones was registered using a terrain map from the Swedish mapping, cadastral and land registration authority for the year 2005 with an accuracy of +/-10 meters.

Meteorological predictors (MT) used in the LUR modelling included, temperature, relative humidity, global radiation and wind vectors. Wind vectors were computed as eastern and northern wind direction components together with a separate variable for wind speed. All meteorological measurements were obtained at a stationary monitoring station positioned at rooftop level in central Stockholm (Torkel Knutssonsgatan). The MT predictors represented temporally resolved 2-week averages covering the same periods as the monitored NO_x concentrations. A descriptive table of the MT predictors can be found in the (Supplemental Table 1).

Regression model development

Two LUR models were developed, one standard LUR model "LUR" based on the above mentioned spatially land-use variables. Another LUR model, denoted "LUR+MT+STAT", included also the temporally defined MT and STAT data. For both models, first, a univariate ordinary least squares regression model was developed for each predictor. The model best explaining the observed variance (R^2) was kept. To this model all the remaining predictors were added separately using a repeated stepwise forward regression method and in each turn, the predictor adding the most additional explained variance was included. Predictors entering the model had to add at least 1% explained variance while having a coefficient with the correct predefined direction of effect. Furthermore, the new predictor should not influence the direction of effect of other predictors (28). Theoretically, we allowed for more

than one buffer size of the same predictor to enter the model although these predictors would be re-written into “doughnut-shaped” buffer forms mutually excluding each other. From the final model, predictors with a p-value larger than 0.1 were removed.

The hybrid model (DM+LUR+MT+STAT) was based on the same predictor data and modeling technique as the LUR+MT+STAT model with the addition of DM exposure estimates offered as a potential predictor. As an intermediate step a DM+MT+STAT model was developed, similar to the above LUR+MT+STAT model.

The variance in the regression models was estimated with a cluster robust method (29), to avoid the underestimation of variance because of repeated sampling. Potential multicolinearity between the predictors in the final models was investigated using the Variance Inflation factor test (VIF). If a predictor variable exceeded a VIF value of 3, the influence on model performance was investigated by removal of this predictor.

The performance of all models was assessed and compared by model-specific proportion of explained variability (R^2), the root mean square error (RMSE) and the best visual fit. To estimate model robustness, leave one out cross-validation (LOOCV) was applied on all the models (7) Differences in model performance were also tested for statistical significance using the Wald test. To assess the degree of association between monitored NO_x and predictors separately in the final hybrid model, partial R^2 were calculated.

Results

During the ESCAPE monitoring period a high temporal variability was found in NO_x concentrations observed from routine monitoring at street level, with a somewhat lower concentration range from late spring to early fall (April-September). The routine urban background NO_x followed a similar pattern while the rural background levels were stable (Figure 1). As described by the black bars in Figure 1 the monitoring campaign had an overall good coverage of the variability in both meteorology and NO_x exposure during the campaign year.

Comparing the spatially distributed measured NO_x values with the dispersion modelling shows that DM performed well with an R^2 of 0.68 and a RMSE of 12 $\mu g/m^3$ (Table1). At NO_x levels below about 30$\mu g/m^3$ there was a tendency of overestimation and for higher levels a tendency of underestimation (Figure 2). The basic LUR model (without any temporal variables) explained 58 % of the variability within the measured NO_x (R^2 = 0.58, RMSE 13.9 $\mu g/m^3$) (Table 1). Similar to the DM, lower levels were overestimated and higher were underestimated (Figure 2). The predictors included in this model were population density within a radius of 300m, and the total number of vehicles per day at the nearest street (Table 1).

Both DM and LUR explained measured values significantly better when also temporal variables were included. In the DM+MT+STAT model, the DM estimates were complemented with, the meteorological predictor global radiation and delta urban NO_x (urban-rural). Global radiation levels had a clear annual pattern with peak levels in the end of April until

September (Figure 1). The inclusion of these predictors increased model performance (R^2 = 0.82, RMSE=9.14 (Table 1, Figure 3). The LUR+MT+STAT model included similar variables: the delta urban NO_x (urban-rural), traffic intensity on the nearest street, population density within a radius of 100m and global radiation. The model performance was also similar to the DM+MT+STAT model, (R^2= 0.80, RMSE = 9.70, (Table 1, Figure 3).

