Imperial College London

# Harmonised ambient air pollution and road traffic noise exposures linked to cardiovascular outcomes in European cohorts

Thesis submitted for Doctor of Philosophy degree in Epidemiology

Yutong Cai

Department of Epidemiology and Biostatistics

School of Public Health

Imperial College London

## Declaration of Originality

*I, Yutong Cai, declare that this thesis is my own work and is based on the research that I conducted during 2012-2016. It has not been submitted in any form for another degree or diploma at any university. Information derived from the published and unpublished work of others has been fully acknowledged in the text and references is listed in the bibliography.*

## Copyright Declaration

## Abstract

Ambient air pollution and traffic-related noise are the two leading environmental risk factors for health in Europe. Associations between long-term exposure to air pollution or noise and cardiovascular diseases (CVD) were not entirely consistent across previous studies in adults. Moreover, noise may confound the relationship between air pollution and CVD, and vice versa.

This PhD project was conducted to study the separate and joint effects of both air pollution and noise on 1) CVD blood biochemistry including C-reactive protein, blood lipids and glucose and on 2) incident CVD outcomes. Health and exposures data were harmonised across four European cohorts (EPIC-Oxford, HUNT, LifeLines, UK Biobank), as part of the EU-funded BioSHaRE project. All harmonised data were virtually pooled for individual-level analyses in DataSHIELD, a novel statistical tool to perform a 'compute to data' statistical approach.

The cross-sectional analyses on biochemistry data generally suggested that both air pollution and noise were significantly associated with adverse changes in markers of systemic inflammation, blood lipids and glucose. The significant association between road traffic noise and C-reactive protein or triglycerides was confounded by air pollution whilst both air pollution and noise were significantly and independently associated with elevated blood glucose levels.

Incident analyses suggested a possible increased risk for both particulate matter (PM) and gaseous air pollution on incident cerebrovascular disease but a null association for ischaemic heart disease (IHD). Daytime noise was associated with a non-significantly increased risk for incident IHD but evidence for cerebrovascular disease was inconclusive. Both air pollution and noise effects on CVD outcomes were independent from each other.

This PhD study provides some novel evidence of both air pollution and noise on CVD biochemistry and incident CVD outcomes, and is a substantial addition to the current knowledge of cardiovascular health effects of both ambient air pollution and traffic noise.

# Contents

# List of Tables

# List of Figures

# Acknowledgements

*This work was supported by funding from the European Union Seventh Framework Programme (FP7/2007-2013) Biobank Standardisation and Harmonisation for Research Excellence in the European Union-BioSHaRE-EU (n° 261433).*

*Acknowledgements to study participants and the data collection teams in each cohort.*

*I would like to thank, wholeheartedly, Dr Susan Hodgson, Dr Anna Hansell, Dr Marta Blangiardo, for their tremendous academic efforts in supervising this PhD and for giving me opportunities, freedom and inspirations in pursing my research career. It has been such an amazing experience working with all of you.*

*I would also like to thank Prof Paul Elliott, for his role as the Imperial's PI for BioSHaRE project and his insightful advice on my PhD and other works.*

*Gratitude also goes to Dr John Gulliver and Dr Kees de Hoogh, for their huge efforts in leading noise and air pollution modelling works.*

*Special mention goes to David Morley, Eric Johnson, Paul Burton, Dany Doiron, Stephane Mbatchou, Yannick Marcon, Demetirs Avraam and Wilma Zijlema for their supports into this work. And to the MRC-PHE Centre, Department of Epidemiology and Biostatistics within the School of Public Health, a friendly place that I have been inspired by so many talented colleagues from all over the world.*

*Finally, thanks go to mum and dad, who will be thrilled and impressed about the journeys I have achieved so far. Thanks also go to Ping, for lives shared, and for her unbelievable patience and support when I was spending time on this work.*

# Introduction

Ambient air pollution is the leading environmental risk factor for health in Europe, followed by traffic-related noise(Hanninen et al., 2014). Associations between long-term exposure to ambient air pollution or traffic-related noise and cardiovascular outcomes in adults have been investigated in various cohorts around the world. However, results are not entirely consistent across these studies, likely due to, different study designs and exposure assessment approaches, quality of the exposure assessment as well as the health outcomes assessment, adjustments made for different sets of covariates. Compared to studies on the effects of air pollution on cardiovascular outcomes, fewer investigated the noise effects. Importantly, as road traffic is the common source of both ambient air pollution and noise, these exposures may confound effects of each other. Increasing numbers of studies have looked at this potential confounding issue in recent years, with most current evidence suggesting an independent effect of each. This however needs further investigation as this issue may depend on the health outcomes studied as well as the specific areas or urban structures where the study is conducted. In **Chapter 1**, most of the key literature have been critically reviewed to give readers an up-to-date knowledge on current air pollution/noise and cardiovascular research.

To address some of the research gaps and limitations identified in the literature review, I proposed to study the separate and joint effects of both long-term air pollution and road traffic noise on cardiovascular health in four European cohorts (EPIC-Oxford, HUNT, LifeLines, UK Biobank) using a harmonised approach, as funded by the European Union Seventh Framework Programme BioSHaRE (Biobank Standardisation and Harmonisation for Research Excellence in the European Union) project (2012-2015).

BioSHaRE was established to develop novel tools and computing infrastructures, with an ambitious aim to facilitate data sharing and pooling across multiple biobanks and cohorts in Europe. This PhD, as one of the cores of BioSHaRE project, was among the first to test and validate these novel tools in a real-life epidemiological project.

Harmonised approach was the spotlight of this PhD project. Questionnaire and health outcome data were harmonised retrospectively across cohorts by the candidate (**Chapter2**), following a validated protocol. These data were harmonised according to standard definitions, using computerised scripts being applied to all cohorts. With strong inputs from the exposure assessment teams within the MRC-PHE Centre for Environment and Health at Imperial College London, common Europe-wide air pollution and noise exposure models were applied across the four cohorts (**Chapter 3**). These common models were used to minimise the differences between cohorts that would otherwise be introduced by having different assessment methods.

A huge amount of efforts were made to harmonise both questionnaire and exposures data across cohorts, allowing an effective pooling of individual-level data for analyses. Yet, a physical sharing of data has always been associated with complex ethico-legal issues, which undermines research potentials and collaborations. To this regard, I conducted the pooled individual-level analyses using a new novel tool, DataSHIELD, as developed by BioSHaRE. In brief, individual-level harmonised data from various cohorts were deposited onto their local secure servers in the respective research centres, and then DataSHIELD was able to link up all these local servers to virtually pool these data to form an integrated database for analyses in DataSHIELD. Therefore, by using this DataSHIELD 'compute-to-data' approach, data are not physically shared and data owner retains governance of their data. The successful applications of DataSHIELD in this PhD project will have useful implications for future epidemiological studies. An introduction of DataSHIELD and statistical methods used in this PhD project is in **Chapter 4**.

Two epidemiological analyses were conducted in this PhD project using DataSHIELD. Results from DataSHIELD were further validated against cohort-specific meta-analyses and analyses on physically pooled data in Stata.

Cross-sectional associations between air pollution or noise and blood biochemistry markers for cardiovascular diseases (CVD) were first examined, pooling data from HUNT3 and LifeLines (**Chapter**

**5**). Results generally suggested that both exposures were significantly associated with adverse changes in blood biochemistry markers of systemic inflammation, blood lipids and blood glucose, providing some mechanistic insights into links between air pollution/noise and cardiovascular disease. I then explored the direct links between long-term air pollution, noise and incident cardiovascular diseases in **Chapter 6**, pooling data from EPIC-Oxford, HUNT2 and UK Biobank. In view of comparisons with other European studies, I further conducted meta-analyses on estimates from all published European studies, including the results from this PhD, to give updated pooled effect estimates on the associations between long-term air pollution and incident CVD.

In the last chapter (**Chapter 7),** an overall summary of findings in this PhD project was given. I also highlighted some important issues which were not addressed in this work and will require careful investigations in future studies. In additions, some implications from this work for environmental health policy making were also noted.

***Hypothesis and research objectives***

The general hypothesis is that long-term exposure to both ambient air pollution and road traffic noise increase the risks for cardiovascular diseases in adult populations. It is speculated that adverse CVD risk profiles may underlie these associations.

This PhD project has several objectives.

1. To critically review the current literatures on long-term air pollution and traffic noise on cardiovascular health;
2. To retrospectively harmonise data across cohorts to ensure an effective data pooling for analyses;
3. To quantify the separate and joint effects of both long-term ambient air pollution and road traffic noise exposures on blood biochemistry markers for cardiovascular diseases, including C-reactive protein, blood lipids and blood glucose;

4. To quantify the separate and joint effects of both exposures on incidence of cardiovascular diseases, including the total cardiovascular diseases, ischaemic heart disease and cerebrovascular disease;
5. To conduct an updated meta-analysis which includes the results from this PhD and other European studies on air pollution effects on incident cardiovascular disease;
6. To validate the effective use of DataSHIELD, using two methods: cohort-specific meta-analysis and directly analysing data which were physically pooled.

# Chapter 1 Ambient air pollution, noise and cardiovascular health

Air pollution was probably first recognised as a public health problem in England in the 17th century. In one of the pamphlets published by John Evelyn in 1661(Mark Jenner, 1995), he described that the then Londoners '*breathe nothing but an impure and thick mist…and filthy vapour*'. He attributed the '*clouds of smoke*' to the immoderate use of coal by the inhabitants and even developed some ambitious plans to improve the quality of London's air. Yet unfortunately his advice was not heeded until the famous episodes of severe air pollution that occurred in London nearly three hundred years later in December 1952 (known as London Great Smog 1952)(Bell, Davis & Fletcher, 2004). This smog directly or indirectly killed thousands of people and made several thousands of citizens ill over the course of only a few days because of the polluted air. This "wake-up call" then led to the Clean Air Act 1956, the first time air pollution was to be regulated systematically by law in the United Kingdom (UK). Today, in the 21st century, although air pollution levels have been substantially lowered from those experienced in earlier centuries, in many parts of the west, including the UK and Europe, poor air quality still poses a significant adverse effect on our health. In 2010, it was estimated that 3.1 million out of 52.8 million deaths worldwide were attributed to air pollution(Newby et al., 2015). Outdoor air pollution was ranked ninth among the modifiable risk factors for disease worldwide(Newby et al., 2015). According to another report published by World Health Organisation (WHO), about 7 million people die from exposure to both indoor and outdoor air pollution each year(Kuehn, 2014). Evidence about the deleterious effects of air pollution on human's health is growing. It is now widely accepted that air pollution increases risk for heart diseases and respiratory diseases, although the pollutants and mechanisms underlying these associations are still under debate.

Environmental noise exposure is also ubiquitous but has been the focus of far fewer studies than air pollution. In modern societies, most people live with environmental noise on a daily basis and are likely to treat it as part of their daily life. A comprehensive report published by WHO in 2010 estimated that 61,000 healthy years were lost from ischaemic heart disease alone as a result of exposure to road

traffic noise in western Europe(European Office WHO, 2011). Indeed, it was not until the last two decades that scientific research has linked environmental noise exposure to various health effects including cardiovascular diseases in adults(Babisch, 2006). In 2002, a European Union (EU)-wide legislation, known as "Environmental Noise Directive (END)-2002/49/EC", was adopted by each member state to assess and regulate the environmental noise(Directive 2002/49/EC, 2002). This was treated as a significant step towards controlling excessive environmental noise and assessing its health effects across Europe. In fact, with reliable and rigorous assessments of environmental noise now available, more research is currently being conducted in Europe and North America, including on less studied health outcomes such as metabolic disorders.

In this first chapter, I present a critical review of existing epidemiological studies to date (assessed up to 30 Nov 2015) linking long-term ambient air pollution and environmental road traffic noise exposures to specific cardiovascular outcomes in adult populations.

## 1.1. Ambient air pollution

What is ambient air pollution?

Air pollution is defined as "*contamination of the indoor or outdoor environment by any chemical, physical or biological agent that modifies the natural characteristics of the atmosphere*" by WHO (http://www.who.int/topics/air_pollution/en/, accessed March 2015). In today's world, common sources of air pollution include traffic, industrial facilities and household combustion activities. Sometimes, natural source of pollution such as ashes from volcanic eruptions, dust storm from deserts and forest fires can also cause local episodes of severe air pollution (WHO). In general, air pollution in modern societies comprises a mixture of pollutants with various complex physical and chemical characteristics. It is nearly impossible to study these mixtures of pollutants. Given the design and scope of this project and based on the previous accumulated evidence, I chose to focus on particulate matter (PM) and gaseous pollutants such as Nitrogen Oxides as general indicators of ambient (outdoor) air pollution.

Particulate matter comprises a mixture of particles of different sizes and with various chemical constituents attached to it. PM can be generated from diverse sources, including natural (dust, volcanic ash, pollens etc.) and man-made activities (fossil fuel combustion, biomass burning, cigarette smoking etc.)(Kelly & Fussell, 2012a). In urban areas, diesel exhaust from motor vehicles is the main source of PM, with industrial factories and power plants also contributing(Riedl & Diaz-Sanchez, 2005). PM is emitted into the atmosphere directly from sources (primary particles) or indirectly via secondary chemical reactions (secondary particles)(Kelly & Fussell, 2012a).

In most scientific literatures, PM is broadly categorised by aerodynamic diameter: particles ≤10μm ($PM_{10}$); particles ≤2.5μm ($PM_{2.5}$, usually be referred to as 'fine particle'); coarse particles 2.5~10 μm (PM coarse)(Brook et al., 2010). These particle fractions are typically measured in their mass per volume of air ($\mu g/m^3$). It is generally believed that particles with diameter 10 μm or less can deposit along the airway, and the smaller the size of PM, the more easily it penetrates deep into the lung and passes the air-blood barrier(Kelly & Fussell, 2012a). Most studies linking PM exposures with specific health outcomes have focused on particle mass as a crude indicator of air pollution; but the toxic components that exert the most harmful effects are still unclear.

Unlike fine particle ($PM_{2.5}$) which can travel long-distances and distribute homogeneously at regional level(Kunzli, 2014), nitrogen oxides (NOx) are gases directly emitted from motor vehicles and are often used as a proxy for near-road traffic-related air pollution(Katsouyanni, 2003). Nitrogen oxides are usually measured in $\mu g/m^3$ or parts per billion(ppb). Nitrogen dioxide ($NO_2$) is one of the most studied gaseous pollutants with regards to its health effects. However, whether it has direct health effects or merely acts an indicator for other health relevant pollutants remains debatable given that correlations between $NO_2$ and other pollutants, for example particulate matter, are usually high(Faustini, Rapp & Forastiere, 2014).

### 1.1.1. Long-term ambient air pollution and cardiovascular effects

There is growing evidence regarding the adverse cardiovascular effects of ambient air pollution. These effects include both mortality and morbidity due to not only short-term exposure to peaks of pollution but also long-term, low level exposures to ambient air pollution. In 2010, an updated scientific statement from the American Heart Association (AHA) had concluded that long-term exposure to $PM_{2.5}$ increases risks for cardiovascular deaths, and overall evidence is supportive that $PM_{2.5}$ exposure is a causal modifiable risk factor for cardiovascular morbidity and mortality(Brook et al., 2010). The REVIHAAP (Review of evidence on health aspects of air pollution) report coordinated by WHO has also documented detailed up-to-date scientific evidence of the health effects of exposure to different air pollutants. In this report, it was concluded that life expectancy was reduced by an average of 9 months in Europe as a result of particulate air pollution(World Health Organisation Europe, 2013).

In the following sessions, current epidemiological evidence of long-term air pollution exposures (i.e., exposures of a year or more) with respect to specific cardiovascular health outcomes were critically reviewed.

#### *Mortality studies: all-cause and cardiovascular cause*

The associations between air pollution and all-cause or specific-cause mortality have been investigated worldwide, although most evidence to date has come from populations in North America and Western Europe.

*Studies in North America*

Two early studies in America in the 1990s, the Harvard Six Cities study(Dockery et al., 1993) and the American Cancer Society (ACS)(Pope et al., 1995) study, were among the first to investigate air pollution effects in prospective cohorts. After the follow-up period (14 to 16 years for Six Cities study and 7 years for ACS study), the researchers compared the mortality rates among recruited adults in different cities with different levels of air pollution based on city-specific central monitoring, after controlling for major individual risk factors including tobacco smoking, occupational exposures and

individual socioeconomic status. Both studies found an increased risk in all-cause mortality in relation to long-term PM exposures in those living in the most polluted cities. In the six cities study, based on the fine particle ($PM_{2.5}$) levels, compared with the least polluted city, the adjusted all-cause mortality rate ratio was 26% higher (95%CI: 8 to 47%) in the most polluted city, while for cardiorespiratory mortality, the rate ratio was 37% higher (95%CI: 11 to 68%). In the ACS study, every 10 $\mu g/m^3$ increase in $PM_{2.5}$ was associated with a 17% increase (95%CI: 9 to 26%) in all-cause mortality in the most polluted cities. Unlike the six cities study, the ACS study found a less strong association with deaths due to a cardiopulmonary cause, with a 6% increase in deaths for every 10 $\mu g/m^3$ increase in $PM_{2.5}$.

Extended re-analyses of both studies have replicated these findings and concluded that mortality due to air pollution was much higher for cardiovascular causes rather than respiratory causes(Lepeule et al., 2012; Laden et al., 2006; Pope et al., 2004a). Laden et al(Laden et al., 2006) followed the mortality rates of the Six Cities study participants for an additional eight years (to the end of 1998), collected historic (pre-1987) $PM_{2.5}$ data from city-specific monitoring sites for all six cities, and developed a crude regression model built on routinely collected $PM_{10}$ data (1988-1998) and humidity-corrected visibility data from local airports to estimate city-specific annual mean $PM_{2.5}$ for years 1987-1998. During the entire follow-up period, each 10 $\mu g/m^3$ increase in $PM_{2.5}$ was associated with 16% (95%CI: 7 to 26%) increase in all-cause mortality, a 28% increased risk (95%CI: 13 to 44%) for cardiovascular mortality, and a 8% increase risk (95%CI: 0.8 to 49%) for respiratory mortality. Moreover, during this extended follow-up period, there was a general reduction of air pollution levels in US cities, and the authors demonstrated that this reduction of $PM_{2.5}$ levels was significantly associated with reduced overall mortality and cardiovascular mortality.

One of the limitations in this study, is that only baseline data on key covariates (e.g. smoking, Body Mass Index) collected decades previously were used in the analysis. These confounding factors and their associations with air pollution would have also likely changed during a long period of follow-up. The same research group carried out a further follow-up up to 2009. They obtained city-specific $PM_{2.5}$

data for the period 1999 to 2009 directly from the air quality monitors (which started in 1999 in the US). Again, similar results for all-cause and cardiovascular deaths were replicated(Lepeule et al., 2012). In the follow-up ACS study, Pope CA 3rd et al(Pope et al., 2004a) also found an increased risk for various causes of cardiovascular deaths of between 8% and 18% with a 10 $\mu g/m^3$ increase in $PM_{2.5}$.