We found the hybrid model (DM+LUR+MT+STAT) to perform better than any other model. The model captured 89% of the variance in the monitored concentrations (R^2 = 0.89) and had the lowest model RMSE value (7.14). Furthermore, the predicted NO_x estimates where more accurate across the whole exposure range (Table 1, Figure 3). The Wald test indicated that the difference in performance by the hybrid model compared to the DM+MT+STAT and LUR+MT+STAT was significant ($p< 0.01$.). The Hybrid model included the following predictors: the DM estimates, traffic intensity on the nearest street, population density within 100m, global radiation and delta urban NO_x (urban –rural)background (Table 1). Except for the DM estimate, traffic intensity on the nearest street was found to be the most correlated predictor according to the partial R^2 (Supplemental Table 2).

The variance inflation factor (VIF) test for multicolinearity did not indicate any strong correlation (VIF > 3) between predictors in any of the models, while a strong correlation could have led to unreliable regression coefficients. In the leave one out cross-validation analysis, the models explained between 2-3% less variance indicating good model robustness (Table 1).

Discussion

We demonstrated the possibility to improve Dispersion modeling using a LUR framework and to evaluate areas of improvement in the DM. As expected, we found the final Hybrid model to perform better than the DM and LUR models separately.

A similar result was found in another study (18) where the performance of a LUR-DM Hybrid was compared to LUR and DM models at monitoring sites describing bi-weekly near road exposure gradients. The performance of the best Hybrid model for NO_x in that study was somewhat lower compared to our model (R^2 = 0.71) but some differences between the Hybrid models could be found. The Dispersion model output was retrieved from a simplified version of the Caline3 model and explained 26 % of the variance in the data (compared to 68% in our study). Additionally, predictors based on meteorological data were also included as predictor variables in our model.

Most of the variance (spatial and temporal) in the monitored data was covered by the DM, but the model had a tendency to overestimate the lowest NO_x levels while underestimating high concentrations. A similar performance was found for the Dutch dispersion modeling tool "CAR" when used to model annual small scale variations in NO_2 in the city of Amsterdam (30). This model mostly underestimated the local traffic contribution and displayed the least accuracy for the highest concentrations. This was in part explained by the authors as a difficulty to model complicated traffic situations such as for example often congested heavily trafficked roads. Several earlier studies have reported that real-world emissions are underestimated particularly for some vehicles (31). As clearly seen from the graph, our

hybrid model performed well at all concentration ranges, possibly due to the incorporation of very local land-use characteristics.

The addition of global radiation was found to be important. Global radiation is a measure of the incoming direct sunlight as well as diffuse and scattered light. In an earlier validation of the Gaussian dispersion models used here for urban sites (and three open space traffic sites), discrepancies between monitored and modeled daily averages of NO_x were proposed to relate to deficiencies in the model's parameterization of mixing processes in the planetary boundary layer. For example, if the mixing is underestimated in the afternoon the dilution of the emissions will be underestimated, resulting in high concentration of pollutants (32). Conceptually, the effect of global radiation should have a smaller impact on NO_x levels at street canyon sites, where local vehicle emissions and wind speed and direction are more influential. We have too few observations to formally test this, but the model with global radiation seemed to provide a better fit at low levels. The performance of the LUR model (R^2= 0.58) compared to the DM (R^2 = 68) was good considering that only spatially related predictors were used to explain a 2-week average. However, we demonstrated that the model could be improved substantially by including also time varying meteorological and routine monitoring data. In earlier LUR models meteorological components such as temperature, relative humidity, wind speed, wind vectors and cloud cover were used on different time scales, but not global radiation (13-17, 33).

The traffic and population based LUR predictors used in our hybrid model are commonly used in LUR modeling (34). The traffic predictor "traffic intensity on the nearest street "may reflect the difficulties to account for the influence of very near traffic lack or indicate an underprediction of vehicle emissions.

The predictor “population density” has been described as a marker of air pollution variability related to the differences in urban and rural living environments including sources of traffic and home heating (35). In Stockholm, this may act as a marker of the amount of traffic in the neighborhood, but could perhaps also reflect other aspects of urbanicity as e. g. street configuration and use of off-road machinery. Finally, a description of the temporal NO_x variability on an urban scale “delta NO_x (urban- rural)” was included in the Hybrid model. This may indicate that the Dispersion modeled 2-week mean levels did not fully cover the temporal changes related to the NO_x concentrations in the City. The traffic-related NO_x contribution is estimated to be about 60% in this region (36) but the measured urban background (less rural background) was only weakly correlated to the DM estimates that included traffic-related NO_x levels and the rural contribution (R^2 = 0.10, data not shown). Therefore it is likely that the time variations in urban background (less rural background) represents influence of non-traffic NO_x sources as energy production, off-road machinery and shipping. Our study indicates that future improvements in the DM could decrease exposure misclassification both for long and -short term exposure assessment.

This type of Hybrid model development needs dispersion modeling, routine monitoring, land use data and a campaign of spatially distributed measurements, and is therefore data and computationally intense. Ideally, spatial models trained on observed air pollution data should be evaluated with completely separate sets of observed air pollution measurement data. We did not have monitoring data for this kind of comparison, and therefore used the leave one out cross-validation technique. The LOOCV method has been suggested to give overly optimistic R^2 statistics compared to validations on external datasets for the same models, but the gap between validation and LOOCV R^2 has been reported to be modest

when using 80 or more observations (37). One mechanism that may contribute to an inflated R^2 is when a good model fit is obtained by including many predictors in relation to the number of observations for which variance should be explained by the model. It has been suggested that linear regression models may suffer from such overfitting when the number of observations is less than 2 to 10 times the number of predictors (38-40). Our models were developed on 93 observations from 31 different sites, during 12 different time periods (up to 10 sites could be monitored in parallel). In the final model 3 spatial and 3 temporal predictors (DM counted in both categories) were included, why there might be some overfitting for temporal variables but probably not for spatial variables.

Conclusions

A hybrid spatio-temporal model, combining dispersion modelling (DM), local land use, and centrally monitored pollutants and meteorology, explained variation of two-week average NO_x concentrations within a metropolitan area significantly better than DM alone. This indicates that there is a potential for improving long-term estimates of air pollutant concentrations based on DM by incorporating further spatial characteristics of the immediate surroundings. In addition our results suggest that the inclusion of data from routine air pollution monitoring and meteorology may improve both DM and land use regression in spatially resolved short-term assessment.

References

1. WHO. Review of evidence on health aspects of air pollution – REVIHAAP project: final technical report. The WHO European Centre for Environment and Health, Bonn (2013)[Available: http://wwweurowhoint/en/health-topics/environment-and-health/air-quality/publications/2013/review-of-evidence-on-health-aspects-of-air-pollution-revihaap-project-final-technical-report [accessed 2 September 2015]]. 2013.
2. Brook RD. Cardiovascular effects of air pollution. Clin Sci (Lond). 2008;115(6):175-87. doi: 10.1042/CS20070444. PubMed PMID: 18691154.
3. Dockery DW, Pope CA, 3rd, Xu X, Spengler JD, Ware JH, Fay ME, et al. An association between air pollution and mortality in six U.S. cities. N Engl J Med. 1993;329(24):1753-9. doi: 10.1056/NEJM199312093292401. PubMed PMID: 8179653.
4. HEI. Traffic-Related Air Pollution: A Critical Review of the Literature on Emissions, Exposure, and Health Effects. HEI Special Report 17Health Effects Institute, Boston (2010) [Available: http://pubshealtheffectsorg/getfilephp?u=553 [accessed 8 June 2015]]. 2010.
5. Bellander T, Berglind N, Gustavsson P, Jonson T, Nyberg F, Pershagen G, et al. Using geographic information systems to assess individual historical exposure to air pollution from traffic and house heating in Stockholm. Environ Health Perspect. 2001;109(6):633-9. Epub 2001/07/11. doi: sc271_5_1835 [pii]. PubMed PMID: 11445519; PubMed Central PMCID: PMC1240347.
6. Johansson C, Burman L, Forsberg B. The effects of congestions tax on air quality and health. . Atmos Environ 43, 4843-4854. 2009.
7. Cyrys J, Eeftens M, Heinrich J, Ampe C, Armengaud A, Beelen R, et al. Variation of NO2 and NOx concentrations between and within 36 European study areas: Results from the ESCAPE study. Atmospheric Environment. 2012;62:374-90. doi: http://dx.doi.org/10.1016/j.atmosenv.2012.07.080.
8. Armstrong BG. Effect of measurement error on epidemiological studies of environmental and occupational exposures. Occup Environ Med. 1998;55(10):651-6. PubMed PMID: 9930084; PubMed Central PMCID: PMC1757516.
9. de Hoogh K, Korek M, Vienneau D, Keuken M, Kukkonen J, Nieuwenhuijsen MJ, et al. Comparing land use regression and dispersion modelling to assess residential exposure to ambient air pollution for epidemiological studies. Environment international. 2014;73:382-92.
10. Jerrett M, Arain A, Kanaroglou P, Beckerman B, Potoglou D, Sahsuvaroglu T, et al. A review and evaluation of intraurban air pollution exposure models. J Expo Anal Environ Epidemiol. 2005;15(2):185-204. doi: 10.1038/sj.jea.7500388. PubMed PMID: 15292906.
11. Briggs DJ, Collins S, Elliott P, Fischer P, Kingham S, Lebret E, et al. Mapping urban air pollution using GIS: a regression-based approach. International Journal of Geographical Information Science. 1997;11(7):699-718.
12. Sellier Y, Galineau J, Hulin A, Caini F, Marquis N, Navel V, et al. Health effects of ambient air pollution: do different methods for estimating exposure lead to different results? Environ Int. 2014;66:165-73. doi: 10.1016/j.envint.2014.02.001. PubMed PMID: 24598283.
13. Maynard D, Coull BA, Gryparis A, Schwartz J. Mortality risk associated with short-term exposure to traffic particles and sulfates. Environ Health Perspect. 2007;115(5):751-5. doi: 10.1289/ehp.9537. PubMed PMID: 17520063; PubMed Central PMCID: PMC1867995.
14. Gryparis A, Coull BA, Schwartz J, Suh HH. Semiparametric latent variable regression models for spatiotemporal modelling of mobile source particles in the greater Boston area. J Roy Stat Soc C-App. 2007;56:183-209. doi: DOI 10.1111/j.1467-9876.2007.00573.x. PubMed PMID: WOS:000245159600005.
15. Mavko ME, Tang B, George LA. A sub-neighborhood scale land use regression model for predicting NO(2). Science of the Total Environment. 2008;398(1-3):68-75. doi: 10.1016/j.scitotenv.2008.02.017. PubMed PMID: WOS:000256979900009.