Although these studies were influential in demonstrating a clear link between air pollution and health, one drawback of both these earlier studies is that they used city-specific ambient air pollution monitor data as a proxy for an individual's mean air pollution exposure, which will likely have introduced exposure errors resulting in uncertainty in the size of studied associations with health.

Subsequent studies have attempted to improve on this assessment of exposure. Jerrett et al assigned individual estimates of $PM_{2.5}$ and $NO_2$ from a Land Use Regression (LUR) model to 73,711 participants of California in the ACS Cancer Prevention II Study(Jerrett et al., 2013). It was found that both LUR-derived $PM_{2.5}$ and $NO_2$ was positively associated with all-cause and ischaemic heart disease mortality. This study, with a more advanced exposure assessment method, was able to demonstrate the associations among individuals with various exposure levels within cities. The LUR model in this study used the government's routine monitoring sites to estimate exposure. These monitoring sites usually represent background sites, and therefore the near-road environment is not well represented. It was likely that the model will not therefore capture local scale spatial variations, including near-road elevated air pollution levels.

In the three cities of Ontario in Canada, Chen et al found that for each 5ppb increase in $NO_2$, mortality from cardiovascular diseases increased by 12% (95%CI: 7 to 17%) in a random sample of 205,440 adults drawn from an income tax database(Chen et al., 2013). In this study, air pollution exposure estimates were assigned to each participant using a LUR model based on individual residential histories. However this study did not directly adjust for individual smoking data as such data were not available. Another study in Vancouver using the same exposure assessment technique, reported that black carbon, but not $PM_{2.5}$ or $NO_2$, was associated with increased deaths from coronary heart

disease(Gan et al., 2011). In the ten largest Canadian cities within the Canadian Census Health and Environment Cohort (CanCHEC), Crouse et al assigned within-city individual $NO_2$ exposures to over 700,000 residents using city-specific LUR models and calculated city-wide annual mean $NO_2$ for each resident (between-city $NO_2$) using fixed monitoring sites in each city(Crouse et al., 2015). This study, after adjusting for both personal and contextual confounders, reported that within-city $NO_2$ contrasts, but not between-city contrasts, were significantly associated with increased mortality from all-cause and ischaemic heart disease during the 16 years of follow-up. This study was one of the first to study, simultaneously, the differential effects of both within and between-city $NO_2$ exposure contrasts for individuals. Results suggested that associations were mainly driven by the residential, within-city $NO_2$ air pollution variations, which may represent a more toxic mixture of air pollution at the local level. While the association between individual $NO_2$ exposure and ischaemic heart disease mortality was strong in this study (5% risk increase per 5ppb $NO_2$, 95%CI: 2 to 8%), the authors did not observe a significant association with mortality from cerebrovascular disease.

The previous six cities study and ACS study did not study the within-city $PM_{2.5}$ effects but reported significant associations with between-city $PM_{2.5}$ exposure contrast. Unlike $PM_{2.5}$ which has different sources other than traffic, the main source for $NO_2$ air pollution, particularly in urban areas, is mainly traffic-related. It is therefore expected that within-city $NO_2$ exposure contrasts is larger than those of between-city.

*Studies in Europe*

In Europe, extensive studies were also conducted in recent years in many countries either at a national or regional level.

An ecological study at local authority level in England found that higher levels of $PM_{10}$ were associated with higher all-cause mortality rates(Janke, Propper & Henderson, 2009). Air pollution data obtained for each local authority was from the local air quality monitoring network. In another study, using a

national database involving 830,000 English patients registered at their General Practice (GP), reported that increased risk of all-cause but not cardiovascular mortality was associated with a range of pollutants including $PM_{10}$, $PM_{2.5}$ and $NO_2$(Carey et al., 2013). This study used an emission-based air dispersion model at a 1km x 1km resolution to assign the individual air pollution estimates at the nearest postcode. This 1km resolution model, which would not capture potentially important local variations in $PM_{10}$ concentrations, may explain in part the weak or null associations found for cardiovascular mortality. Interestingly, in contrast to the findings in North America, this study reported larger effects on respiratory mortality rather than cardiovascular mortality.

Historical black smoke exposure was associated with all-cause and cardiovascular mortality in a Scottish study(Yap et al., 2012) and in an ecological small-area study in Great Britain(Elliott et al., 2007), both of these studies also reported that effects for respiratory mortality were stronger than for cardiovascular mortality.

There are mixed reports in the Netherlands. In a study of approximately 120,000 participants of the Netherlands Cohort Study (NLCS) on diet and cancer, a 10 $\mu g/m^3$ increase in $PM_{2.5}$ was associated with 6% higher risk for natural-cause mortality, although it was not statistically significant (95%CI: -0.3 to 16%)(Beelen et al., 2008). Associations between $NO_2$ and black smoke with natural-cause mortality were borderline statistically significant (Rate ratio (RR): 1.08, 95%CI: 1.00 to 1.16; RR: 1.05, 95%CI: 1.00 to 1.11 respectively). No associations were found for cardiovascular mortality with any of these pollutants. In the same cohort, but with a much smaller random sample of 5000 participants, Hoek et al reported no association between all-cause mortality and $NO_2$ exposure(Hoek et al., 2002). Conversely, in the same study, the authors found cardiopulmonary mortality was significantly associated with traffic indicators, such as living near a major road. Another study in the Netherlands also reported a significant positive association between traffic intensity on the nearest road and deaths from ischaemic heart disease(Beelen et al., 2009).

A study in Rome involving over a million Italian adults reported a small increase in all-cause mortality with LUR modelled $NO_2$ exposure (Hazard ratio (HR): 1.03, 95%CI: 1.02 to 1.03) and dispersion modelled $PM_{2.5}$ exposure (HR: 1.04, 95%CI: 1.03 to 1.05) per 10 μg/$m^3$ increase(Cesaroni et al., 2013). Both exposures were also associated with cardiovascular deaths with similar effect sizes. Associations with $NO_2$ persisted after co-adjustment for $PM_{2.5}$. Although individual risk factor data such as obesity, smoking habits and diet were not directly available for the full cohort, the approach to exposure assessment (a dedicated measurement campaign for $NO_2$ LUR model and a dispersion model for $PM_{2.5}$) was an advantage for this study.

Significant associations between $NO_2$ and all-cause and cardiovascular mortality among 52,061 participants were reported in a Danish cohort study(Raaschou-Nielsen et al., 2012). A Norwegian cohort study of 16,209 men living in Oslo found that each 10 μg/$m^3$ increase in NOx was associated with 8% (95%CI: 6 to 11) and 8% (95%CI: 3 to 12) higher risk of deaths from all-cause and ischaemic heart diseases respectively(Nafstad et al., 2004). No clear associations were found for $SO_2$(Sulphur dioxide). Neither the Danish nor Oslo study found associations with cerebrovascular mortality. Participants in both studies were followed for years and risk estimates obtained were adjusted for very detailed individual risk factors. It should be noted that exposure assessment methods were not comparable in these two studies. While the Danish study deployed a sophisticated dispersion model with detailed inputs, the Oslo study had calculated annual concentration (1974-1998) for each participant based on historical air pollution data, emissions and meteorological and topographical data.

Research in Germany(Heinrich et al., 2013), France(Filleul et al., 2005) and Spain(Boldo et al., 2011) also reported harmful effects of long-term air pollution exposure on mortality. Few cohort studies have been published outside North America and Europe. Emerging research is seen in Asian-Pacific countries, for example in China(Zhang et al., 2014; Zhou et al., 2014), Japan(Ueda et al., 2012; Yorifuji et al., 2010), Australia(Wang, Hu & Tong, 2009) and New Zealand(Hales, Blakely & Woodward, 2012).

Results in these settings have been inconsistent with some studies reporting an association, while others did not.

*The ESCAPE findings*

The recently concluded European Study of Cohorts for Air Pollution Effects (ESCAPE) project, using standardised individual exposures modelled at home addresses and harmonised health variables for a number of cohorts across Europe, also studied associations between air pollution exposure and various mortality outcomes. Beelen et al observed that for each 5 $\mu g/m^3$ increase in $PM_{2.5}$, risk for natural-cause mortality increased by 7% (95%CI: 2 to 13%) in a meta-analysis from 22 European cohorts including over 360,000 participants (average follow-up 13.9 years), with no evidence of heterogeneity in effect estimates between individual cohorts(Beelen et al., 2014a). Association remained even at levels of $PM_{2.5}$ exposure less than 25 $\mu g/m^3$, the current annual European limit. In contrast, in the same cohorts, using the same exposure and statistical methods, the authors did not find any significant associations for overall cardiovascular mortality; but there was some suggestive evidence for an increased risk for cerebrovascular disease mortality(Beelen et al., 2014b). The authors speculated that the risk profiles for cardiovascular diseases have changed over time, resulting in lower fatality rates. For non-malignant respiratory mortality, no associations were found for either PM or $NO_2$ in the 16 cohort studies included in the ESCAPE analysis(Dimakopoulou et al., 2014).

*The quantitative reviews on current evidence*

An overview of current evidence was updated by two recent review papers on particulate matter and $NO_2$ respectively. One, conducted by Hoek et al, reported that most, but not all, studies found significant associations between particulate air pollution and all-cause mortality(Hoek et al., 2013). An overall excess risk of cardiovascular mortality of 11% (95%CI: 5 to 16%) and all-cause mortality of 6% (95%CI: 4 to 8%) for a 10 $\mu g/m^3$ increase in $PM_{2.5}$ was reported, although heterogeneity existed across effect estimates from different studies. The authors concluded that the excess mortality was more

associated with cardiovascular diseases (particularly ischaemic heart disease) than with respiratory diseases. Several issues were discussed in the review. First, effect modifications by sex, smoking, education, obesity and diabetes would usually require a larger statistical power to detect differences within a single study. Second, exposure assessment varied across studies, in terms of the methods used to assign exposure estimates (estimates from city-wide or regional monitoring sites, LUR model, air dispersion model etc.), and the period of exposure (some studies secured exposures estimates before/during follow-up, while others could only obtain estimates years after the end of follow-up). Third, particle mass ($PM_{2.5}$ or $PM_{10}$) is only a crude indicator of regional air pollution, which may have many chemical compositions and contributing sources. As such, it was recommended that future research should establish which toxic component of the particle contributes most to mortality and identify the sources of these most harmful particles.

Another review paper concluded that the long-term effect of $NO_2$ on mortality is of a similar magnitude to $PM_{2.5}$(Faustini, Rapp & Forastiere, 2014). The authors found that the pooled effect on all-cause, cardiovascular and respiratory mortality for $NO_2$ and $PM_{2.5}$ were all consistently positive and statistically significant, with strongest associations seen for cardiovascular mortality. Based on only four eligible bi-pollutants analyses at the time of the review, effects of $NO_2$ seem to be independent from that of fine particle, highlighting the importance of considering the likely causal role of $NO_2$ independently to PM in future research.

### *Incident cardiovascular events*

Coronary heart disease (also known as ischaemic heart disease) and cerebrovascular disease are the two major types of cardiovascular diseases. It was recently revealed in 2010 that ischaemic heart disease (IHD) ranked at the top of 291 diseases in terms of global disease burden expressed by disability-adjusted life years (DALYs), and had increased by 29% from 1990(Murray et al., 2012). The American Heart Association statement based on a comprehensive review of scientific papers published between 2004 and 2009, concluded that particulate air pollution may have a role on

IHD(Brook et al., 2010). Few studies have focused on incident cardiovascular disease rather than mortality in association with ambient air pollution.

*Studies in America: findings from three women cohorts*

The main suggested evidence of long-term air pollution effects on incident cardiovascular disease in America was based on three women cohort studies.

The Women's Health Initiative (WHI) was one of those which first reported long-term effects of ambient air pollution on incident cardiovascular events(Miller et al., 2007). In this WHI study, conducted by Miller et al, 65,893 postmenopausal women aged 50-79 years from 36 metropolitan areas in America were enrolled during 1994-1998 and followed through to August 2003. It was observed that, for each 10 $\mu g/m^3$ increase in $PM_{2.5}$, risks for the first cardiovascular event increased by 24% (95%CI: 9 to 41%), first coronary heart disease event by 21% (4 to 42%), first cerebrovascular disease event by 35% (95%CI: 8 to 68%) and first stroke by 28% (95%CI: 2 to 61%). There was no statistically significant association between $PM_{2.5}$ and first myocardial infarction (MI) event alone. The authors also concluded that within-city effect estimates were larger than between-city effects for $PM_{2.5}$. This finding, in line with a previous report for $NO_2$ in Canada(Crouse et al., 2015), suggested that within-city exposure contrast may be more important to detect an association if any.

This important study had several strengths, including considering many potential confounders and the robust ascertainment of outcomes through medical records. Although annual mean $PM_{2.5}$ was assigned to each women based on levels recorded at the nearest monitoring site based on their postcode, and is likely therefore to introduce exposure error, the study nonetheless provides compelling evidence that long-term exposure to fine particle is associated with incident cardiovascular events in this subpopulation.

Unlike the WHI study, which used a single year average estimate of air pollution, the California Teachers Study (CTS) Cohort estimated monthly average $PM_{10}$ and $PM_{2.5}$ concentrations via inverse

distance-weighting from nearest monitoring sites during 1996-2005 and then linked the cumulative exposure to incident cardiorespiratory diseases(Lipsett et al., 2011). None of $PM_{2.5}$, $PM_{10}$ or $NO_2$ was associated with MI incidence among over 120,000 female teachers; $PM_{10}$ but not $PM_{2.5}$ was associated with stroke incidence (HR: 1.06, 95%CI: 1.00 to 1.13). This study did not find any associations between all-cause mortality, cardiovascular mortality, cerebrovascular mortality and any of the exposures. The simple inverse distance-weighted interpolation method used in this study may in part explain the observed weak associations, particularly for $NO_2$ which is mainly driven by local traffic volumes and may vary considerably across a short distance, variability not likely to be captured by the distance-weighted interpolation approach.

Another USA study of 66,250 female nurses, using model-derived individual air pollution estimates, observed no significant association between $PM_{10}$ and incident non-fatal coronary heart disease events during a 10-year follow-up(Puett et al., 2008). However, fatal CHD was significantly associated with increased $PM_{10}$. One limitation of this study is that it did not adjust for individual-level socioeconomic status, although the authors did include two indicators of area-level socioeconomic status. Also, some of the outcomes were derived from self-reported information, which may have introduced bias, such as underreporting. Both this study, and the WHI study, reported stronger associations between PM exposure and cardiovascular incidence with increasing levels of obesity and in those who were never-smokers. Future studies are needed to replicate these findings.

All the above studies in America were conducted in specific populations (postmenopausal women, female teachers and female nurses), therefore results were not generalisable to other group. One study of general populations has reported that higher traffic density near home, as a proxy of ambient traffic-related air pollution, was associated with higher incidence of CHD(Kan et al., 2008). Stronger associations were found in men and current/ex-smokers, but no effect modification was observed for BMI.

*Studies in Europe*

In Europe, only a few studies were conducted, with some focusing on either $NO_2$ or particulate matter only and others focusing on both exposures.

In Greece, in a small study of 2752 participants, it was estimated that risk of an incident cardiovascular event was increased by 50% (95%CI: 5% to 116%) per each 10 μg/$m^3$ increase in $PM_{10}$ but not $NO_2$ (Katsoulis et al., 2014). The study also found that both $PM_{10}$ and $NO_2$ were positively associated with incident IHD among women but not men, with the association with $NO_2$ being statistically significant (HR: 1.54, 95%CI: 1.01 to 2.37).

A meta-analysis of three large cross-sectional health surveys in England indicated a positive although not statistically significant association between elevated $PM_{10}$ and increased prevalence of self-reported doctor-diagnosed cardiovascular diseases in men and women (2.9% (95%CI: -0.6 to 6.5%) and 1.6% (95%CI: -2.1 to 5.5%) respectively)(Forbes et al., 2009a). In an English cohort of over 830,000 patients drawn from general practice databases, Atkinson et al did not find any significant associations with $PM_{10}$ or $NO_2$ and incident MI or stroke except for heart failure(Atkinson et al., 2013). One of the reasons to explain the inconsistencies with previous studies in North America might be due to the fact that modelled within study variations in levels of particulate air pollutants were much lower in the UK, making it difficult to observe an effect.

Weak associations between $NO_2$ and incident cardiovascular events were reported in two studies. In a study of all residents in Rome, Rosenlund et al found a small increase of 3% (95%CI: 0 to 7%) in risk for incident coronary event per 10 μg/$m^3$ increase in $NO_2$(Rosenlund et al., 2008). In a Danish cohort of 52,215 participants, a weak association was found between $NO_2$ and incident stroke (HR: 1.05, 95%CI: 0.99 to 1.11); and it was reported (for the first time) that associations were stronger for ischaemic stroke than haemorrhagic stroke(Andersen et al., 2012a). However, a case-control study in the Stockholm County did not report associations between $NO_2$, $PM_{10}$ and first-time MI cases(Rosenlund et al., 2006).

*The ESCAPE findings*

The most recent evidence of long-term air pollution effects on cardiovascular morbidity is from the ESCAPE project. In an analysis involving 99,446 participants from 11 European cohorts, Stafoggia et al found that $PM_{2.5}$ was not associated with increased risk for incident stroke (HR: 1.05, 95%CI: 0.88 to 1.62)(Stafoggia et al., 2014). No significant association was found for $NO_2$ either. However, it is of note that in this study, the authors found stronger significant effects in older people (aged>60 years) and never-smokers, indicating some modification effects exist. Another ESCAPE analysis showed that a 10 $\mu g/m^3$ increase in $PM_{10}$ was associated with 12% (95%CI: 1 to 25%) increase in risk for incident acute coronary events (MI and unstable angina)(Cesaroni et al., 2014). A similar association for $PM_{2.5}$ was also observed but it was not significant (HR: 1.13, 95%CI: 0.98 to 1.30). Importantly, both ESCAPE analyses also reported positive significant associations at levels of air pollution below the current annual European limits (25$\mu g/m^3$ for $PM_{2.5}$ 40 $\mu g/m^3$ for $PM_{10}$)(European Union, ), indicating an urgent need to review the current standards of air quality in Europe.

### *Summary*

In summary, while it is well documented that long-term ambient air pollution (both particulate matter and $NO_2$) increased the overall mortality risk and specifically mortality from cardiovascular causes, its effects on cardiovascular morbidity are less certain, partly due to a smaller number of available studies. The current evidence is suggestive of an air pollution effect on incident stroke and incident MI, but to understand which pollutants drive these effects and in which susceptible groups, needs further investigations.

It should be noted that it is difficult to compare results across previous studies covering a span of 20 or more years, due to changes in air pollution sources and improvements in approaches to exposure assessment. For example, in the 1990s, air pollution estimates based on nearest or central fixed-site regional-level monitoring stations assigned to each participant based on their home addresses or postcode. In more recent years dispersion model or land use regression model derived estimates at a

finer spatial scale were assigned to individual home address. Also, adjustment for confounding factors, methods of outcome ascertainment, and study populations differed across different countries and studies, which may contribute to the heterogeneity across findings.