16. Ainslie B, Steyn DG, Su J, Buzzelli M, Brauer M, Larson T, et al. A source area model incorporating simplified atmospheric dispersion and advection at fine scale for population air pollutant exposure assessment. Atmospheric Environment. 2008;42(10):2394-404. doi: 10.1016/j.atmosenv.2007.12.021. PubMed PMID: WOS:000255163200011.
17. Su JG, Brauer M, Ainslie B, Steyn D, Larson T, Buzzelli M. An innovative land use regression model incorporating meteorology for exposure analysis. Science of the Total Environment. 2008;390(2-3):520-9. doi: 10.1016/j.scitotenv.2007.10.032. PubMed PMID: WOS:000252538300024.
18. Wilton D, Szpiro A, Gould T, Larson T. Improving spatial concentration estimates for nitrogen oxides using a hybrid meteorological dispersion/land use regression model in Los Angeles, CA and Seattle, WA. Sci Total Environ. 2010;408(5):1120-30. doi: 10.1016/j.scitotenv.2009.11.033. PubMed PMID: 20006373.
19. Johansson C, Hadenius A, Johansson P-Å, Jonson T. SHAPE: THE STOCKHOLM STUDY ON HEALTH EFFECTS OF AIR POLLUTION AND THEIR ECONOMIC CONSEQUENCES. PART 1: NO2 AND PARTICULATE MATTER IN STOCKHOLM. CONCENTRATIONS AND POPULATION EXPOSURE. VAEGVERKET PUBLIKATION. 1999(1999: 41).
20. Johansson C, Eneroth K. TESS Traffic Emissions, Socioeconomic valuation and Socioeconomic measures-Part 1: Emissions and exposure of particles and NOx. 2007.
21. Gruzieva O, Bellander T, Eneroth K, Kull I, Melen E, Nordling E, et al. Traffic-related air pollution and development of allergic sensitization in children during the first 8 years of life. J Allergy Clin Immunol. 2012;129(1):240-6. Epub 2011/11/23. doi: 10.1016/j.jaci.2011.11.001

S0091-6749(11)01739-8 [pii]. PubMed PMID: 22104609.
22. Rosenlund M, Berglind N, Pershagen G, Hallqvist J, Jonson T, Bellander T. Long-term exposure to urban air pollution and myocardial infarction. Epidemiology. 2006;17(4):383-90. doi: 10.1097/01.ede.0000219722.25569.0f. PubMed PMID: 16699471.
23. Panasevich S, Leander K, Ljungman P, Bellander T, de Faire U, Pershagen G, et al. Interaction between air pollution exposure and genes in relation to levels of inflammatory markers and risk of myocardial infarction. BMJ Open. 2013;3(9):e003058. doi: 10.1136/bmjopen-2013-003058. PubMed PMID: 24056475; PubMed Central PMCID: PMC3780315.
24. Johansson C, Andersson C, Bergström R, Krecl P. Exposure to particles due to local and non-local sources in Stockholm: estimates based on modelling and measurements 1997-2006. ITM RAPPORT. 2008(175).
25. Omstedt G, Andersson S, Gidhagen L, Robertson L. Evaluation of new model tools for meeting the targets of the EU Air Quality Directive: a case study on the studded tyre use in Sweden. International Journal of Environment and Pollution. 2011;47(1-4):79-96.
26. Beelen R, Hoek G, Vienneau D, Eeftens M, Dimakopoulou K, Pedeli X, et al. Development of NO 2 and NO x land use regression models for estimating air pollution exposure in 36 study areas in Europe–the ESCAPE project. Atmospheric Environment. 2013;72:10-23.
27. Su JG, Jerrett M, Beckerman B. A distance-decay variable selection strategy for land use regression modeling of ambient air pollution exposures. Sci Total Environ. 2009;407(12):3890-8. doi: 10.1016/j.scitotenv.2009.01.061. PubMed PMID: 19304313.
28. Eeftens M, Beelen R, de Hoogh K, Bellander T, Cesaroni G, Cirach M, et al. Development of Land Use Regression models for PM(2.5), PM(2.5) absorbance, PM(10) and PM(coarse) in 20 European study areas; results of the ESCAPE project. Environ Sci Technol. 2012;46(20):11195-205. doi: 10.1021/es301948k. PubMed PMID: 22963366.
29. Williams RL. A note on robust variance estimation for cluster-correlated data. Biometrics. 2000;56(2):645-6.
30. Dijkema MB, Gehring U, van Strien RT, van der Zee SC, Fischer P, Hoek G, et al. A comparison of different approaches to estimate small-scale spatial variation in outdoor NO(2) concentrations. Environ Health Perspect. 2011;119(5):670-5. doi: 10.1289/ehp.0901818. PubMed PMID: 21193385; PubMed Central PMCID: PMC3094419.

31. Franco V, Kousoulidou M, Muntean M, Ntziachristos L, Hausberger S, Dilara P. Road vehicle emission factors development: A review. Atmospheric Environment. 2013;70:84-97.
32. Eneroth K, Johansson C, Bellander T. Exposure comparison between measurements and calculations based on dispersion modelling (EXPOSE). Stockholm Environment and Health Protection Administration. 2006.
33. Arain MA, Blair R, Finkelstein N, Brook JR, Sahsuvaroglu T, Beckerman B, et al. The use of wind fields in a land use regression model to predict air pollution concentrations for health exposure studies. Atmospheric Environment. 2007;41(16):3453-64. doi: 10.1016/j.atmosenv.2006.11.063. PubMed PMID: WOS:000246430000014.
34. Ryan PH, LeMasters GK. A review of land-use regression models for characterizing intraurban air pollution exposure. Inhal Toxicol. 2007;19 Suppl 1:127-33. doi: 10.1080/08958370701495998. PubMed PMID: 17886060; PubMed Central PMCID: PMC2233947.
35. Brauer M, Hoek G, van Vliet P, Meliefste K, Fischer P, Gehring U, et al. Estimating long-term average particulate air pollution concentrations: application of traffic indicators and geographic information systems. Epidemiology. 2003;14(2):228-39. doi: 10.1097/01.EDE.0000041910.49046.9B. PubMed PMID: 12606891.
36. Eneroth K, Johansson C, Bellander T. Exposure, comparison between measurements and calculations based on dispersion modeling (EXPOSE)(2006) Report from the stockholm and Uppsala air quality management http://slb.nu/slb/rapporter/pdf8/lvf2006_012.pdf. [Available: http://slbnu/slb/rapporter/pdf8/lvf2006_012pdf[accessed 2 September 2015]].
37. Wang M, Beelen R, Eeftens M, Meliefste K, Hoek G, Brunekreef B. Systematic evaluation of land use regression models for NO(2). Environ Sci Technol. 2012;46(8):4481-9. doi: 10.1021/es204183v. PubMed PMID: 22435498.
38. Chinn S. Statistics for the European Respiratory Journal. European Respiratory Journal. 2001;18(2):393-401.
39. Basagaña X, Rivera M, Aguilera I, Agis D, Bouso L, Elosua R, et al. Effect of the number of measurement sites on land use regression models in estimating local air pollution. Atmospheric Environment. 2012;54:634-42.
40. Austin PC, Steyerberg EW. The number of subjects per variable required in linear regression analyses. Journal of clinical epidemiology. 2015;68(6):627-36.