### 1.1.2 Long-term ambient air pollution and biochemistry markers

Although air pollution has been linked to cardiovascular diseases in various populations as reviewed in the last section, the underlying mechanisms behind these associations are not fully understood. Proposed mechanisms include, but are not limited to, systemic inflammation, systemic oxidative stress, vascular dysfunction, atherosclerosis, epigenetic modifications and promotion of traditional risk factors such as high blood lipids and diabetes(Brook et al., 2010). In the following sections, experimental and epidemiological evidence with regards to associations between air pollution and some common biochemistry markers of cardiovascular diseases, as well as the possible underlying mechanisms for each association, are reviewed.

#### *C-reactive protein (CRP)*

According to Brown et al, fine particle ($PM_{2.5}$) can travel deep into respiratory tract and initiate a local immune reaction(Brown et al., 2001), inducing the release of pro-inflammatory cytokines from lung cells. Also, because of the tiny size of these particles, they are able to enter the circulation system by crossing cell membranes, where the same inflammation reactions occur in different tissues throughout the body.

C-reactive protein (CRP) is a protein which has been used, as one of many possible markers, to reflect the presence and intensity of systemic inflammation(Pope, 2001). It is widely measured in clinical settings and levels of CRP are a predictor of both manifested cardiovascular diseases (e.g. MI, stroke)(Ridker, 2003) and sub-clinical markers for CVD such as atherosclerosis(Libby, 2002). It is therefore hypothesised that effects of air pollution on cardiovascular diseases may act partially via the pathway of systemic inflammation, as reflected by elevated CRP level.

Serum CRP concentration increases rapidly from low, normal levels in response to stimuli such as inflammation(Li et al., 2012). Animal studies provided compelling evidence that air pollution, particularly PM-related air pollution, increases systemic inflammation(Rohr et al., 2010; Upadhyay et al., 2010; Niwa et al., 2008). It should be made clear that the exposure routes, doses and reactions against exposure differ between these animal experiments and typical human exposure scenarios. Therefore it is difficult to judge the potential effects on humans from experiments on animals, although these studies do provide hints for further epidemiological investigations in humans.

Stronger and more consistent evidence of harmful effects of ambient air pollution on CRP level are seen in children compared to adults(Li et al., 2012). In adult populations, findings were inconsistent across a number of cross-sectional and panel studies(Li et al., 2012). Short-term effect of PM-related air pollution on CRP level has been investigated in panel studies of various populations(Ruckerl et al., 2014; Khafaie et al., 2013; Tsai et al., 2012; Huttunen et al., 2012; Rudez et al., 2009; Panasevich et al., 2009; Steinvil et al., 2008; Sullivan et al., 2007; Chuang et al., 2007; Ruckerl et al., 2007; Ruckerl et al., 2006; Pope et al., 2004b), with sample sizes ranging from less than 10 individuals to several dozen. Among these, some reported an association between increased levels of air pollution and elevated levels of serum CRP(Ruckerl et al., 2014; Huttunen et al., 2012; Chuang et al., 2007; Ruckerl et al., 2006; Pope et al., 2004b), though associations depended on what time lag was chosen in the analysis and which air pollutant was assessed. For example, one study observed associations with $PM_{10}$ but not $PM_{2.5}$(Ruckerl et al., 2006). Other studies did not observe a statistically significant association, although results were suggestive of an effect of air pollution on systemic inflammation, as indicated by markers other than CRP(Khafaie et al., 2013; Tsai et al., 2012; Rudez et al., 2009; Panasevich et al., 2009; Steinvil et al., 2008; Sullivan et al., 2007; Ruckerl et al., 2007). One should be aware that most research has focused on short-term effects (i.e. within days of air pollution exposure) on levels of CRP, though this is likely appropriate as elevated CRP levels represent an acute response to air pollution or other stimuli. However, it is also likely that longer-term exposure to air pollution may induce a chronic

low-grade systemic inflammation, which has been suggested a role in the pathogenesis of heart diseases(Kaptoge et al., 2014).

Medium-term (weeks to months) and long-term (years) average outdoor air pollution levels in relation to levels of CRP have been assessed in only a few studies, and results are less clear. Zeka et al found a positive, though not significant, association between CRP levels and a 4-week averaged $PM_{2.5}$(Zeka et al., 2006). Diez-Roux et al also reported weak, non-significant associations between CRP concentrations and 30-day mean (1.03, 95%CI: 0.98 to 1.10) or 60-day mean (1.04, 95%CI: 0.97 to 1.11) $PM_{2.5}$ level prior to blood extraction(Diez Roux et al., 2006).

In a German cross-sectional study which included nearly 5,000 individuals it was reported that every 3.9 $\mu g/m^3$ increase of annual $PM_{2.5}$ was associated with a 24% (95%CI: 4 to 47%) increase of CRP level in men but not in women(Hoffmann et al., 2009). However, this association was sensitive to adjustment for short-term $PM_{10}$ exposure. A longitudinal analysis with repeated measurements of CRP was conducted later in the same cohort(Viehmann et al., 2015). It was found that a 2.4$\mu g/m^3$ increase of annual mean $PM_{2.5}$ was associated with 5.4% (95%CI: 0.6% to 10.5%) increase of CRP, independent of short-term air pollution. An extended analysis of the same study populations has confirmed that $PM_{2.5}$ from traffic sources was strongly associated with CRP levels. $PM_{2.5}$ from industry sources was not associated with CRP(Hennig et al., 2014). In an analysis of 1923 mid-life women in USA, Ostro et al observed that prior-year $PM_{2.5}$ exposure was strongly associated with increased CRP levels, particularly in older diabetics, smokers and unmarried persons(Ostro et al., 2014).

Relatively fewer studies have investigated the role of $NO_2$ on CRP levels in adults. No statistically significant associations were found between long-term exposure to $NO_2$, $PM_{10}$ and CRP in two cross-sectional surveys conducted in general populations in England, which led the investigators to conclude that systemic inflammation might not play a role in the link between air pollution and cardiovascular diseases(Forbes et al., 2009b). In a study in Stockholm, long-term exposure (assessed over by 1, 5 or 30 years) to traffic-related $NO_2$ was not associated with CRP levels(Panasevich et al., 2009).

The two largest studies so far are the ESCAPE and MESA (Multi-Ethnic Study of Atherosclerosis) studies. In the ESCAPE cross-sectional analysis, which involved 22,561 participants from six European cohorts (all from central and Northern Europe), annual exposure to particulate air pollution was not associated with increased levels of CRP, but high traffic intensity on the nearest road was(Lanki et al., 2015). Higher NOx was associated with increased CRP in the main model, but significance was lost after further adjustments for area-level socioeconomic status. The authors argued that exposure to locally emitted near-road traffic exhausts may have contained more toxic substances, which might explain why associations were only seen for traffic indicators such as traffic intensity or NOx, the latter which may serve as an indicator of vehicle exhausts. In the MESA study of 11,190 participants in America, while the authors found some positive significant associations between other inflammation markers and annual averaged $PM_{2.5}$, association was not found for CRP(Hajat et al., 2015). Only the $PM_{2.5}$ level measured on the day of blood extraction was associated with CRP levels.

Studies on chronic effects of air pollution on serum CRP are still relatively rare and current results are mixed. Overall, in the extensive review in 2010 carried out by Brook et al, the epidemiological evidence of long-term effects of air pollution on systemic inflammation biomarkers was marked as "*limited or weak*"(Brook et al., 2010). Further research is warranted regarding the effects of chronic exposure to particulate air pollution or $NO_2$ on levels of serum CRP, particularly with a focus on effects in different susceptible groups of the general populations.

### *Blood lipids*

Emerging studies have suggested an association between air pollution and atherosclerosis, assessed by intima-medial thickness (IMT)(Adar et al., 2013; Kunzli et al., 2010), yet few studies have investigated directly the link between air pollution and blood lipids level, which is also a risk predictor for atherosclerosis and the subsequent development of cardiovascular diseases.

Experimental studies in animals have suggested that exposure to air pollution might change the blood lipid profiles adversely. A novel study carried out by Suwa et al in 2002 reported that exposure to $PM_{10}$

was associated with increases not only in extracellular lipid pools but also in the total amount of lipids in aortic lesions in rabbits(Suwa et al., 2002). In a mouse model study conducted by Sun et al a 1.5 fold increase in lipid content in the aortic arch was observed in mice that were high-fat fed and exposed to $PM_{2.5}$ for a total period of six months, compared with those which were also high-fat fed but exposed to filtered air(Sun et al., 2005). These studies, among others(Soares et al., 2009; Araujo et al., 2008; Nonogaki et al., 1995), suggested that air pollution-induced systemic inflammation can alter lipid metabolism, lipid oxidation and accelerate hepatic secretion of triglycerides.

Total cholesterol (TC), low-density lipoprotein cholesterol (LDL-C) and high blood pressure were identified as risk factors for cardiovascular diseases in the Framingham Heart Study(Wilson et al., 1998). However, there is a paucity of data linking long-term air pollution to blood lipids in humans. A few studies have examined the role of short-term air pollution on blood lipids levels.

A study conducted in Taiwan with nearly 7,600 participants (70% non-smokers) reported that every interquartile increase (34 $\mu g/m^3$) of 1-day averaged $PM_{10}$ was associated with a reduction in high-density lipoprotein cholesterol (HDL-C) by 0.90 mg/dL (95%CI: 0.34 to 1.46)(Chuang, Yan & Cheng, 2010). However, no statistically significant findings were reported for either LDL-C or triglycerides (TG). This is one of the first studies of its kind conducted in a relatively large population. Two earlier studies reported similar results in different research settings. One, a case-control study, compared blood lipid levels of 118 traffic police personnel and 118 office personnel in Italy(Tomao et al., 2002). Case and controls were matched by age and time in service. Increased levels of HDL-C and TG were observed in the traffic police personnel group, leading the investigators to conclude that those frequently exposed to traffic-related air pollution might suffer from dyslipidaemia. The other, a panel study, involved only 12 adult asthmatics in North Carolina followed for a 12-week period(Yeatts et al., 2007). It was found that every 1 $\mu g/m^3$ increase of PM coarse was associated with small increases in LDL-C (1.2%, 95%CI: 0.3 to 2.0%) and TG (4.8%, 95%CI: 0.8 to 8.7%) respectively. There were no statistically significant associations found for $PM_{2.5}$. The authors suspected that coarse PM contained

rich biologic materials such as endotoxin and therefore may play a greater role in changing the lipids profile(Poursafa et al., 2014a). More research is needed to replicate this finding, particularly in a large general population.

More recently, results from two large population-based studies were published. A study of nearly 40,000 adults in Copenhagen aged 50-64 years reported statistically significant associations between non-fasting total cholesterol and both ESCAPE-LUR modelled $PM_{2.5}$ (0.101 mmol/L, 95%CI: 0.028 to 0.173, per 5 $\mu g/m^3$) and dispersion modelled $NO_2$ (0.026mmol/L, 95%CI: 0.008 to 0.045, per 10 $\mu g/m^3$)(Sorensen et al., 2015). However, data on other lipid measures were not available.

In the Third National Health and Nutrition Examination Survey (NHANES III, 1988-1994) in America, it was found that among 11,623 adults (median age: 41 years) long-term inter-quartile higher $PM_{10}$ air pollution was significantly associated with greater total cholesterol levels (1.43%, 95%CI: 1.21% to 1.66%) and triglycerides levels(2.42%, 95%CI: 1.09% to 3.76%)(Shanley et al., 2016). Air pollution exposure in the NHANES III study was assigned from monitoring networks based on the centroid of resident census blocks of the participants, which was likely to introduce more exposure errors compared to the model-based approaches at individual address level used in the Copenhagen study. In additions, estimates for other PM size fractions and $NO_2$ were not available for the NHANES III study.

For both studies in Copenhagen and America, albeit of a cross-sectional design, it was suggested that a possible dose-response relationship exists between long-term air pollution and total cholesterol level.

### *Blood glucose*

Glucose intolerance is another established risk factor for cardiovascular diseases, as noted in the Framingham Study(Kannel, 1990). A meta-analysis of randomised controlled trials on the effects of glucose control on cardiovascular outcomes concluded that intensive glycaemic control significantly reduces coronary events(Ray et al., 2009). However, it remains uncertain whether glucose impairment

mediates the association between exposure to long-term ambient air pollution and development of cardiovascular diseases.

Sun and colleagues carried out an experiment investigating the link between air pollution and insulin resistance in mice(Sun et al., 2009). It was demonstrated that whole-body insulin resistance was linked with exposure to $PM_{2.5}$. In population studies, the link between air pollution and blood glucose remains obscure, partly because there are not many epidemiological studies addressing this link.

A study of 363 women in the SALIA cohort in Germany reported some suggestive evidence that exposure to $NO_2$ may contributed to impaired glucose metabolism(Teichert et al., 2013). However, as the authors acknowledged, results from this cross-sectional study were limited as it was conducted in a small sample of elderly women, which may be subject to survivor bias. This association certainly needs to be confirmed in future studies with a much larger sample drawn from wider population groups. Another study, among 9,102 newly diagnosed diabetes patients, found that in both sexes adjusted HbA1c (Haemoglobin A1c) was significantly lower in those exposed to the lowest quartile of $PM_{10}$(Tamayo et al., 2014). HbA1c is often used as a marker of glucose control and is associated with risks for developing hard arterial plaques(Jorgensen et al., 2004). To date, only one study was conducted in the general population to access the association between long-term air pollution and HbA1c(Liu et al., 2016). This particular cross-sectional study in China reported that among 11,847 adults, each inter-quartile (41.1 $\mu g/m^3$) higher annual mean $PM_{2.5}$ was significantly associated with elevated levels of HbA1c (0.08%, 95%CI: 0.06% to 0.10%). $PM_{2.5}$ levels were estimated at home address level using a satellite-based spatial statistical model. Within-study exposure contrast for $PM_{2.5}$ was large in this study, contributing to the observed significant association.

Increasing numbers of studies have suggested an association between long-term air pollution and development of type 2 diabetes in humans. For example, a study conducted by Brook et al in Canada, examined cross-sectional associations between air pollution and prevalence of diabetes among 7,600 patients from two respiratory clinics(Brook et al., 2008). Significantly increased prevalence of diabetes

(OR: 1.04, 95%CI: 1.00 to 1.08) was found in women but not in men for each 1ppb increase of $NO_2$ exposure. A Danish prospective cohort also found a small increase in risk for incident diabetes cases in association with $NO_2$ exposure; with effects stronger in people with relatively healthy lifestyles such as non-smokers and physically active people(Andersen et al., 2012b). A study in America reported that diabetes prevalence increased with $PM_{2.5}$ levels(Pearson et al., 2010). This finding was further supported by some mechanistic investigations, which found exposures to $PM_{2.5}$ could reduce insulin sensitivity among healthy adults(Brook et al., 2013; Xu et al., 2011). A German study reported that diabetes incidence increased by 15-42% per interquartile higher of $NO_2$ and several traffic-related PM exposures; most associations were statistically significant(Kramer et al., 2010). However, a study in the Netherlands did not observe an association between air pollution and diabetes(Dijkema et al., 2011).

A recent meta-analysis including studies conducted in Europe and North America concluded that both $PM_{2.5}$ (based on three eligible studies) and $NO_2$ (based on four eligible studies) were associated with increased risk of type 2 diabetes(Eze et al., 2015). However, the authors acknowledged the high diversity among the included studies with unclear and high risk of bias. The authors also highlighted the importance for future studies to apply comparable models in assigning exposures to participants and to report the scales of exposure assessment.

### *Summary*

In summary, data on the effects of long-term ambient air pollution on some CVD intermediate biochemistry markers are still limited in population studies. Studies investigating direct associations between long-term air pollution and serum CRP levels are far from conclusive and needs more investigations. In the last few years, evidence is starting to emerge regarding the long-term air pollution effects on blood lipid and blood glucose levels, which all suggested a positive relationship. More research, especially in large-scale population studies, is warranted to provide more robust evidence and to explain the likely biological mechanisms underlying associations between air pollution and CVD outcomes.

## 1.2. Environmental noise pollution

Environmental noise is defined by the EU Environmental Noise Directive (END)-2002/49/EC as "*unwanted or harmful outdoor sound created by human activities, including noise emitted by means of transport, road traffic, rail traffic, air traffic and from sites of industrial activity*"(Directive 2002/49/EC, 2002). Noise was regarded to be a major problem by more than half of city dwellers in many European cities surveyed about quality of life (with the figure ranging by city from 51% in Rotterdam to 95% in Athens)(European Commission, 2009). A recent survey conducted by the European Commission in 2010 showed that 44% of Europeans believed that noise affects their health to "*a large extent"*, an increase of 3% since the last survey in 2006(European Commission, 2010). Transport is still the main source of noise in urban areas(European Environment Agency, 2013), and noise from road traffic may potentially affect a large part of the population. In fact, a report published by European Environment Agency (EEA), has revealed that around 70 million people in the largest European cities were exposed to long-term road traffic noise above 55 decibels (dB) (the EU threshold)(European Environment Agency, 2012).

Environmental noise is commonly measured in A-weighted decibel (dB(A)) level. dB is a logarithmic scale to measure sound pressure level and A-weighted is a frequency-weighting of sound pressure levels that simulates the subjective response of the human ears(Basner et al., 2014). Different noise metrics have been used for different purposes(Passchier-Vermeer & Passchier, 2000). The maximum A-weighted sound level, Lmax, measures the highest level of noise in a single event (e.g. a single aircraft noise) to reflect how intrusive the noise is. However, Lmax does not take into account the duration of a noise event. Another noise metric for a single event is sound exposure level (SEL), which measures both duration and magnitude of an entire noise event. In epidemiological studies, the equivalent sound level (Leq), is usually used to represent the average sound level over a given period of time. Leq can be calculated for any time period. For example, Ldn measures day-night average sound level, with a 10dB penalty added for 22:00-07:00 because of the heightened sensitivity to night-

time noise. The common noise metric proposed in the Environmental Noise Directive (END)-2002/49/EC is Lden (day-evening-night 24-hour average sound level, with a penalty of 5 dB added for the evening hours or 19:00 to 22:00, and a penalty of 10 dB added for the night-time hours of 22:00 to 07:00).

Environmental noise is usually not associated with any significant auditory effects because hearing damage will require a higher sound level of noise exposure (i.e. >70 dB) for significant periods of time(Murphy M, 2014). As a result, research on environmental noise has mainly focused on the detrimental effects on non-auditory health. According to the WHO (2011) *Burden of Disease from Environmental Noise* study, it was estimated that each year at least 1 million healthy life years were lost due to environmental noise in the western part of Europe(European Office WHO, 2011). Sleep disturbance and annoyance are the two main outcomes which contribute to the burden of diseases from environmental noise. In addition, other outcomes such as cognitive impairment in children and cardiovascular diseases in adults are of increasing concern. Within the scope of this project, evidence on effects of road traffic noise on adult cardiovascular outcomes (ischaemic heart disease and stroke) is presented.

### 1.2.1. Long-term road traffic noise and cardiovascular diseases

Ischaemic heart disease (IHD) and hypertension are the two most frequently examined cardiovascular outcomes in relation to road traffic noise exposure in epidemiological studies(European Office WHO, 2011). There have been fewer studies on other outcomes (e.g. stroke) but increasing evidence is starting to emerge. In general, there is increasing evidence in support of an effect of road traffic noise on the risk for IHD, hypertension and possibly stroke.