Figures and Tables

Figure 1. Mean levels of daily NO_x observed at a rural, urban and traffic site and the daily mean global radiation during the years of the monitoring campaign in Stockholm County

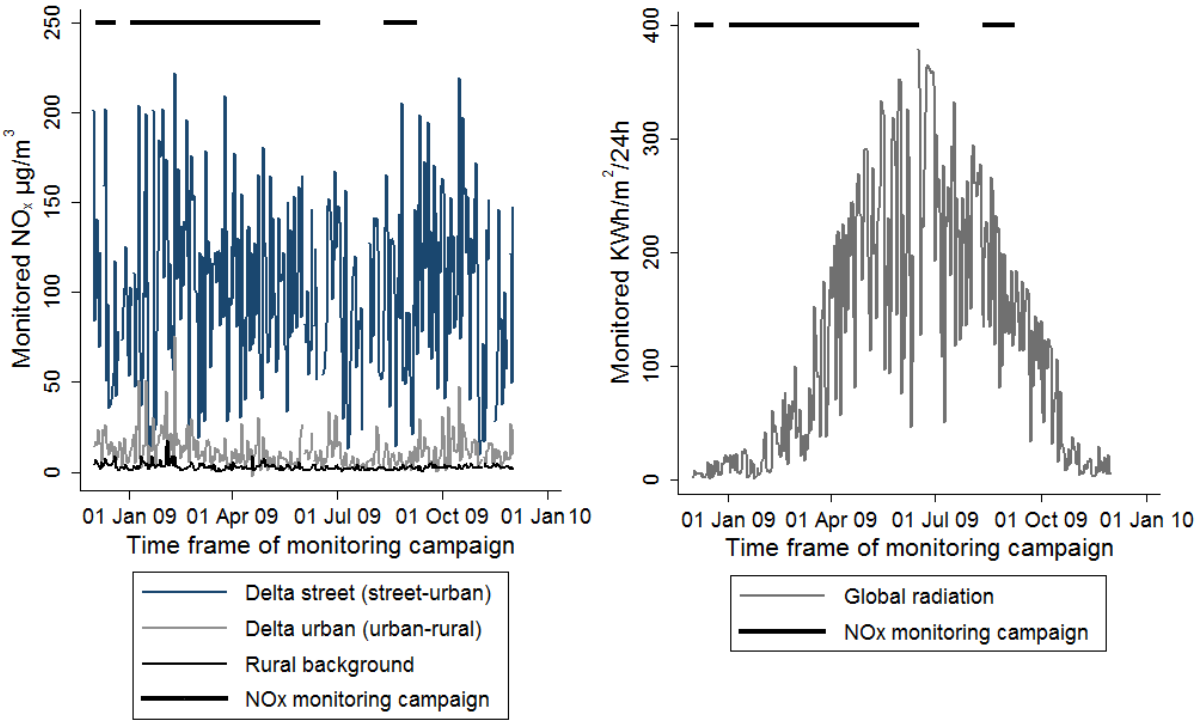
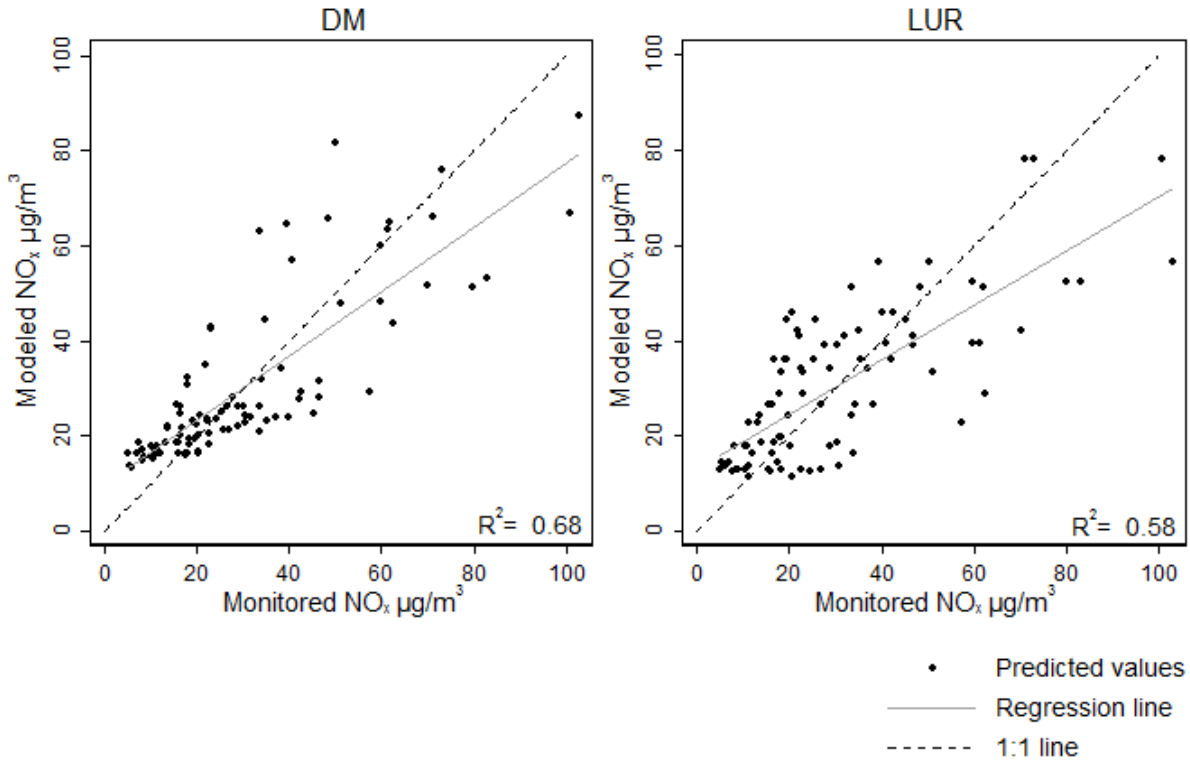