One of the underlying biological mechanisms for the associations between noise and cardiovascular diseases is that noise acts as a general environmental stressor, activating the autonomous nervous system and the endocrine system(Maschke, Rupp & Hecht, 2000; Lercher, 1996). Long-term exposure to noise might persistently activate these systems and lead to metabolic dysfunction (e.g. increased

blood pressure and blood glucose) and subsequent cardiovascular diseases may occur. Direct and indirect pathways linking noise exposure and cardiovascular endpoints are shown in Figure 1.1(Babisch, 2014).

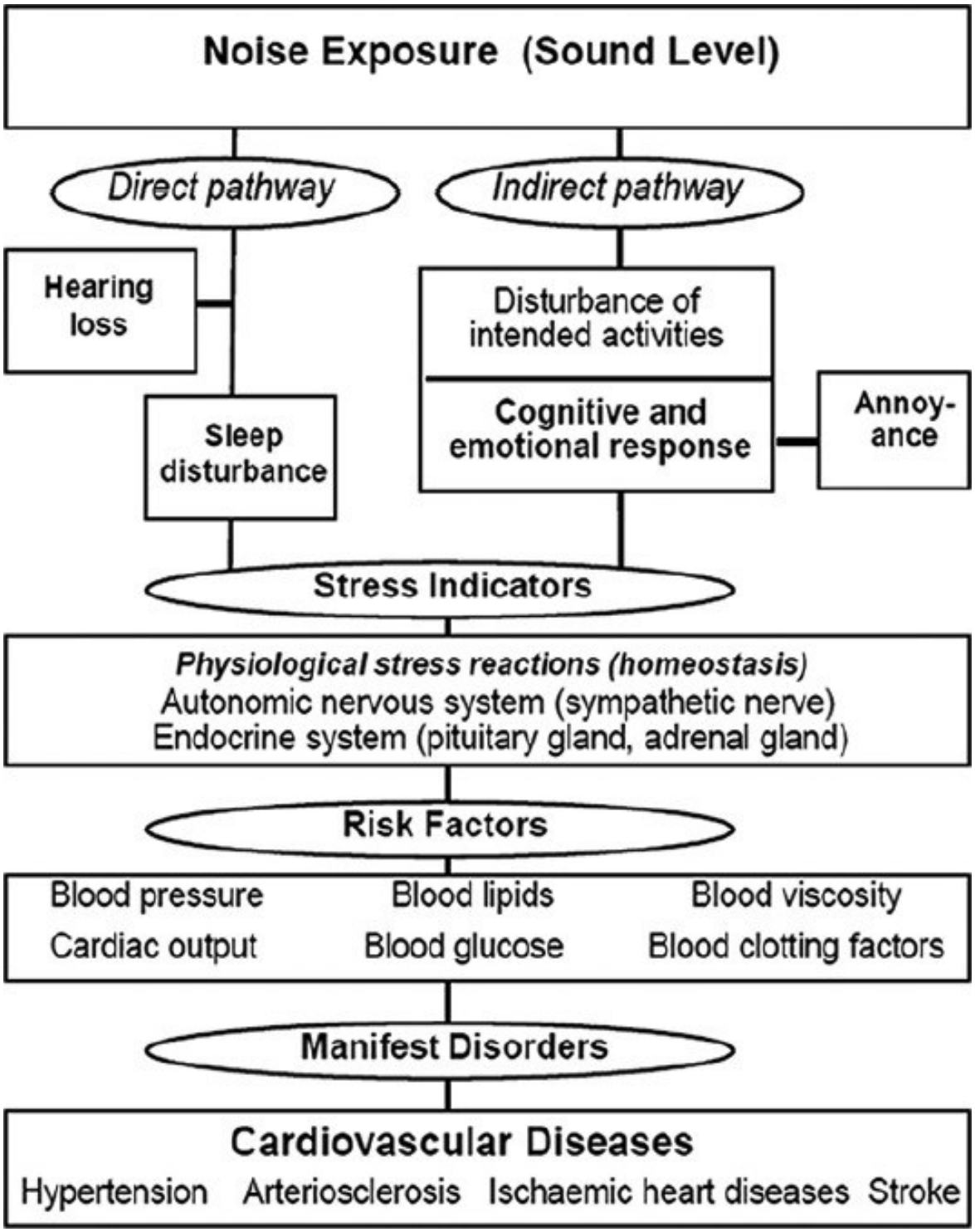



Figure 1.1 Possible biological mechanisms between noise exposure and cardiovascular diseases as published by Babisch(Babisch, 2014).

### *Ischaemic heart disease*

Most studies published before the year 2008 used a comparable daytime (16hour) noise indicator. But since 2008, increasing numbers of noise studies have been published using other noise indicators. Therefore, I grouped the reviewed literatures into two sections: published before and after year 2008.

#### Studies published up to year 2008

There were two landmark studies of the associations between road traffic noise and cardiovascular health in the late 1980s and early 1990s: the Berlin study(Babisch et al., 1994) and the Caerphilly-Speedwell study(Babisch et al., 1990).

The Berlin study was a prospective case-control study which recruited over 4,000 male participants (aged 31-70 years) from the former West Berlin area to investigate the association between long-term exposure to road traffic noise and incidence of MI. Compared to the lowest noise exposure group (<=60 dB(A), Leq, 6-22 h), the odds ratios (OR) of MI incidence in the highest noise exposure category (71-80 dB(A), Leq, 6-22 h) was 1.2 (95%CI:0.8 to 1.7), although this was not statistically significant. Restricting the sample to those who had lived at the same address for 15 years to reduce exposure error, the OR slightly increased to 1.3 (95%CI: 0.9 to 2.0) in the same highest noise exposure category. This study was limited by several factors, e.g. a men-only sample and crude assessment of noise exposure (the nearest street to home address was assigned a noise exposure based on a city-wide noise map). The same research team retested their hypothesis in a sample which also included female participants and used more refined noise assessment methods. Traffic noise levels were calculated for the most affected facades of each participant's home address for day and night using a noise map. This later study, known as The Noise and Risk of Myocardial Infarction (NaRoMI) study, found a positive dose-response relationship in MI incidence with increasing noise level in men but not in women(Babisch et al., 2005). Those men exposed to the highest noise level (>70dB (A), Leq, 6-22 h) had a very similar OR to the previous study of 1.3 (95%CI: 0.9 to 1.8) when compared with the lowest noise exposure group (<=60 dB (A), Leq, 6-22 h). Statistical significance was only reached in those who remained at the same address for at least 10 years and were exposed to the highest noise level (OR:

1.81, 95%CI: 1.02 to 3.21). This study has several strengths compared to the previous study. Results were adjusted for many potential confounders and/or effect modifiers, noise from rail and air traffic, occupational noise exposure assessed by interviews and noise sensitivity scores. Residence time was considered as a means of limiting the impact of exposure error in this study as stronger associations were seen among participants who remained in the current address for a longer period (e.g. >10 years). In these participants noise exposure is less likely to be misclassified, compared to those who had lived elsewhere in the previous decade. However, this study did not report any effects in women, possibly due to a lack of statistical power.

The second phase of Caerphilly and Speedwell study, comprising a total of 2,512 and 2,348 middle-aged men from south Wales and Bristol respectively, examined the impact of environmental noise on health. Men in the highest noise exposure group (=66-70 dB (A),Leq, 6-22 h) had slightly higher prevalence of IHD compared with those in the lowest noise exposure group of 51-55 dB(A) but an association was not seen with regards to incidence of IHD(Babisch et al., 1993a). In the third phase, after a 10-year follow-up from the recruitment (first phase), increased incidence of IHD was seen in the Caerphilly sample (OR: 1.07 95%CI: 0.60 to 1.91) but not in the Speedwell sample (OR: 0.92, 95%CI: 0.61 to 1.41), comparing those in the higher noise exposure group of 66-70 dB (A) with the lowest noise exposure group of 51-55 dB(A)(Babisch et al., 1999). However, when the two samples were combined and restricted to only those followed from phase two to phase three (follow-up period of 6 years), increased risks were seen in those who provided information on room orientation and window opening habits (OR: 1.31, 95%CI: 0.78 to 2.21) and in those who had a residency at the same address of over 15 years (OR: 1.59, 95%CI: 0.85 to 2.97). This may indicate that room location, window insulation and restricted mobility are possible effect modifiers of the noise effects.

Based on these earlier findings, Babisch conducted several epidemiological reviews of the associations between road traffic noise and cardiovascular outcomes since year 2000. The evidence of road traffic noise on IHD prevalence and incidence is increasing, being marked as “limited/sufficient” in an earlier

review in year 2000(Babisch, 2000) to “sufficient” in an updated paper in 2006(Babisch, 2006). The WHO report based on data from these review papers reported that there was a clear dose-response relationship between risk of MI and road traffic noise level above 60 dB (A)(European Office WHO, 2011). In a meta-analysis of five studies in 2008, Babisch reported a pooled estimate for incident IHD in men of 1.17 (95%CI: 0.87 to 1.57) per 10 dB(A) increase of the day noise (Leq 16h) level at the most exposed façade(Babisch, 2008).

### Studies published after year 2008

Two prospective cohort studies from Scandinavian countries, using a 24-hour weighted noise indicator as calculated by the common Nordic prediction methods in Scandinavia, investigated the association between long-term road traffic noise and incidence of MI.

In Stockholm county, Selander et al reported a non-significant positive association for MI incidence in those exposed to a noise level of ≥50 dB (A), $L_{Aeq}$ (A-weighted equivalent sound level) 24h (OR: 1.12 95%CI: 0.95 to 1.33) and a significant association (OR: 1.38, 95%CI: 1.11 to 1.71) after excluding participants with hearing loss and with other sources of noise exposure (e.g. noise from air traffic, rail traffic, neighbourhood) rather than road-traffic noise(Selander et al., 2009). In Denmark, Sorensen et al reported a clear dose-response relationship based on a cohort of over 50,000 participants(Sorensen et al., 2012a). For each 10 dB increase in road traffic noise ($L_{den}$), risk of incident MI increased by 12% (OR: 1.12, 95%CI: 1.02 to 1.22). Association remained when using the accumulated $L_{den}$ covering the 5-years preceding the diagnosis.

Unlike previous studies, both studies used modelled noise exposures with detailed inputs accounting for residential locations and co-adjusted for $NO_2$ or $NO_X$ in the statistical models. However, results from these two studies were somewhat inconsistent. For example, Sorensen et al reported stronger associations in men, and elderly people (aged 65 years or more)(Sorensen et al., 2012a), whereas Selander et al reported no differences in the effects by sex or age(Selander et al., 2009).

These two studies, together with a few others published after year 2008, were included in an updated meta-analysis conducted by Babisch in 2014(Babisch, 2014). In this updated analysis, including 14 studies of men and women, a significant pooled estimate of the risk for IHD (OR: 1.08, 95%CI: 1.04 to 1.13) per 10 dB(A) increase of weighted day-night road noise level within the range of 52-77 dB(A) was reported. Babisch drew up two major points to consider in future studies related to sex and age.

First, sex may be an effect modifier. Indeed associations with ischaemic heart disease were previously reported in men only, but this is likely to be partly due to the limitation of initial study design (e.g. the Caerphilly-Speedwell study recruited men only) or the fact that cardiovascular diseases are more common in middle-aged men(Babisch, 2008) increasing the statistical power for analyses in men. Babisch suggested that, at this stage, at least we can reasonably assume that the risk of noise on cardiovascular health in men and women should be similar, providing that all related confounding or modifying variables are considered in the analyses(Babisch, 2008). For example, in the analyses of women only, it worth considering the hormonal or menopausal status as a confounding factor along with other usual factors.

Second, age may be potential effect modifier. Sleep disturbance may explain this in part. Elderly people are more susceptible to sleep disturbance at night as sleep structure become more unstable with age(Sateia et al., 2000). Sleep disturbance could also contribute to adverse cardiovascular health(Jackson, Redline & Emmons, 2015a). To this end, it is possible that noise-induced sleep disturbance at night may have effects on cardiovascular health in elderly populations. It is still unclear whether sleep disturbance modifies the association between night-time noise and cardiovascular diseases and, if so, among which population groups. Few studies specifically investigated the effects of night-time noise alone on cardiovascular disease and information about participant's sleep quality is sparse.

Vienneau et al further updated Babisch's 2014 review by including studies on aircraft noise(Vienneau et al., 2015). In this review, a pooled relative risk for IHD was 1.06 (95%CI: 1.03 to 1.09) per 10 dB

increase in noise exposures (both road and air traffic noise). The review reported that a possible threshold of 50dB, above which a linear dose-response was seen. Again, men and those over 65 years of age possibly had a higher risk, but the review was based on a limited number of studies.

A series of review papers were published regarding the cardiovascular effects of traffic noise in several European countries(Argalasova-Sobotova et al., 2013; Belojevic et al., 2011; Bluhm & Eriksson, 2011; Lercher et al., 2011; Kempen, 2011; Stansfeld & Crombie, 2011; Maschke, 2011). Whereas in some countries or regions studies only assessed hypertension in relation to noise(Argalasova-Sobotova et al., 2013; Belojevic et al., 2011; Bluhm & Eriksson, 2011), others also investigated ischaemic heart diseases although inconsistent associations were found across these countries(Lercher et al., 2011; Kempen, 2011; Stansfeld & Crombie, 2011; Maschke, 2011).

### *Stroke*

Few studies have examined the association between long-term road traffic noise exposure and stroke.

In a Dutch cohort, road traffic noise was not associated with cerebrovascular mortality, regardless of adjustment for air pollution(Beelen et al., 2009). Self-reported heart disease and stroke (as a combined outcome) was associated with 24-hour average road traffic noise exposure (OR: 1.19, 95%CI: 1.00 to 1.41) in the Hypertension and Environmental Noise near Airports (HYENA) study(Floud et al., 2013), but an association was not seen in a subsample analysis in which further adjustments were made for air pollution exposures. In the GLOBE study in the Netherlands, no associations were reported between Lden and 'IHD or cerebrovascular disease'(de Kluizenaar et al., 2013).

A prospective cohort study of 57,053 participants (1881 incident stroke cases in 13 years of follow-up) reported the first evidence for a specific effect of road traffic noise on stroke: the incident rate ratio(IRR) for stroke was 1.14 (95%CI: 1.03 to 1.25) per 10 dB(A) increase of road traffic noise ($L_{den}$), independent of NOx exposures and railway and aircraft noise exposures(Sorensen et al., 2011). This study also found that the risk was even higher (IRR: 1.27, 95%CI: 1.13 to 1.43) among those above aged 64.5 years, possibly again because elderly people were more susceptible to noise-induced sleep

disturbance which could contribute to stroke risk. The same research team re-analysed the data by investigating subtypes of stroke in a separate study(Sorensen et al., 2014). Ischaemic stroke was significantly associated with $L_{den}$ (IRR: 1.16, 95%CI: 1.07 to 1.24), possibly because most stroke cases were of the ischaemic type rather than haemorrhage type. This association was independent from air pollution indicated as $NO_2$ or $NO_X$ level.

In a small-area level study in Greater London of 8.6 million residents, daytime road traffic noise increased the risk of hospital admission for stroke in both adults (≥25 years) and elderly (≥75 years) groups, with a relative risk of 1.05 (95%CI: 1.02 to 1.09) and of 1.09 (95%CI: 1.04 to 1.14) respectively, comparing areas exposed to a noise level of >60 dB with those exposed to <55 dB(Halonen et al., 2015). Night-time road traffic noise exposure was associated with hospital admissions for stroke only among the elderly. This study, albeit the largest to date, is ecological study in design, and residual confounding at the individual level may remain.

Recently, two studies have linked aircraft noise with hospital admissions for stroke. In the UK, hospital admission rates for various cardiovascular diseases, including stroke, with regards to exposure to aircraft noise in different boroughs and districts near London Heathrow airport were compared(Hansell et al., 2013). The authors found that in comparison with those experiencing the lowest level of daytime aircraft noise (<51 dB), those in the highest noise exposure group (>63 dB) had an increased risk of hospital admissions for stroke (RR: 1.24, 95%CI: 1.08 to 1.43). Results were robust to adjustments of road traffic noise. In America, a study involving only elderly people (aged ≥65 years) who lived near one of the 89 airports across the country, also found a positive association between aircraft noise and hospital admission for stroke(Correia et al., 2013). Results were also adjusted for road traffic noise and zip-code level air pollution. For both studies, road traffic noise effects were not specifically assessed. Hypertension is one of the most common risk factors for stroke and has been associated with exposure to road traffic noise(van Kempen & Babisch, 2012). Some studies also documented effects of aircraft noise on hypertension. In the HYENA study, a 10-dB increase in night-

time noise was associated with a 14% increase (95%CI: 1 to 29%) of prevalent hypertension among 4,861 residents living near one of the six major European airports(Jarup et al., 2008). A similar finding from France was also reported recently, in which a 10-dB increase in night-time noise was associated with increased risk for hypertension (OR: 1.34, 95%CI: 1.00 to 1.97) in men but not in women(Evrard et al., 2016). A longitudinal study in Sweden reported that for those exposed to a level of aircraft noise greater than 50 dB(A), risk of incident hypertension increased by 19% (95%CI: 3 to 37%), compared to those exposed to a level less than 50 dB(A)(Eriksson et al., 2007).

It should be noted though aircraft and road traffic noise are qualitatively different and therefore may have different health impacts on certain outcomes, e.g. children's reading ability, cognitive performance(Clark et al., 2006). Compared with road traffic noise, aircraft noise is transient, more intense in a short period and usually causes a higher arousal level for areas which are directly under the flight paths. The exposure assessment approaches may be different for both types of traffic noise, for example, data about traffic flow and fleet are usually required for road traffic noise modelling. In addition to that, road noise propagation routes will also need to be considered. Nevertheless, given that aircraft noise is relatively independent from air pollution, more aircraft noise studies are warranted to investigate the independent effect of traffic noise on cardiovascular outcomes.

### *Summary and future research gaps*

In summary, there is accumulating evidence to link road traffic noise exposure to cardiovascular diseases, but evidence remains inconclusive especially regarding, for example, effect modification by certain factors (e.g. room orientation, sex, age), and effects in some vulnerable groups (e.g. elderly people, people with long-term ill-health). Additional issues include dealing with different noise sources and characteristics (e.g. noise from road, rail, aircraft, neighbourhood and occupation) and using different noise indicators (daytime noise, night-time noise, 24-hour weighted noise, etc.).

The ENNAH (European Network on Noise and Health) project brought together researchers from 33 research centres in Europe to carry out comprehensive reviews of noise and health and to draw future

research directions in this field(European Commission, 2013). Certain gaps in current research on cardiovascular effects have been identified, as described below.

1. More evidence is needed to strengthen the previously presented dose-response curves between road traffic noise and cardiovascular endpoints. Prospective cohort studies are needed to provide robust effect estimates.
2. It is still unclear if effect estimates differ by sex, daytime versus night-time noise, in chronically ill participants, and the elderly, and more evidence is needed about potentially important exposure modifiers, including window opening habits and shielding effects. A dedicated study design and a large sample size will be needed to investigate these effects.
3. Studies of combined noise sources other than road traffic noise alone are still very few. Also, noise acoustic characteristics (e.g. sound level, frequency spectrum, time course) need to be specified in future studies(Basner et al., 2014).
4. The role of noise sensitivity or noise annoyance needs to be addressed. Questions about whether noise sensitivity/annoyance serves as an effect modifier or a predictor of CVD morbidity and mortality by itself requires further investigations. In a UK-wide cross-sectional survey, higher noise sensitivity was seen among older participants (middle-aged and above), females, people with a home mortgage and in a higher social class(Van de Ker Ckhove, 2016). In future studies it may worth investigating associations between noise and cardiovascular diseases in these subgroups. In the Whitehall study of British civil servants, noise sensitivity as assessed by a single question in the baseline survey at 1985-1988 was not associated with incident coronary heart diseases during the follow-up period up to 2008-2009(Stansfeld & Shipley, 2015). As this study did not have objective noise estimates, the role of noise sensitivity on the association between road noise and incident coronary diseases could not be assessed. In the HYENA study, where data on both measured noise and noise annoyance were available, association between aircraft noise and hypertension was stronger in more annoyed persons, suggesting effect modification by annoyance(Babisch et al., 2013).

5. More evidence is needed for the associations between road traffic noise and stroke. It would be worthwhile to also examine the specific subtypes of stroke, for example haemorrhage and ischaemic stroke.
6. Co-exposure of road traffic noise and air pollution and their impacts on cardiovascular health has not been extensively explored and inconsistent results were reported across studies.

The emphasis of current research in this field should be placed on finding the potential threshold above which an increase in cardiovascular risk is observed, the magnitude of such an effect, and the susceptible groups(Babisch, 2014).

## 1.3. Joint cardiovascular effects of ambient air pollution and road traffic noise

Road traffic is the main common source of air pollution and noise. Both exposures have been linked to similar cardiovascular endpoints, as detailed in the aforementioned epidemiological studies. It remains unclear though whether these two exposures contribute to cardiovascular effects independently or act as a confounder for each other in the causal link with cardiovascular endpoints(Foraster, 2013).

*Studies on cardiovascular mortality and ischaemic heart disease*

There are increasing numbers of studies, mainly in Europe, that have investigated both exposures with respect to cardiovascular diseases in the same study.

Beelen et al studied the joint effect of air pollution and noise on cardiovascular mortality in a Dutch cohort (N=117,528) during a 9-year follow-up period(Beelen et al., 2009). Both air pollution and noise exposure estimates were model-derived and assigned to participant's home address at baseline. Air pollution indicators, namely background black smoke and traffic intensity on the nearest road, and traffic noise ($L_{den}$) were assessed separately and jointly in the statistical models. Correlations between noise and air pollution measures were moderate (r=0.24 for black smoke, r=0.30 for traffic intensity

on the nearest road). Both traffic intensity and the highest traffic noise exposure group (>65 dB (A)) were positively associated with increased mortality from total cardiovascular diseases and heart failure, and significance was reached for those with the highest noise exposure. However, when mutually adjusted, risk ratios for traffic intensity remained similar while significance for highest noise exposure was lost, indicating a confounding effect by air pollution.

Mortality from overall cardiovascular diseases in association with air and noise pollution was studied in a Danish cohort(Raaschou-Nielsen et al., 2012). The study followed over 50,000 participants aged 50-64 years for up to 16 years and traced participant's residential address since 1971. Model-derived $NO_2$ concentration at all addresses since 1971 was time-weighted and road traffic noise (Lden) was calculated for baseline residence based on the Nordic prediction method. It was reported that, in the full model which was adjusted for several personal and lifestyle characteristics, mortality from cardiovascular diseases increased by 33% (95% CI: 16 to 54%) per doubling of $NO_2$ concentration. Following further adjustment for road traffic noise, although the point effect estimate decreased slightly, the significant association with $NO_2$ remained (26%, 95%CI: 6 to 51%), suggesting effects of $NO_2$ on cardiovascular mortality were independent of road traffic noise.

In Stockholm, incidence of myocardial infarction increased with increased long-term exposure to road traffic noise above 50 dB (A) (OR: 1.12, 95%CI: 0.95-1.13) after accounting for $NO_2$ exposure(Selander et al., 2009). In the Netherlands, de Kluizenaar et al found that either road traffic noise or air pollution was associated with hospital admissions for IHD, after mutual adjustment(de Kluizenaar et al., 2013). The series of ESCAPE studies also reported that associations between long-term traffic-related air pollution and different cardiovascular endpoints were independent of environmental noise, although the latter exposure was defined locally in each individual cohort(Dimakopoulou et al., 2014; Cesaroni et al., 2014; Stafoggia et al., 2014).

A population-based cohort (N=445,868) was linked to a health insurance database covering nearly all residents in the metropolitan area of Vancouver(Gan et al., 2012). After eight years of follow-up, a 10

dB(A) increase of noise was associated with 9% increase (95%CI: 1 to 18%) in mortality from coronary heart disease (CHD), adjusting for covariates and several air pollutants ($PM_{2.5}$, $NO_2$, black carbon). When assessed as a categorical variable, those exposed to the highest level of noise (>70 dB (A)) had consistently increased mortality in a sequence of nested models, compared to those in the lowest noise level group (≤ 58 dB (A)). In the fully adjusted model including air pollutants, a 22% increase of deaths (95% CI: 4 to 43%) was observed in those exposed to noise level of above 70 dB (A). Black carbon was found to be independently associated with CHD mortality as well. Although confounders such as smoking and individual socioeconomic status were not available for analysis, and there were some limitations in exposure assessment (e.g. noise exposure was estimated at community-level), the authors concluded that both exposures (noise and black carbon) were likely contributing to the CHD mortality independently.

*Studies on other cardiovascular outcomes*

Other than studies on cardiovascular mortality and ischaemic heart disease, some studies have investigated associations between other cardiovascular outcomes and both air and noise pollution.

As discussed in section 2.1.2, Floud et al reported that the association between self-reported heart disease and stroke (as a combined outcome) and 24-hour average road traffic noise exposure may have been confounded by air pollution(Floud et al., 2013). In contrast, Sorensen et al found a 14%~16% increase in stroke incidence for each 10 dB increase in road traffic noise, after adjusting for noise from railway and air traffic and also $NO_2$ or $NO_x$(Sorensen et al., 2011).

Fuks et al reported that increased mean systolic and diastolic blood pressure was associated with long term exposure to PM-related air pollution in a German population-based sample, independent from road traffic noise(Fuks et al., 2011). After additionally adjusting for neighbourhood noise exposures (<70 dB (A) vs. ≥ 70 dB (A)), Coogan et al found increased level of NOx was significantly associated with incident hypertension (RR: 1.14, 95%CI: 1.03 to 1.25, per interquartile change) among black women

living in Los Angeles(Coogan et al., 2012). In a population-based cohort in Spain, a 10 μg/$m^3$ increase in annual averaged $NO_2$ was associated with 1.34mmHg (95%CI: 0.14 to 2.55) higher SBP, adjusting for noise from road, rail and air traffic(Foraster et al., 2014). Sorensen et al found that long term exposure to NOx was inversely associated with self-reported hypertension in a study after adjusting for road traffic noise(Sorensen et al., 2012b).

A more recent study of 4238 adults in Germany showed that both long-term exposures to $PM_{2.5}$ and night-time noise were significantly associated with subclinical atherosclerosis, measured by thoracic aortic calcification (TAC)(Kalsch et al., 2014). This study suggested an independent effect of each exposure on atherosclerosis.

In summary, most the above studies have suggested an independent effect of both traffic-related air pollution and noise on cardiovascular outcomes, with only two(Floud et al., 2013; Beelen et al., 2009) of the reviewed studies suggesting a confounding effect.

*Current knowledge and research gaps*

A systematic review based on only nine publications up to 2013 has inconclusively suggested that confounding effect is likely minimal, but it should also be noted that the review reported heterogeneity not only across studies but also across areas within a study(Tetreault, Perron & Smargiassi, 2013). Nevertheless, the review has raised several issues which need to be explored further in future studies.

1. It was found that correlations between noise and air pollution exposure do not necessarily influence the confounding effects in the reviewed studies. The authors argued that the wide range of correlations observed in the reviewed studies partly reflects different urban structures across the study areas and therefore any confounding effect is perhaps a study-specific issue.

2. To better account for the confounding effect by road traffic noise, it was suggested that traffic-related indicators of air pollution (e.g. $NO_2$, NOx) should be used.
3. More studies are needed to investigate the potential confounding or independent effect in different types of cardiovascular diseases (IHD, stroke, hypertension etc.).
4. Compared to outdoor air pollution, noise is generally considered as a nuisance for many people and could be modified by personal efforts such as wearing ear plug to mitigate excessive noise(Foraster, 2013). In addition, perception of noise is also affected by hearing impairment. These differences could potentially influence the confounding effect and should be better addressed in future studies.

# Chapter 2 Cohort descriptions and data harmonisation

## 2.1. Cohort profiles

Four European population-based cohorts contributed to this project. Two were from the United Kingdom, namely *EPIC-Oxford*(Davey et al., 2003) and *UK Biobank*(Allen et al., 2014); one was from the Netherlands, the *LifeLines* cohort(Scholtens et al., 2015); one was from Norway, the *HUNT* study(Krokstad et al., 2013). These four cohorts were all partners in the BioSHaRE (Biobank Standardisation and Harmonisation for Research Excellence in the European Union) consortium. They were selected for this PhD project for several reasons: i) they are the largest and most recently established population-based cohorts in their countries; epidemiological analyses on air pollution or noise are still limited in these cohorts; ii) data for exposure modelling already exist or are able to be obtained within the timeframe of this project; iii) together, they increase the environmental exposure contrasts that are usually limited within specific geographical regions and landscapes. Detailed descriptions of each cohort profiles have been published elsewhere(Scholtens et al., 2015; Allen et al., 2014; Krokstad et al., 2013; Davey et al., 2003), herein a brief description of each cohort with respect to baseline recruitment and follow-up is summarised.

***EPIC-Oxford***: The EPIC-Oxford cohort is one of the components of the European Prospective Investigation into Cancer and Nutrition (EPIC) study(Riboli & Kaaks, 1997), which recruited over half a million people across 10 European countries. During 1993-1999, 57,446 participants aged>=20 years living throughout the United Kingdom were recruited into the study and completed baseline assessments. The study population of the EPIC-Oxford cohort consist of two groups based on their source of recruitment. 7,421 participants were successfully recruited from general populations (*"general population" group*). Men and women aged ≥ 35 years on the list of collaborating General Practices (GPs) in Oxfordshire, Buckinghamshire and Greater Manchester were invited to participate in the study. Consenting participants were examined and interviewed by trained nurses in the GP surgeries. The nurses also conducted anthropometric measurements, took a 30ml blood sample and

checked the completed questionnaire. The remaining 50,025 participants of EPIC-Oxford cohort were recruited by post, in a campaign that aimed to recruit as many vegetarians as possible from across the country (*"health-conscious" group*). Men and women aged 20 years and above were eligible to participate. The main questionnaire was sent to all members of the Vegetarian Society of the UK. Relatives and friends of these members were also encouraged to participate through a "snowballing" method. Members of The Vegan Society who expressed an interest in the study were also mailed the main questionnaire to complete. Participants recruited by post who were willing to provide a blood sample were contacted by their GPs and samples were collected in the GP surgeries. About one third of this cohort were vegetarians and vegans. A subset of 19,500 participants also provided blood samples for analysis.

***HUNT***: The HUNT (Helseundersøkelsen i Nord-Trøndelag) study is mainly based in the Nord-Trøndelag County in the central part of Norway. Nord-Trøndelag County is one of the 19 counties in Norway, consisting of 24 municipalities located in both inland and coastal areas. To date, three surveys have been conducted in the HUNT study, HUNT1 (1984-1986), HUNT2 (1995-1997) and HUNT3 (2006-2008). In each survey, data were collected via questionnaires, interviews and measurements by teams of trained staff located in health examination sites within each of the 24 municipalities. Every citizen living in the county aged 20 years or older at the time of each survey was invited to participate. The original aims of the HUNT study, as designed in HUNT1, were to investigate hypertension, diabetes and quality of life. Over time, the aims of the HUNT study have expanded to include more health-related lifestyles and outcomes in accordance with national health priorities. For example, CVD-related items were extensively examined in HUNT2 and HUNT3, hence data from both surveys were analysed with regards to specific research aims in this PhD project. 93,898 citizens were invited to HUNT2 survey, 65,232 of them (70%) participated (as defined by completing the main Questionnaire 1) and 65,007 of these had physical measurements and blood/serum samples collected. In HUNT3, 93,860 citizens were invited, 50,805 of them (54%) participated and 50,666 of these had physical

measurements and blood/serum samples collected. A total of 37,070 participants took part in both HUNT2 and HUNT3 surveys.

***LifeLines:*** The LifeLines cohort study was piloted in 2006 and started in 2007, and is now the largest population-based study in the Netherlands. The main focus of this study is to investigate universal risk factors (genetic, environmental, biomedical and psychosocial factors) and their modifiers in relation to a range of multifactorial diseases and healthy ageing, using its three-generation design. The cohort recruited from the three provinces (Groningen, Friesland, Drenthe) in the north of the Netherlands. This part of the country has a highly homogeneous population and low migration rates. Participants (index persons, or probands) aged 25 to 50 years who were registered in general practitioners' (GP) practices in one of the three provinces were randomly invited to participate. After expressing interest in participating and signing a consent form, participants were sent a baseline questionnaire (part one) and an invitation to attend one of the 12 LifeLines research sites where they underwent for health assessment, and completed the second part of questionnaire. In addition, participants were asked if their family members including partners, parents, parents-in-law and children were willing to participate. The same recruitment procedures were followed for these family members, thus the LifeLines cohort consists of a representative population-based cohort across three generations. The recruitment was completed in December 2013, by which time a total of 167,729 participants had been recruited.

Information on demography, socioeconomic position, lifestyle, general health, depression and medication use were collected via questionnaires at baseline. Completeness and correctness of data were checked by staff at the research sites. In addition, each participant underwent a physical examination, and had blood and urine samples collected.

Given the timeframe of this PhD project, only 93,277 out of the 167,729 LifeLines participants were able to be included, as data quality for these 93,277 participants had been checked. Record linkages

were still being processed at the time this thesis was written and therefore hospitalisation for incident CVD data were not available for this cohort.

Due to its three-generation study design, it is possible that some of the 93,277 LifeLines participants included in this project are indeed from the same family/household, which implies some clustering effects (e.g. shared genetic and environmental factors within the same family/household) may exist and therefore potentially influence the statistical estimates. However, I was not able to investigate specifically this clustering effect in the LifeLines cohort in the proposed pooled analysis in this PhD study. Hence, participants in the LifeLines cohort was treated as independent individuals in statistical analyses (Chapter 4) and limitations of doing so are acknowledged in section 6.2.4 of Chapter 6.

***UK Biobank***: UK Biobank is one of the largest population-based prospective cohorts in the world, with over 500,000 participants recruited from all over the UK. It was set up to study the genetic, environmental and lifestyle determinants of common diseases in middle-aged and older populations. Baseline assessment was conducted in 2006-2010. During this period, targeted participants aged 40-69 years were invited to visit one of 22 assessment centres throughout the country. At the assessment centre, participants signed a consent form and completed a touch-screen questionnaire. This computerised questionnaire allowed direct data entry to facilitate checks for completeness and consistency. A rich array of data of public health and research importance were collected. For data which were inconvenient to collect via touch-screen questionnaires (e.g. names of any specific medications taken), participants were interviewed briefly by trained staff at the centre to record these information. Participants also underwent physical and functional measurements, and collections of blood, urine and saliva samples for biobanking. By the end of the recruitment period, a total of 502,656 participants had provided data for this project initially, after seven participants withdrew in 2014, a sample size of 502,649 was available for analysis.

An overview of the participating cohorts is provided in Table 2.1. As stated in the “research objective” section in Introduction, this PhD study aims to investigate two specific outcomes - CVD and blood

biochemistry markers for CVD - in association with exposure to both ambient air pollution and road traffic noise. Blood biochemistry data were only available for LifeLines and HUNT3 within the timeframe of this project. EPIC-Oxford, HUNT2 and UK Biobank had data available to contribute to the incident CVD analyses.

Table 2.1 Overview of participating cohorts in this project

| | EPIC-Oxford | LifeLines | HUNT | UK Biobank |
|---|---|---|---|---|
| **Recruitment regions** | Throughout the UK | North of the Netherlands | Nord-Trøndelag, Norway | Throughout the UK |
| **Baseline periods** | 1993-1999 | 2006-2013 | 1995-1997 (HUNT2) 2006-2008 (HUNT3) | 2006-2010 |
| **Targeted populations** | General and vegetarian populations | General populations | General populations | General populations |
| **Targeted age groups** | ≥20 years | 25-50 years; children and older family members of the participant were also invited | ≥20 years | 40-69 years |
| **Sample size available for this project** | 57,446 | 93,277 | 65,232 (HUNT2) 50,805 (HUNT3) | 502,649 |
| **Contributed to analyses in this PhD** | Incident CVD | CVD blood biochemistry | CVD blood biochemistry (HUNT3) Incident CVD (HUNT2) | Incident CVD |

## 2.2. Cohort data harmonisation

There have been debates in recent years over the issue of whether to merge the existing established biobanks/cohort studies into a large consortium (e.g. the ESCAPE project) or to establish a new large-scale purpose-designed cohort to enable the study of the effects of genes and the environment on population health(Collins & Manolio, 2007; Willett et al., 2007). Indeed, both approaches have pros and cons. Clearly, combining data from several existing cohorts could quickly and efficiently produce the large sample sizes needed to study subtle but important environmental effects. This method requires only modest investments, making it particularly attractive in today's research funding climate. However, there are several challenges that need to be addressed when merging different studies as pointed out by Collins and Manolio(Collins & Manolio, 2007). The first concerns, the standardisation of collected data or measures, for instance, how valid is it to merge measures which were not uniformly collected? The second concern is that, existing studies may not represent the general populations from which they are drawn, and are more likely to underrepresent the younger age groups, as most cohorts are established to study disease-onset from mid-life. Third, ethical-legal issues (e.g. control of data, consent limitation) may block collaboration across various cohort studies. Collins and Manolio acknowledged that while huge resources may need to be invested in establishing a new large-scale cohort and that the first major results for certain diseases (e.g. cancers) will not emerge until years, if not decades, after recruitment, they propose that these two approaches should not necessarily be mutually exclusive. Merging data from various cohorts should be treated as an interim and effective way to study the gene-environment effects on some common diseases (e.g. CVD, diabetes), but in the long run, a rigorously-designed cohort equipped with new tools to capture a rich array of data is needed for future generations.

Combining data from various cohorts could be achieved in two ways: i) combining the results generated from various cohorts in which variables were standardised within the cohort using meta-analysis methods, also known as study-level meta-analysis; ii) combining the individual-level data from

each cohort to form a pooled individual-level database, providing that variables were already standardised across the cohorts. Examples are given below for these two approaches.