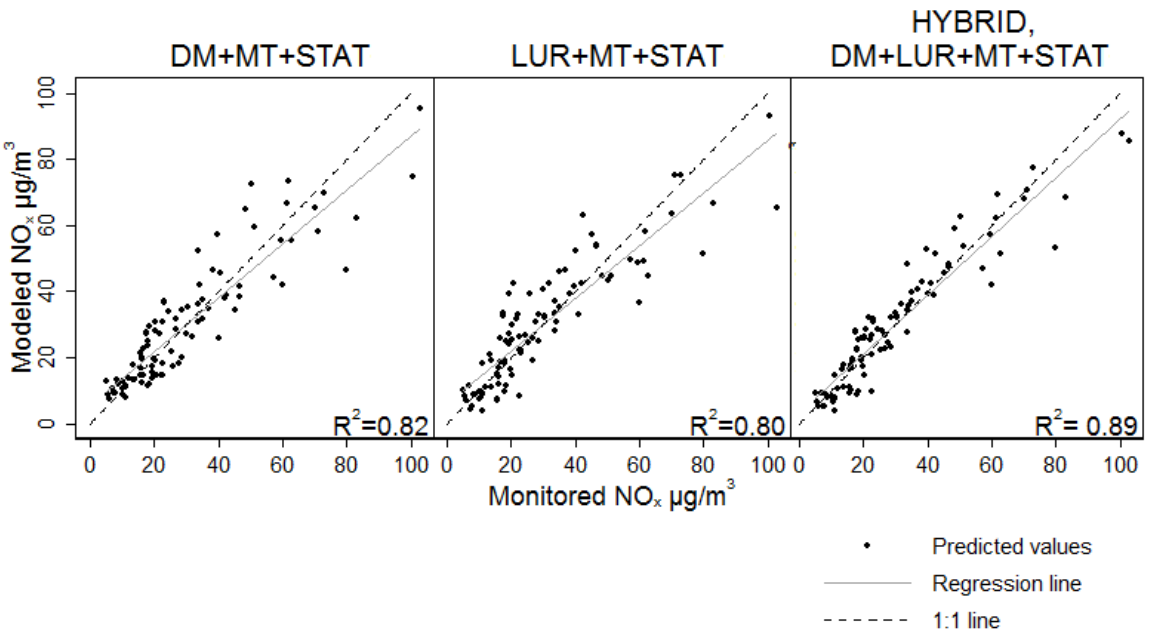
Figure 2. Dispersion and Land use regression modeled predictions of NO_x concentrations related to 92 bi-weekly monitored NO_x observations by univariate regression



DM = Airviro Gauss and SIMAIR road dispersion model

LUR = land use regression model

Figure 3. Comparison of Model specific NO_x predictions from three modeling scenarios; 1, Dispersion modeling with additional information on global radiation, 2, Land use regression modeling including global radiation and 3, a Hybrid model including Dispersion modeling, global radiation and LUR components.