The Pooling Project of Prospective Studies of Diet and Cancer(Smith-Warner et al., 2006) in North America and the ESCAPE project in Europe are two good examples of study-level meta-analyses. Both projects involved many cohorts drawn from different regions and countries. Both exposures and confounding variables were standardised within each cohort, then uniform statistical models were applied to each cohort separately, and cohort-specific results were pooled using meta-analysis techniques to obtain the overall effect estimate. In both projects, findings from each cohort were largely consistent and heterogeneity across cohorts was not reported for most of the examined health outcomes.

Pooling individual-level data across studies has not been extensively attempted due to the foreseeable technical, ethical and legal issues that need to be considered thoroughly. The dedicated EPIC consortium is one of the few examples in pooling individual-level data across many EPIC component cohorts of which EPIC-Oxford is part. Data from EPIC cohorts were collected based on centre-specific questionnaires but all shared some major components. Standardised variables with universal definitions were derived from each local cohort based on a joint protocol, enabling pooled analyses (either individual-level pooled analysis or study-level pooled analysis) using these standardised variables. By doing this, the EPIC investigators hoped to resolve two major problems: statistical power and study size(Riboli & Kaaks, 1997).

Sample size and statistical power remain critical in epidemiological research, and are particularly important when studying effect modifications by several factors, or gene-environment interactions. Clearly, both methods of combining data across cohorts can achieve a large sample size and/or statistical power. Yet, most attempts to date are still targeted at the study-level meta-analysis, despite the fact that pooled individual-level analysis usually offers greater flexibility and statistical power(Roetzheim et al., 2012).

In response to the aforementioned debates, the EU-funded BioSHaRE consortium, in which this PhD project sits, was commissioned to give new insights into the effective pooling of individual-level data from various cohorts to study common complex diseases. Investigators from the BioSHaRE consortium aim to provide solutions to some of the concerns noted earlier by Collins and Manolio(Collins & Manolio, 2007). For example, new computational infrastructures have been developed by BioSHaRE for retrospectively harmonising key measures of lifestyle, social circumstance and environment across cohorts. In addition, novel statistical tools have been developed to enable virtual individual-level data pooling from various cohorts, allowing data custodians retain control over their data. This PhD project was designed to use and test these new tools and to demonstrate the value of pooled individual-level analysis.

### *Data Harmonisation Platforms in BioSHaRE*

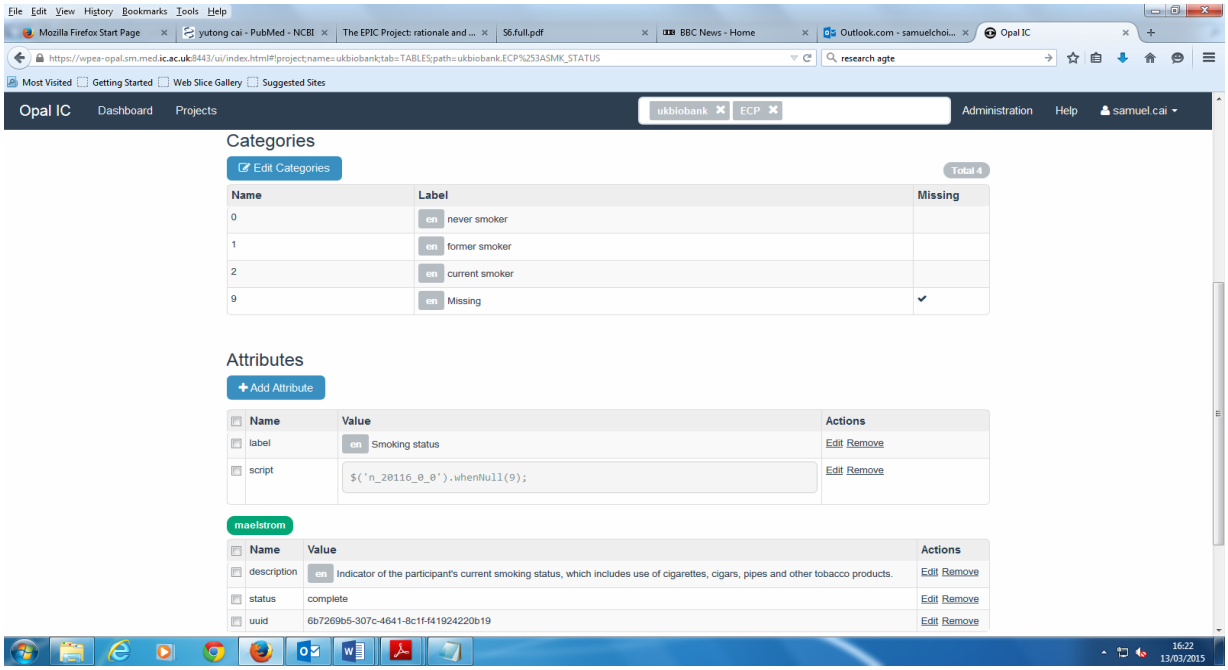
Fortier and colleagues from the BioSHaRE consortium have successfully streamlined and tested the data harmonisation methodology in a pilot study of 53 individual cohorts around the world, using their DataSHaPER (DataSchema and Harmonization Platform for Epidemiological Research) approaches(Fortier et al., 2011). This method, as its name suggests, is structured by two components, a dataschema platform and a harmonisation platform. First, a dataschema was developed with standard annotations attached to each variable requiring harmonisation. Second, based on the defined *harmonised* variable, data from each original cohort were examined and harmonised to achieve this defined harmonised variable, using computer-generated scripts. The DataSHaPER aims to provide a certain degree of flexibility with regards to harmonisation while also allowing meaningful research with a set of comprehensive data.

As a crucial first step, data compatibility across the participating cohorts in this PhD project were assessed, based on the DataSHaPER protocol, to inform whether these data could be retrospectively harmonised meaningfully such that they could be pooled validly for analyses to address my research objectives.

### *Dataschema Platform overview*

Using the DataSHaPER format, I developed an online dataschema for this PhD project. A core set of variables were identified for priority harmonisation across all cohorts (referred as "***Core harmonised variables***" hereafter). The strategy in choosing these core harmonised variables was based on my knowledge of the required analyses derived from current scientific literature, advice from senior colleagues, and general criteria as described in the DataSHaPER conceptual paper(Fortier et al., 2010). Briefly, variables were selected to capture information about demography, social circumstance, lifestyle and general health. These variables are either key components in epidemiological research, straightforward and reliable to measure, or are potentially important effect modifiers. Some data variables were not able to be harmonised in each cohort, partly because the original cohort did not collect the relevant data, or only collected partial information. Nevertheless, these data variables were still added to the dataschema for future research (referred as "***Other harmonised variables***" hereafter). Other data, for example, on physical activity, diet habits and area-level contextual variables (e.g. area deprivation index, income by neighbourhood) may be important confounders and/or effect modifiers, but were too heterogeneous to enable a valid synthesis within this project. Therefore, these data were not added to the dataschema. Figure 2.1 displays a screenshot of the working version of the online dataschema, using smoking status ("SMK_STATUS") as an example. Basically, for each harmonised variable, in this case "SMK_STATUS", the harmonised name, standard definition, scripts used to derive the variable from each cohort, label of each category (1=never-smoker; 2=ex-smoker;3=current-smoker), and harmonisation working progress (completed/pending/impossible) are shown on the screen, which allows researchers to follow the working status closely.

Figure 2.1 A screenshot of online working version of dataschema-"smoking status" as an example




In the following sections, procedures to derive harmonised variables (both core and other) from each of the four cohorts are described in detail.

### *Harmonisation Platform for core variables*

These core variables were harmonised across all cohorts and were used in the main analyses as important confounders and/or effect modifiers. Most variables were straightforward to define with some following a universal definition (e.g. smoking pack years) or international standard classification (e.g. highest education level). Some degree of flexibility was allowed in the harmonisation process. Harmonisation of core variables from each cohort are described as below.

- **Participant's characteristics variables**

All cohorts had information about age at recruitment ("**AGE_YRS**"), sex ("**GENDER**") and year of interview ("**ADM_YRINT**"). For age at recruitment, data were directly recorded for HUNT and EPIC-Oxford, while for UK Biobank and LifeLines this variable needs to be derived using date of birth and date of completion of the questionnaire form or date of attendance at the assessment centre.

Height ("**PM_HEIGHT**"), weight ("**PM_WEIGHT**"), waist ("**PM_WAIST_SIZE**") and hip circumference ("**PM_HIP_SIZE**") were all objectively measured in HUNT, LifeLines and UK Biobank, although measurement protocols differed. These measurement differences were permitted in the harmonisation process and as such data were harmonised in line with definitions set out in the Table 2.2.

Only a small proportion (7,388 out of 57,446, 13%) of EPIC-Oxford participants had height, weight, waist and hip circumference measured through attendance at a General Practice. Most other EPIC-Oxford participants self-reported their measures in the questionnaire by post (N=49,065). Approximately 7,000 participants had both anthropometry measured at a GP visit and also self-reported these data in the questionnaire. In these participants, there was a high correlation between self-reported and measured height(r=0.97) and weight(r=0.99)(Spencer et al., 2002). For self-reported and measured waist or hip circumference, the correlation was 0.91(Spencer, Roddam & Key, 2004). Based on both measured and self-reported data, colleagues in EPIC-Oxford had previously developed sex-specific prediction equations(Spencer, Roddam & Key, 2004; Spencer et al., 2002). It was found that using these predictive equivalents for height and weight, misclassification of BMI reduced from 22% to 15% in men and 18% to 14% in women, compared with using self-reported height and weight(Spencer et al., 2002). Therefore, the EPIC consortium recommended that these predicted values derived from the sex-specific equations should be used in my analyses where anthropometry measures are included as confounding variables, instead of the self-reported values derived from the questionnaire. In this PhD project, as anthropometry data were mainly used to adjust for confounding effects, predicted anthropometry values based on the sex-specific prediction equations were therefore used for all participants in EPIC-Oxford and were harmonised as if they were objectively measured values.

Waist-Hip ratio (WHR, "**PM_WHR**"), Body Mass Index (BMI, "**PM_BMI_CONTINUOUS**") were subsequently computed for all cohorts based on the relevant measures, as defined in Table 2.2.

Table 2.2 Harmonisation of core variables: participants' basic characteristics

| Harmonised name | Harmonised definition | Unit (continuous variables)<br>Categories (categorical variables) |
|---|---|---|
| **AGE_YRS** | Age of the participant in years (continuous) at recruitment | Years |
| **AGE_YRS_CATEGORICAL** | Age of the participant in years (categorical) at recruitment | 1: 18-30 years;<br>2: 30-40 years,<br>3: 40-50 years;<br>4: 50-60 years;<br>5: 60-70 years;<br>6:70 years and over |
| **GENDER** | Sex of the participant | 0: Male; 1:Female |
| **ADM_YRINT** | Calendar Year of the interview | calendar year |
| **PM_HEIGHT** | measured or self-reported height | cm |
| **PM_WEIGHT** | measured or self-reported weight | kg |
| **PM_HIP_SIZE** | measured or self-reported distance around the hips | cm |
| **PM_WAIST_SIZE** | measured or self-reported distance around the waist | cm |
| **PM_WHR** | Waist-Hip ratio (WHR): calculated from waist and hip size in each cohort (waist/hip) | % |
| **PM_BMI_CONTINUOUS** | Body Mass Index: calculated using measured or self-reported weight and height ($kg/m^2$) | $kg/m^2$ |
| **PM_BMI_CATEGORIAL** | Body Mass Index calculated using measured or self-reported weight | 1: less 25 $kg/m^2$;<br>2: 25 to 30 $kg/m^2$;<br>3: over 30 $kg/m^2$ |

| | | |
|---|---|---|
| | and height (Mass in Kg / (Height in metre)$^2$). | |

- **Participant's lifestyle variables**

**Smoking**

Smoking data including an indicator of current smoking status ("**SMK_STATUS**") and a quantitative marker of cigarette consumption measured across life, smoking pack-years ("**SMK_PACKYRS**"), were defined as with most epidemiological studies (see Table 2.3).

Smoking status was directly recorded for EPIC-Oxford, HUNT and UK Biobank in three categories (never, ex, and current smoker) and were subsequently harmonised for these three cohorts. For LifeLines, additional work was needed to secure this crucial variable. Three original variables in LifeLines were used to derive "SMK_STATUS", namely "SMK11" (*Have you ever smoked for a full year? Yes/no*), "SMK31" (*Do you smoke now, or have you smoked in the past month? Yes/No*) and "SMK51" (*Have you stopped smoking? Yes/No*). Never-smokers were defined as those who answered "No" to both SMK11 and SMK31 while current-smokers were defined as those who answered "Yes" to SMK31 and "No" to SMK51. Ex-smokers were defined as those who answered "No" to SMK31, "Yes" to SMK11 and "Yes" to SMK51.

The smoking pack-years ("**SMK_PACKYRS**") is a unit for measuring the amount a person has smoked over the life course. It is calculated by multiplying the numbers of packs of cigarettes smoked per day by the number of years the person has smoked. For example, 1 pack-year is equal to smoking 20 cigarettes (1 pack) per day for 1 year.

Equation 1: Calculation of smoking pack-years

**Number of pack-years= (number of cigarettes smoked per day x number of years smoked)/20**

As shown in the above Equation 1, information about daily number of cigarettes smoked and number of years of smoking are needed to derive the variable smoking pack-years for each cohort. Smoking

pack-years was readily available for EPIC-Oxford; the harmonisation was conducted in other three cohorts as follows.

Regarding the daily number of cigarettes smoked, UK Biobank asked separately in two questions for different time periods, now (if they still smoke) and before they quit. Answers to these questions were then applied to the calculations of smoking pack-years for current-smokers and ex-smokers respectively.

In HUNT, participants were asked-"*How many cigarettes do/did you usually smoke daily? (if now or earlier daily smoking)*"-but exactly the same information was obtained as those in UK Biobank.

For LifeLines, daily number of cigarettes smoked was recorded in current-smokers only. No specific data were recorded for ex-smokers regarding numbers of cigarettes they previously smoked daily. Instead, total number of cigarettes, roll-up cigarettes, cigarillos, cigars and pipe tobacco per day were recorded. It was assumed that most participants in LifeLines with a habit of smoking mainly smoked cigarettes rather than other types of tobacco. Therefore, this information was used as a proxy of number of daily cigarettes smoked previously for ex-smokers in LifeLines.

All three cohorts (HUNT, LifeLines, UK Biobank) had recorded "age started smoking" and "age quitted smoking" (applicable to smoking quitters). Number of years of smoking was therefore calculated separately for current-smokers and ex-smokers. For the current smokers, it was calculated as age at recruitment minus age started smoking while for ex-smokers, it was calculated as age quit smoking minus age started smoking.

For all known never-smokers, pack-years was set to zero.

**Alcohol consumption**

Current alcohol consumption at recruitment ("**ALC_CURRENT_QTY_TOTAL**") was quantified as average grams of total alcohol (beers, wines, spirits) consumed per week (see Table 2.3). Serving size

(grams of alcohol per drink) of alcoholic drinks usually differs from country to country and different serving size standards were reported in previous publications for each cohort. The harmonisation for these alcohol consumption variables is in line with these reported serving size standards for each cohort.

In EPIC-Oxford, beer, wines, spirits and total alcohol consumption (grams) at recruitment *per day* was respectively recorded for each participant. To derive the defined harmonised variable, the 'total alcohol consumption' in grams was multiplied by seven to represent averaged consumption *per week*.

In HUNT, average alcohol consumption was collected via three questions ("*How many glasses of beer (wine/spirits) do you usually drink in the course of two weeks*?"). Information about the serving size (e.g. how many grams of beer in a serving glass) was not directly mentioned in the questionnaire. However, in line with a previous HUNT publication(Rasouli et al., 2013), the average serving of beer is equivalent to *16* grams of alcohol, the average serving of wine is equivalent to *12* grams of alcohol and the average serving of spirits is equivalent to *12* grams of alcohol. Thus, reported glasses of each beverage consumed were multiplied by the above alcohol contents respectively in grams. These numbers were then summed up to give a total average consumption of alcohol over a two-week period. To derive the defined harmonised variable, this total average consumption of alcohol was therefore divided by two to represent average consumption of total alcohol per week.

In LifeLines, general alcohol consumption questions (regardless of alcoholic beverage types) were asked in the questionnaires. Participants were asked "*How often did you drink alcoholic beverages in the past month, also think of non-alcoholic beer*?" There were seven categories to choose from for this question: *not this month (non-drinker); 1 day per month (equivalent to 1/4 day per week); 2-3 days per month (equivalent to 2.5/4 per week); 1 day per week; 2-3 days per week; 4-5 days per week; 6-7 days per week*. For the latter three categories, drinking was averaged to 2.5days per week, 4.5 days per week and 6.5 days per week respectively. Participants were then asked "*On days that you drank alcoholic beverages, how many glasses did you drink on average*?". In the Netherlands, one standard

serving size of an alcoholic drink is equivalent to 9.9 grams of alcohol, regardless of beverage type(Slagter et al., 2014). As a result, total average consumption of alcohol per week was derived by multiplying the number of drinking days per week by number of glasses of alcoholic beverages drank on these drinking days by a standard serving unit of 9.9 grams of alcohol.

In UK Biobank, participants were asked about the number of drinks of each alcoholic beverage consumed weekly. For beer consumption, participants were asked "*how many pints of beer or cider would you drink in an average week*?"; for both red wine and white wine (including champagne), participants were asked "*how many glasses would you drink in an average week (typically there are six glasses per bottle)?*"; for spirits, participants were asked "*how many measures of spirits or liqueurs would you drink in an average week (there are 25 standard measures in a normal sized bottle)?*". In the UK, a unit of alcohol is used as a measure to quantify alcohol consumption and one unit is equal to eight grams of alcohol, as defined in a guideline published by the House of Commons Science and Technology committee in 2012 (http://www.publications.parliament.uk/pa/cm201012/cmselect/cmsctech/1536/1536.pdf, accessed in March 2015). As with a previous UK Biobank publication(Dawes et al., 2014), one pint of beer or cider is equal to 2.5 units of alcohol, one medium-sized glass of wine or champagne is equal to 2.3 units of alcohol and 1 measure of spirits or liqueurs is equal to 1 unit alcohol. As such, actual alcohol content of one serving size for each beverage was calculated by multiplying the units of alcohol by eight, which is *20* grams of alcohol per pint of beer or cider, *18.4* grams of alcohol per one medium-sized glass of wine or champagne, and *8* grams of alcohol per one measure of spirits or liqueurs. To derive the harmonised variable, the number of drinks per week for each beverage was multiplied by the above mentioned grams of alcohol content per drink for each beverage. Then these numbers were summed up to give an average consumption of total alcohol per week for each participant.

It should be noted that for all non-alcohol drinkers in each cohort, a value of zero was assigned to this harmonised variable.

Table 2.3 Harmonisation of core variables: participants' lifestyle variables

| Harmonised name | Harmonised definition | Unit (continuous variables)<br>Categories (categorical variables) |
|---|---|---|
| **SMK_STATUS** | Indicator of the participant's current and past smoking status, which includes use of cigarettes, cigars, pipes and other tobacco products. | 0: Never-smoker;<br>1: Ex-smoker;<br>2: Current-smoker |
| **SMK_PACKYRS** | The pack-years is a unit for measuring the amount a person has smoked over a long period of time. It is calculated by multiplying the number of packs of cigarettes smoked per day by the number of years the person has smoked. | Pack-years |
| **ALC_CURRENT_QTY_TOTAL** | Current quantity (grams) of total alcohol consumed on average in beer, wine and spirits per week. Current grams of total alcohol taken = (current quantity of beer taken*Average grams of alcohol in a bottle/glass/pint of beer) + (current grams of wine taken* Average grams of alcohol in a serving of wine) + (current grams of spirits taken *Average grams of alcohol in a serving of spirits). | grams/week |

- **Participants' socioeconomic position variables**

Two common indicators of individual socioeconomic position, current work status ("**WORK_STATUS_CURRENT**") and highest education attainment ("**EDU_HIGHEST_1**"), were identified and harmonised (see Table 2.4).

For "**WORK_STATUS_CURRENT**", both employees (either full-time or part-time) and self-employed persons were classified as "currently at work". An indicator of whether the participant was retired ("**WORK_RETIRED**") at the time of recruitment was harmonised in all cohorts except for HUNT.

Education systems are not directly comparable in the three countries (UK, the Netherlands, and Norway) from where the four cohorts originate. As expected, answers to the question about education attainment varied across cohorts. To simplify the harmonisation work, a harmonised variable ("**EDU_HIGHEST_1**") with three broad categories (no education or primary education; secondary education; post-secondary/vocational/college/university education) was created, adapting from the UNESCO Revision of the International Standard Classification of Education, 2011 (http://www.uis.unesco.org/Education/Documents/isced-2011-en.pdf, accessed in Dec 2015).

In EPIC-Oxford, there is a standardised EPIC variable "l_school" which recorded the highest school level the participant attained. This variable has six categories: *none; Primary school completed; Technical/professional school; secondary school; Longer education (incl. University degree); Not specified*. In EPIC-Oxford, no participants identified themselves in the groups of 'none' or 'primary school completed'. 19% of participants (N=10,791) did not specify their highest education level and therefore these participants were treated as having missing information. After liaising with the EPIC-Oxford data team, those who finished technical or professional school were grouped as having secondary education for this harmonisation. It should be noted that the then technical school or professional school in UK that EPIC-Oxford participants (mostly born in 1920s-50s) attended may not be the same level as those that later generations attended. Indeed, the mean age of leaving school for those who reported attaining technical or professional education in EPIC-Oxford was 16.8 years, which

justified grouping these participants in the secondary education group. Obviously, those with longer education were grouped as having post-secondary/vocational/college/university education.

In HUNT2, participants were asked about their highest level of education. There were 5 options (translated from Norwegian to English) and participants were asked to select the most suitable one: *1) primary school 7-10 years, continuation school, folk high school; 2) high school, intermediate school, vocational school, 1-2 years high school; 3) university qualifying examination, junior college, A levels; 4) University or other post-secondary education, less than four years; 5) university/college, 4 years or more*. To derive the harmonised variable, participants who selected option 1 were classified as having primary school education only; participants who selected options 2 and 3 were all considered as having secondary education; participants who selected options 4 and 5 were those who had post-secondary/vocational/college/university education.

In LifeLines, participants were asked about their highest education level. Participants were asked to select one of the eight options (translated from Dutch to English): *1) No education (primary education not completed); 2) Primary education; 3) Lower or preparatory vocational education; 4) Lower general secondary education; 5) Intermediate vocational education or apprenticeship; 6) Higher general secondary education or pre-university secondary education; 7) Higher vocational education; 8) University*. To derive the harmonised variable, participants who selected options 1 and 2 were classified as having no education or primary school only; those selected options 3 to 6 were classified as having secondary education; participants who selected options 7 and 8 were classified as having post-secondary/vocational/college/university education.

In UK Biobank, participants were asked about all educational qualifications obtained, by selecting more than one of the following eight options: *college or university degree; A-level/AS levels or equivalent; O-level/GCSEs or equivalent; CSEs or equivalent; NVQ or HND or HNC or equivalent; other professional qualifications (e.g. nursing); none of the above; prefer not to answer*. If the participant had selected more than one option, the highest level of qualification was then chosen for that

participant. To derive the harmonised variable, procedures were adopted as follow: i) if participants selected "*none of above*" only then they were classified as having no education or had only finished primary school; ii)*O-level, GCSEs and CSEs* are all qualifications of secondary education (usually obtained at age 16 years) in different generations in the UK whilst *A-level/AS levels* is a pre-university secondary education qualification (usually obtained at age 18 years), if participants had selected one of these as the highest qualification attained, then they were classified as having secondary education; iii) *NVQ/HND/HNC/Professional* are all vocational training qualifications. Participants who selected one of these as the highest qualification as well as participants who had a university degree were classified as having post-secondary/vocational/college/university education. Those who preferred not to answer this question were classified as having missing information.

Table 2.4 Harmonisation of core variables: participants' socioeconomic position

| **Harmonised name** | **Harmonised definition** | **Unit (continuous variables)**<br>**Categories (categorical variables)** |
|---|---|---|
| **WORK_STATUS_CURRENT** | Indicator of whether the participant is currently in paid employment or is self-employed. | 0: No paid employment or not self-employed;<br>1: Paid employment or self-employed |
| **EDU_HIGHEST_1** | Highest level of education completed by the participant. Categories are adapted from the UNESCO Revision of the International Standard Classification of Education, 2011 | 0: No education or primary education<br>1: Secondary education;<br>2:Vocational/college/university (post-secondary education) |

### *Harmonisation Platform for other variables*

Other variables were also identified for harmonisation across cohorts to enable additional sensitivity or stratified analyses. However, not all variables were collected by each cohort or were collected only

in a subgroup of each study population. A summary of these harmonised variables is listed in Table 2.5.

Length of current residency ("**RES_LENGTH**") at the time of recruitment was obtained from the questionnaire in UK Biobank by asking participants how many years they had lived at their current (at baseline recruitment) address. For HUNT and LifeLines participants, multiple residential addresses were obtained, and the date of moving to last address was confirmed. Date at recruitment minus date of moving to last address was then used to calculate how many years the participant had lived in their current address. History of residential addresses or length at current residence was not obtained for EPIC-Oxford.

Information about environmental exposure to tobacco smoking at home ("**ETS_HOME**") was only available in LifeLines and UK Biobank, but only for non-smokers. In LifeLines, an open question was asked in the questionnaire: "*How many people smoke regularly in your household",* if participants indicated no such persons, then they were classified as not having exposure to ETS; if participants indicated one or more than one persons smoke, then they were classified as having exposure to ETS. A similar question was asked in UK Biobank, "*Does anyone in your household smoke",* participants indicated yes or no.

Information about the number of people in the household ("**HOUSEHOLD_NUM_PPL**") was only recorded for LifeLines and UK Biobank. In UK Biobank, participants were asked "*how many people were living together in your household (including yourself)*". If participants only answered one, then it was assumed that they were living alone ("**LIVING_ALONE**"). Exactly the same question regarding number of people in the household was asked in LifeLines, and the same procedure was followed to derive the "**LIVING_ALONE**" variable in LifeLines. The HUNT cohort had information on living alone, but not on the number of people in the household. EPIC-Oxford asked no questions that would enable either harmonised variable to be derived.

Data on measured systolic blood pressure (SBP, "**PM_SYSTOLIC_MEASURE**") and diastolic blood pressure (DBP, "**PM_DIASTOLIC_MEASURE**") were available in all cohorts, although in EPIC-Oxford these data were available for only 19,500 participants (33% of the total 57,446). Measurement protocols varied by cohort. In EPIC-Oxford, measurement of blood pressure was undertaken by trained health professionals at general practice. Two measurements of SBP and DBP were planned for each participant, although some only had one measurement recorded. If two measurements of SBP and DBP were recorded, then an averaged value of these two measurements was adopted. For HUNT, three measurements of SBP and DBP were undertaken for each participant. In line with a previous publication(Fagernaes et al., 2015), the mean of the second and third measurement of SBP and DBP was adopted as the final value. In LifeLines, 10 measurements of SBP and DBP were conducted, but only the final two measurements were averaged as the final values according to the LifeLines protocol. In UK Biobank, two measurements of SBP and DBP were undertaken and an averaged value was used for the harmonised variable.

History of ever-had specific diseases was reported in questionnaires by participants from all cohorts. Four main diseases were identified for harmonisation: diabetes (type 1 or type 2, "**DIS_DIAB**"), stroke ("**DIS_CVA**"), myocardial infarction ("**DIS_AMI**") and hypertension ("**DIS_HBP**"). For both EPIC-Oxford and UK Biobank, similar questions were asked: "*have you been told by a doctor/Has your doctor told you that you have the [disease]? Yes/no*." But for LifeLines and HUNT, the question was worded as "*have you had or do you have the following [disease]? Yes/no.*" Data were harmonised assuming these are equivalent questions. Information about history of the hypertensive conditions was not available in HUNT. Here data on current medication use for hypertension was used as an indicator of whether the participant had hypertension. Those who were currently using anti-hypertensive medication were grouped as "has had hypertension" of for the "DIS_HBP" variable. There have been some debates about the validity of these self-reported data to identify prevalent cases(Woodfield et al., 2015; Huerta et al., 2009). In general, the accuracy of self-reported diabetes is usually higher, compared with that of, for instance, self-reported hypertension or stroke, because participants may misinterpret the

diagnosis with other similar cardiovascular diseases. Nevertheless, in combining with other data sources such as medical records, these data are useful for case ascertainment.

Additional variables describing current medication use against hypertension ("**MEDI_HBP**") or diabetes ("**MEDI_DIAB**") were available for LifeLines and UK Biobank, collected via questionnaire or interviews. ATC codes (Anatomical Therapeutic Chemical (ATC) Classification System) were used in LifeLines to identify the exact medication that the participant has taken. Medication data was not available in EPIC-Oxford.

Table 2.5 Harmonisation of other variables

| Harmonised name | Harmonised definition | Unit(continuous variables); Categories (categorical variables) | Not available in |
|---|---|---|---|
| **RES_LENGTH** | Length of current residency | years | EPIC-Oxford |
| **LIVING_ALONE** | Indicator of whether the participant was living alone | 0: Not live alone<br>1: live alone | EPIC-Oxford |
| **HOUSEHOLD_NUM_PPL** | Number of people who live with the participant in the same household (including the participant) | - | EPIC-Oxford, HUNT |
| **ETS_HOME** | Indicator of whether the participant (not current-smokers) was regularly exposed to someone's | 0: No<br>1: Yes | EPIC-Oxford, HUNT |

| | | | |
|---|---|---|---|
| | smoking at home or was living with someone who were a regular smoker. | | |
| **PM_SYSTOLIC_MEASURE** | Measured systolic blood pressure. | mmHg | - |
| **PM_DIASTOLIC_MEASURE** | Measured diastolic blood pressure. | mmHg | - |
| **DIS_DIAB** | Occurrence of diabetes at any point during the life of the participant (not including gestational). Can be collected using an assessment item asking about the history of diabetes specifically, or from a list of disease history using ICD-10 Codes E10-E16 | 0: Never had diabetes<br><br>1: Has had diabetes | - |
| **DIS_HBP** | Occurrence of high blood pressure at any point during the life of the participant. Can be collected using an assessment item asking | 0: Never had high blood pressure<br><br>1: Has had high blood pressure | Not directly available in HUNT, but use the current medication of hypertensive (MEDI_HBP) as a proxy |

| | | | |
|---|---|---|---|
| | about the history of hypertension specifically, or from a list of disease history using ICD-10 Codes I10-I15 | | |
| **DIS_AMI** | Occurrence of myocardial infarction at any point during the life of the participant. Can be collected using an assessment item asking about myocardial infarction or heart attack or from a list of disease history using ICD-10 Codes I20-I22 | 0: Never had myocardial infarction<br><br>1: Has had myocardial infarction | - |
| **DIS_CVA** | Occurrence of stroke at any point during the life of the participant. Can be collected using an assessment item asking about the history of stroke or from a list of disease history using ICD-10 Codes I60-I69 | 0: Never had stroke<br><br>1: Has had stroke | - |

| | | | |
|---|---|---|---|
| **MEDI_HBP** | Indicator of whether the participant currently uses antihypertensive medication. Self-reported by a question targeting the use of antihypertensive medication or extracted from a list of medication using the following ATC Codes or equivalent in other classifications: C02, C03, C04, C07, C08, C09 | 0: Not currently using antihypertensive medication<br><br>1: Currently using antihypertensive medication | EPIC-Oxford |
| **MEDI_DIAB** | Indicator of whether the participant currently uses blood glucose lowering medication. Self-reported by a question targeting the use of glucose lowering medication or medication to treat diabetes or extracted from a list of medication using the following ATC Codes or equivalent in | 0: Not currently using glucose lowering medication<br><br>1: Currently using glucose lowering medication | EPIC-Oxford<br><br>HUNT |

| | | | |
|---|---|---|---|
| | other classifications: A10. | | |
| **WORK_RETIRED** | Indicator of whether the participant is currently retired. | 0: Not retired<br>1: Retired | HUNT |

### *Harmonisation platform for outcome variables*

Outcome variables in this project were identified and harmonised across the cohorts to enable me to meet the two research objectives.

### *Blood biochemistry data*

Very little blood biochemistry data was available in EPIC-Oxford and data from UK Biobank were not released in the timeframe of this PhD project (data due to be released mid-2016). Therefore, laboratory measurements for blood biochemistry markers were obtained from HUNT3 and LifeLines only (Table 2.6). Both cohorts have measured a range of different blood biochemistry markers. After searching through the measured biomarkers in each cohort, five blood biochemistry markers for cardiovascular diseases were found to have been measured in both cohorts for the whole population. These include: total serum cholesterol ("**LAB_TSC**"), triglycerides ("**LAB_TRIG**"), high-density lipoprotein (HDL) cholesterol ("**LAB_HDL**"), High-sensitivity C-reactive protein ("**LAB_HsCRP**") and blood glucose. The procedures for collection, storage, transport and analysis of blood samples in the two cohorts are described in full elsewhere(Scholtens et al., 2015; Hveem, 2011). Both cohorts followed the highest standards in processing these blood samples. However, regarding the blood sample itself, non-fasting blood samples were collected in HUNT3 whilst fasting blood samples were collected in LifeLines. In the process of harmonisation, such differences were allowed for all other four biochemistry makers, but not for blood glucose, as it is known that fasting versus non-fasting blood samples may differentially affect the levels of measured blood glucose. In this PhD project, to assess

the effects of air pollution or noise, I used data on fasting blood glucose from LifeLines cohort only (see details in Chapter 4).

### *Incident CVD outcomes*

Record linkages to incident cardiovascular outcomes from hospital admission records and/or mortality registers were made in EPIC-Oxford, HUNT2 and UK Biobank. Ascertainment of incident CVD outcomes in each cohort is described below, and a summary displayed in Table 2.6.

**EPIC-Oxford**: Each participant was followed for first CVD admission to hospital following recruitment using the unique National Health Service (NHS) number. In England, information on hospital admissions for each patient are available from 1 April 1997 from the Hospital Episode Statistics (HES) database. In Scotland, the equivalent database is called the Scottish Morbidity Records (SMR) which started in 1 January 1981. Hospital admission data for each participant from England (from 1 April 1997 to 31 December 2012) and Scotland (from 1 January 1981 to 31 December 2008) were obtained via linkages to the respective database and provided by EPIC-Oxford(Crowe et al., 2013). The diagnosis codes used WHO 9$^{th}$ version of International Classification of Diseases (ICD9) before 1 April 1996, and since then the 10$^{th}$ version (ICD10). The respective ICD codes used to identify incident cardiovascular diseases were ICD9 390-459 and ICD10 I00-I99. For participants dying during follow-up, cause of death up to 30 December 2009 was obtained from the NHS central register. Incident CVD death were ascertained when CVD (ICD9 390-459; ICD10 I00-I99) was either a primary or a secondary cause of death, providing that the participant had never had CVD diagnosed/reported prior to the death.

Following internal 'standard exclusions' in EPIC-Oxford, linkages were not possible for 12,342 of the recruited 57,446 participants (*Email correspondence from EPIC-Oxford team*). In addition, participants whose nations of residence was Wales or Northern Ireland as well as those participants who had pre-existing self-reported CVD-related history (heart attack, angina, stroke, hypertension and hyperlipidaemia) were excluded (N=4720), making linkages possible for a total of 40,384 participants.

**HUNT**: The HUNT cohort has been further enhanced by linkage to registers covering all participants in all three surveys. Information from national or local registers is linked to each HUNT participant using the unique Norwegian Personal Identification Number (PIN). In this PhD project, after recruitment to HUNT2, participants were followed for first CVD admission to hospital up to 31 March 2015. Incident CVD cases were identified by linkages with medical records at the only two local hospitals serving the population of the county of Nord-Trøndelag. Before 1 January 1999, ICD9 codes were used as diagnosis codes in the medical records, and since then ICD10 codes were used. The respective ICD codes used to identify incident cardiovascular diseases are ICD9 390-459 and ICD10 I00-I99. Mortality data until 31 December 2013 were obtained from the National Cause of Death Registry. Participants were ascertained as incident deaths if CVD (ICD9 390-459; ICD10 I00-I99) was one of the causes of death and participants had never had CVD diagnosed/reported before. Participants who reported at the time of HUNT2 recruitment that they had ever had heart attack, angina pectoris, stroke, previously taken medication for hypertension as well as those who were currently taking medication for hypertension were excluded (N=12,037), making linkages possible for a total of 66,844 participants.

**UK Biobank**: The follow-up for disease occurrence in UK Biobank was made possible through linkages to various national registers. As for EPIC-Oxford, data on any hospital admission for each participant are available from HES in England and SMR in Scotland. Hospital data were not yet available for Wales to be included in this PhD project. As for the EPIC-Oxford and HUNT studies, participants were followed for first CVD admission to hospital or CVD death following recruitment. ICD10-coded hospital data were available for UK Biobank from 1996-1997 to March 2010 in England and to December 2012 in Scotland. In this PhD project, the censored date for hospital CVD admission was defined as 31 March 2010 for UK Biobank. Cause of death up to 30 December 2013 was obtained from the Health & Social Care Information Centre (HSCIC) for each participant in England and Wales; while for Scotland, mortality data up to 30 November 2012 were obtained from Information Services Department (ISD). The respective ICD codes used to identify incident cardiovascular diseases from all registers are ICD10 I00-I99. After excluding those who had CVD diagnosed in their medical records before recruitment to

UK Biobank and those who reported in recruitment that they had ever had heart attack, stroke, hypertension and angina (N=165,426), record linkages were possible for a total of 337,223 participants.