DM = Airviro Gauss and SIMAIR road dispersion model

LUR = land use regression model

STAT= stationary monitoring, delta urban NO_x (urban –rural)

MT =meterological variables, (global radiation)

Table 1. Performance evaluation (coefficient of determination/root means square error and leave one out cross-validation) and model structures of the DM, LUR and Hybrid model explaining observed levels of NO_x.

Multivariate linear regression				LOOCV	
Model	intercept + (slope x predictors)	R^2 [a]	RMSE [b]	R^2 [a]	RMSE [b]
DM	9.67 +(1.14x DM)	0.68	12.05	0.66	12.4
LUR	10.12917 +(.004 x population density 300m) +(.001 x traffic intensity in the nearest street)	0.58	13.90	0.55	14.2
DM+MT+STAT	9.29+ (1.10 x DM)+ (-.059x global radiation) + (0.70 x delta urban NO_x (urban-rural))	0.82	9.14	0.80	9.5
LUR+MT+STAT	1.00+(1.40x delta urban NO_x (urban-rural))+(.001xtraffic intensity on nearest street) + (0.025xpopulation density 100m)+(-0.046xglobal radiation)	0.80	9.7	0.77	10.1
HYBRID	2,92+ (0.67 x DM)+ (-.054xglobrad)+ (.0008 x traffic intensity on nearest street)+ (.015x population density 100m) +(0.99 x delta urban NO_x (urban-rural)	0.89	7.15	0.87	7.6

DM = dispersion modeled NO_x estimates, MT = meteorological predictors, final models included levels of global radiation from continues monitoring, STAT = NO_x levels from continuous monitoring, final models included delta urban NO_x (urban-rural), LUR = land use regression data, final models included population density (calculated within buffers with specified radii) and traffic intensity. LOOCV = leave one out cross-validation

[a] Coefficient of determination

[b] Root mean square error

Supplementary Materials

Can Dispersion modeling of air pollution be improved by land use regression?

Supplementary Table 1. Levels of temporally defined variables in 2 week means corresponding to the 93 bi-weekly NO_x observations and the predictor specific direction of effect

Monitored data	R^2 [a] with monitored NO_x	Mean	Std. Dev.	Min	Max	Predefined direction of effect on NO_x levels
NO_x ($\mu g/m^3$)	1	30.3	21.1	5.0	103.0	
Dispersion modeled NO_x	0.82	18.0	15.1	3.2	68.0	+
Rural NO_x ($\mu g/m^3$)	0.37	3.3	1.3	1.5	6.8	+
Delta Urban NO_x ($\mu g/m^3$) [b]	0.42	12.2	4.7	6.5	24.7	+
Delta Traffic NO_x ($\mu g/m^3$) [c]	0.10	102.8	14.9	81.0	132.4	+
Global radiation (Kwh/m^2 /24h)	-0.43	122.4	85.4	4.8	252.3	-
Eastern wind component (degrees°) [d]	0.23	-0.12	0.3	-0.7	.5	?
Northern wind component(degrees°) [d]	0.01	-0.1	0.3	-0.7	0.5	?
Wind speed (m/s /24h)	-0.08	2.9	0.5	1.9	3.7	-
Temperature (°C /24h)	-0.40	6.4	6.4	- 3.1	16.5	-
Relative humidity (% /24h)	0.40	74.6	10.0	54.4	90.8	+

[a] Pearson correlation between 93, bi-weekly observations and temporally defined data

[b] (urban NO_x-rural NO_x)

[c] (traffic NO_x-urban NO_x)

[d] Eastern and Northern components represents eastern and northern winds calculated from wind direction data. In the LUR model, together with wind speed, wind vectors can be calculated.

Supplementary Table 2. The degree of association (explained variance) between observed NO_x levels and predictors in the hybrid model as defined by the partial R^2 and predictor specific beta estimates and Confidence intervals.

Spatial and temporal predictors in the hybrid model			
	Partial $R^{2\,a}$	β [b]	CI [b]
DM ($\mu g/m^3$)	0.68	0.67	(0.48 - 0.85)
Traffic intensity on nearest street Veh. day^{-1}	0.58	0.0007	(0.0006 - 0.0010)
population density 100m (inhabitants/100m radius)	0.53	0.015	(0.008 -0.023)
Global radiation (KWh/m^2/24h)	-0.38	-0.05	(-0.07 - -0.04)
Delta Urban NO_x (urban-rural) ($\mu g/m^3$)	0.38	0.98	(0.42 - 1.55)

[a] Partial R^2 = the proportion of shared variance between observed NO_x concentrations and separate predictors while controlling for other predictors in the regression

[b] level of NO_x per 1 unit change