Table 2.6 Descriptions of studied outcomes in each cohort

| | EPIC-Oxford | LifeLines | HUNT | UK Biobank |
|---|---|---|---|---|
| **Blood biochemistry data for this project** | n/a | High sensitivity C-reactive protein, cholesterol, triglycerides, High-density lipoprotein, fasting blood glucose and HbA1c Available for up to 93,277 participants | High sensitivity C-reactive protein, cholesterol, triglycerides, High-density lipoprotein, non- fasting blood glucose Available for up to 50,666 participants from HUNT3 | n/a |
| **Follow-up of incident CVD** | | | | |
| Data source | Hospital Episode Statistics (HES); Scottish Morbidity Records (SMR); NHS central register for death; | n/a | Two local hospitals in the Nord-Trøndelag county; National Cause of Death Registry; | Hospital Episode Statistics (HES); Scottish Morbidity Records (SMR); NHS central register for death; |
| Start date of medical records | 01-Apr-1997 (England HES) 01-Jan-1981 (Scotland SMR) | n/a | 01-Jan-1995 | 01-Apr-1997 (England HES) 01-Jan-1981 (Scotland SMR) |
| Censored date for first hospital CVD admission | 31-Dec-2012 (England HES) 31-Dec-2008 (Scotland SMR) | n/a | 31-Mar-2015 | 31-Mar-2010 (England and Scotland) |
| Censored date for death | 30-Dec-2009 | n/a | 30-Dec-2013 | 30-Dec-2013 |
| Pre-existing CVD cases before recruitment | 4,720 | n/a | 12,037 | 165,426 |
| Linkages possible to confirm incident CVD cases | 40,384 | n/a | 66,844 (includes participants from HUNT1, 2, and 3. | 337,223 |

Three harmonised outcomes including incident total cardiovascular events ("**FAILURE_CVD**"), incident ischaemic heart diseases ("**FAILURE_MI**") and incident cerebrovascular diseases ("**FAILURE_STROKE**") and one harmonised variable describing the follow-up period (i.e. the person-years of follow-up) ("**ENDTIME**") were then defined (Table 2.7). All three harmonised CVD-related outcomes are binary

variables (yes vs. no), indicating if the participant had developed CVD or not during the follow-up period. Calculations of follow-up period (person-years) varied across the three cohorts as each had different censor dates for the record linkages (Table 2.6). In general, in each cohort, person-years were calculated from the date of recruitment until the date of first incident CVD hospital admission, death from an incident CVD cause (either underlying or contributing), death from non-CVD causes, emigration abroad or the end of follow-up (censored date), whichever came first. Some EPIC-Oxford participants (n=5,220, 13% of the record linkages) in England were recruited before the beginning of HES medical records on 1-April-1997, therefore person-years for these participants were calculated from 1-April-1997.

Table 2.7 Harmonisation of outcome variables

| Harmonised name | Harmonised definition | Unit(continuous variables); Categories (categorical variables) |
|---|---|---|
| **LAB_TSC** | Laboratory measurement of total serum cholesterol | mmol/L |
| **LAB_HsCRP** | Laboratory measurement of High-sensitivity C-reactive protein (hsCRP) | mg/L |
| **LAB_TRIG** | Laboratory measurement of triglycerides | mmol/L |
| **LAB_HDL** | Laboratory measurement of high-density lipoprotein (HDL) cholesterol. | mmol/L |

| | | |
|---|---|---|
| **FAILURE_CVD** | indicator of incident total CVD event (ICD10:I00-I99) during the follow-up periods | 0: No<br>1: Yes |
| **FAILURE_MI** | indicator of incident ischaemic heart diseases event (ICD10:I20-I25) during the follow-up periods | 0: No<br>1: yes |
| **FAILURE_STROKE** | indicator of incident cerebrovascular diseases event (ICD10: I60-I69) during the follow-up periods | 0: No<br>1: yes |
| **ENDTIME** | persons-years of follow-up | Person-years |

# Chapter 3 Exposures modelling

Environmental exposure modelling is now increasingly used in health studies to estimate exposure to environmental hazards including ambient air pollution and traffic noise. Indeed, these techniques are practical in the case of large population-based studies, where personal measurement is almost impossible. An exposure model can be defined as "*a conceptual or mathematical representation of the exposure process*"(WHO, 2005). Generally there are two kinds of exposure assessment model(WHO, 2005). One is a mechanistic model which uses real physiochemical characteristics to simulate the exposure's behaviour in the environment and its key exposure pathways. This model is a mathematical construct and results can be calculated without any measurements. The other is an empirical model which is a mathematical representation of the relationship between input and output variables. This model is based on measurements and both input and output variables should be known before developing the model. Many models are currently being used in the rapidly evolving discipline of exposure science for applications in different fields. For example, various models have been described and used for ambient air pollution assessments as well as for traffic noise. Modelling techniques for ambient air pollution and road traffic noise were not reviewed in detail for each model type, as this was considered beyond the scope of this PhD project. Herein, I focused on the two specific models used in this project to assign individual ambient air pollution and traffic noise exposure to each cohort participants. Particular attention was paid to issues which may potentially affect the subsequent health impact assessments.

## 3.1. Land Use Regression model for ambient air pollution assessment

In air pollution epidemiological studies undertaken in the earlier 1990s(Pope et al., 1995; Dockery et al., 1993), exposure estimates for study participants were usually characterised by using average estimates over the study period from nearby fixed-site air pollution monitoring stations. While these exposure estimates were convenient to use in health studies, they did not adequately capture exposure variability at the individual level due to, for example, the outdoor-indoor air pollution level relationship, time activity patterns and in particular the small-scale spatial variations in intra-urban

areas(Ryan & LeMasters, 2007). As a result, this crude approach to air pollution assessment contributed to uncertainty in effect estimates in epidemiological analyses. As new exposure assessment methodology emerges, there are now new models available which address some of these limitations. These new models are able to generate more refined estimates for ambient air pollutants at the individual address level(Baxter et al., 2013). Intra-urban traffic-related air pollution has high spatial variability, particularly for gaseous pollutants (e.g. $NO_2$) at a small-scale(Kunzli, 2014). It was reported that concentrations of emitted pollutants from vehicles reduced by up to 90% within 50-500m of a busy road(Zhu et al., 2002). Even within 50m of a busy road, concentrations may substantially differ(Nicholas Hewitt, 1991). There is clearly a need to capture these spatial differences in air pollution levels using models capable of delivering output at finer geographic resolutions.

Briggs et al first reported that these variations over small distances can be predicted using regression-based methods, now referred to as Land Use Regression (LUR)(Briggs DJ, Collins S, Elliott P, Fischer P, Kingham S, Lebert E, et al, 1997).

LUR models are a form of empirical model and are being increasingly applied in air pollution epidemiological studies around the world. By definition, LUR predicts exposure to a specific pollutant at unmeasured locations (i.e. home addresses of cohort participants) using parameters derived from a multivariate regression model(Ryan & LeMasters, 2007). More specifically, in the model, measured level of the specific pollutant at certain locations forms the dependent variable, and a range of location-specific variables obtained from a geographic information system (GIS) are included as independent variables. These location-specific variables usually consist of variables such as road networks (e.g. traffic type, traffic load, road length), the surrounding environments of the measurement site (e.g. population density, area-level socioeconomic status), altitude and the land use (e.g. industrial land use, open space). The best performing model, defined as the one with the highest % variation ($R^2$) explained, is subsequently applied to predict levels of that specific air pollutant at the unmeasured sites.

This LUR method was first proposed as “regression mapping” in the SAVIAH (Small Area Variations in Air Quality and Health) project in the mid-1990s(Briggs DJ, Collins S, Elliott P, Fischer P, Kingham S, Lebert E, et al, 1997). In that study, Briggs et al conducted four 2-week measurement surveys of $NO_2$ across 80 sites in Amsterdam, Huddersfield and Prague. They then used these measured estimates, along with local variables derived from a GIS, to develop a separate regression model in each of the three cities. By validation against measured concentrations at eight to ten ‘reference’ sites, these models had a good performance in predicting annual mean concentration of $NO_2$ ($R^2$ ranged from 0.79 to 0.87). Two key findings were suggested for future studies to replicate: first, estimates based on a short measurement campaign in each of the study areas could be used for modelling long-term (e.g. annual) air pollution concentrations; second, spatial patterns of $NO_2$ likely remained broadly stable year to year in urban areas.

It is still debatable though whether LUR model developed in a specific location is transferable to other areas, as the relationships between model parameters may differ between different geographic and traffic settings. Briggs et al found that the model developed for $NO_2$ in Huddersfield in England could also be applied to other areas in the same country (Northampton, Sheffield, Hammersmith and Ealing of Greater London), with $R^2$ ranging from 0.60 to 0.76(Briggs et al., 2000). However, the authors did point out that without local calibrations, estimates from the model for these areas could be either underestimated or overestimated. Following the successful SAVIAH project, Brauer et al conducted the first LUR model for $PM_{2.5}$ and $PM_{2.5\ absorbance}$ in three European areas (the Netherlands, Munich and Stockholm County) in the TRAPCA project(Brauer et al., 2003). They measured both air pollutants across 40-42 sites in each area during four separate 2-week campaigns in one year. The $R^2$ for the models were 0.73, 0.56 and 0.50 for $PM_{2.5}$ in the Netherlands, Munich and Stockholm County respectively. For $PM_{2.5absorbance}$, the models performed slightly better with $R^2$ of 0.81, 0.67 and 0.66 respectively, likely due to the fact that $PM_{2.5\ absorbance}$ is more related to local traffic whilst $PM_{2.5}$ could also be generated by sources other than local traffic and is generally seen as an indicator of regional air pollution(Kelly & Fussell, 2012b).

After these two pioneering works, LUR methods have been increasingly adopted in other parts of the world, although mostly in North America and Europe. In the Montreal city areas of Canada, Gilbert et al developed a LUR model for $NO_2$ concentrations(Gilbert et al., 2005). They monitored $NO_2$ in a single consecutive two-week period in May 2003 and the model was developed based on these measured values and variables from GIS. The $R^2$ for the resultant model was 0.54. Another three Canadian studies in Hamilton(Sahsuvaroglu et al., 2006), Vancouver(Henderson et al., 2007) and Toronto(Jerrett et al., 2007) measured $NO_2$ at over 90 sites in also a one two-week period in 2002-2003, and their LUR model achieved $R^2$ of 0.76, 0.56 and 0.69 respectively. In the APMOSPHERE project in Europe, in which routine air pollution data were obtained from monitoring networks to develop the LUR model at a 1km resolution in 15 countries, the LUR yielded a $R^2$ of 0.61(Hoek et al., 2008). LUR techniques were applied in the pan-European ESCAPE project in 2008-2011, in which air pollution estimates from the models were used in epidemiological studies across cohorts in Europe. The LUR methodology for the ESCAPE project was also applied in this PhD project and was described later in Section 2 of this Chapter.

A review in 2008 listed some key issues with regards to measurement sites, measurement campaigns, applying LUR modelled estimates to earlier established cohort studies and model development of the LUR methodology(Hoek et al., 2008). I briefly summarised each of these below.

**1. Selection and distribution of measurement sites**

The first issue is the strategy of selecting measurement sites in purpose-designed measurement campaigns for LUR. One should question how many sites are required to develop a model and how these sites should be distributed across the study area. Current published LUR studies have most often included up to 100 sites, however more measurement sites do not necessarily guarantee better model performance(Clougherty et al., 2008; Henderson et al., 2007). Hoek et al suggested that 40-80 sites are ideal, but also noted that the number depends on the sizes of specific study areas(Hoek et al., 2008). Basagana et al tested the effect of the number of measurement sites on LUR model

performance in the urban area of Girona, Spain(Basagaña et al., 2012). Their analyses were based on data from 148 measurement sites across the city area, in which they held 28 sites as a validation dataset while the other 120 sites were randomly selected for model development. The authors concluded that models developed from a small number of measurement sites (e.g. 20 sites or so) resulted in an inflated $R^2$. In a compact urban area such as Girona, it was suggested that more measurement sites (80 sites in the Girona study) should be deployed to improve the model performance. Interestingly, the authors also suggested that this should be done in line with restricting the number of predictor variables in the regression model. The distribution of the measurement sites is also crucially important. It is desirable that the measurement sites reflect the variations of air pollution across the study areas(Hoek et al., 2008). In most previous LUR studies, the distribution of these sites was based on investigators' own judgements on the study areas (e.g. population size, traffic flows), with sites usually located at busy traffic locations, urban residential locations and regional background locations. It was reported that model performance could be enhanced if site type was included in the model when other data (e.g. traffic data) were not available(Gulliver et al., 2011).

**2. Temporal issues related to dedicated measurement campaigns**

A second key issue concerns the temporal dimension of measurement campaigns- how long should they last and/or how frequently should they be repeated? After first work on LUR methods from Briggs and co-investigators was published, many studies have adopted the similar strategies, usually including several rounds of 1-week or 2-weeks air pollution measurements during a specific study period (i.e. 1 calendar year). Preferably these measurement periods should fall into each season to account for seasonal variations of air pollution. Ideally, in each measurement period, air pollution data should be collected simultaneously at all measurement sites(Hoek et al., 2008). There have been concerns that, given the short time-frame of measurements, these measured values will not represent the annual mean values of the year of measurement. Results from a study in Canada by Henderson et al provided some insights into this(Henderson et al., 2007). They suggested that the average

concentration from two rounds of 14-day measurement campaigns were similar to the annual mean obtained from regulatory monitors. This however is rather study-specific, and Hoek et al recommended that up to four rounds (each of two weeks) of such campaigns would be preferable to generate reliable information about long-term averages(Hoek et al., 2008).

### 3. Applying LUR modelled estimates to earlier established cohort studies

In established cohort studies, it is common that baseline data collections and/or health outcome ascertainment occurred prior to the air pollution monitoring campaign and accompanying LUR modelling. This raises the issue of how to assign exposures modelled years after the baseline years. This is especially important for certain types of epidemiological study, for example, birth outcome studies. An example of "*temporal transferability*" was reported by Mölter et al(Mölter et al., 2010). LUR models for $NO_2$ and $PM_{10}$ for the year 2005 in the Greater Manchester area were developed. The authors then recalibrated these models using interpolated $NO_2$ and $PM_{10}$ as dependent variables for each year between 1996 and 2008. To validate results from these LUR models for each year, the authors compared them with the measured values from monitoring stations. Their results showed the mean prediction errors (MPE) were consistently low for both $NO_2$ and $PM_{10}$ models each year except for the year 1996 when data from monitoring stations were very limited (only one or two stations). This study demonstrated that temporal variation in air pollution can be modelled by LUR as long as there are some reliable historical air pollution data to allow recalibration. However, it is not always feasible to obtain such historical data and there are research questions about how far back a LUR model could be applied. Gulliver et al reported that LUR model for $NO_2$ developed in 2009 for Great Britain could be back-extrapolated to 1991, and reported mean-squared-error-based $R^2$ from hold-out validation of 0.52-0.55(Gulliver et al., 2013). In the Netherlands, Eeftens et al developed two LUR models, one in 1999-2000 and the other in 2007(Eeftens et al., 2011). $NO_2$ measurements were collected in 35 sites on both occasions. They found that the 2007 model predicted 77% of the spatial variability of the 1999-2000 $NO_2$ measurements while the 1999-2000 model predicted 81% of the

spatial variability for the 2007 measurements. LUR, in this particular setting, and over this 8-year gap, predicted $NO_2$ variations well both backward and forward in time. A similar outcome was reported in Oslo, Norway where two LUR NOx models developed for the year 2005 and 2008 predicted well each other's spatial variation(Madsen et al., 2011).

**4. Model development**

The fourth issue is the strategy of selecting predictor variables for the regression model. Clearly quality as well as availability of the GIS-derived predictor variables will affect LUR model performance. These data are usually collected by local authorities for regulatory purposes and are sometimes collected prior to the period of the measurement campaign(Hoek et al., 2008). The numbers and types of predictor variables that have been included in previous studies have differed substantially. Some variables which have a plausible relationship with air pollution are nearly always included, for example, traffic intensity on major roads. In the absence of traffic intensity data, some LURs have used length of main roads or distance to main roads as a surrogate(Morgenstern et al., 2007; Briggs DJ, Collins S, Elliott P, Fischer P, Kingham S, Lebert E, et al, 1997). In fact these three indicators of traffic were the top three most significant traffic-related predictors in most previous LUR studies as seen in Hoek's 2008 review(Hoek et al., 2008). Other variables, such as land cover and population density are usually, but not always, included in LUR models. It should be noted that, as Hoek et al pointed out, inclusion of certain variables (e.g. population density) in the LUR model may confound the relationship between air pollution exposure and health outcome. For example, population density may be related to area socioeconomic status, which may in turn have an effect on disease status. To minimise this potential confounding effect, it is therefore usual to adjust for area-level socioeconomic status in LUR-based epidemiological studies.

As mentioned before, the number of predictors included in the final model should also be considered carefully to avoid the model being over-specified(Basagaña et al., 2012; Wang et al., 2012). Another key problem here is how to decide the buffer size to generate a meaningful measure of each predictor

variable in GIS. In practice, researchers need to think carefully about the decay patterns of traffic-related air pollution in the study area. Hoek et al suggested that using buffer sizes which are more than 100-200m for traffic intensity may be meaningless in most European compact urban areas, as exposure beyond 100m of a major urban road may be rather homogeneous(Hoek et al., 2008). An exception are the so-called "street canyons"(Vardoulakis et al., 2003), where air pollutants are trapped by adjacent buildings resulting in a higher level of pollution at one side of the road than the other. To deal with this, Eeftens et al developed a GIS-based method to derive quantitative 'canyon indicators' which could be added to the LUR to capture this pollution trapping effect in street canyons(Eeftens et al., 2013).

As with any approach to exposure assessment, there are strengths and limitations in applying LUR modelling techniques for exposure assessment in epidemiological studies. In terms of strengths, the dedicated measurement campaign can be conducted at selected or preferred locations with rigorous algorithms in site selection and distribution. Secondly, LUR makes use of GIS techniques to extract location-specific information at each location and as such it captures local spatial variations well, especially in urban areas where GIS data are usually available in some detail. In fact, a recent study across most of Europe found that LUR model performs well, as does the conventional dispersion model, at least for $NO_2$ (r=0.76)(de Hoogh et al., 2014). Correlation was moderate (r=0.58) for PM estimates from both models. In terms of the limitations, LUR models usually provide long-term annual estimates rather than short-term estimates, as they usually do not take into account meteorology in the modelling. Also, as with other models, for instance, dispersion models, LUR do not reflect time-activity patterns or personal exposures, which may confound the associations explored in the epidemiological studies.

## 3.2. Land Use Regression models applied in this project

LUR modelling techniques were applied in this PhD project to estimate address-level exposure to ambient air pollution for the four participating cohorts. Air pollution estimates from this LUR modelling work were provided by Dr Kees de Hoogh from the MRC-PHE Centre for Environment and

Health at Imperial College London who led this environmental exposure modelling work for BioSHaRE and for ESCAPE. Details regarding the process of developing these LUR models have been published elsewhere(Beelen et al., 2013; Vienneau et al., 2013; Eeftens et al., 2012a), herein an overview summary of the specific LUR models developments for this project are briefly presented.

LUR modelling methods from the ESCAPE project (ESCAPE-LUR) were applied to participants of EPIC-Oxford, LifeLines and UK Biobank. Both the UK and the Netherlands were actively involved in the ESCAPE project at a national scale, and LUR models developed from the ESCAPE project were therefore applicable country-wide for both these countries. The study areas of the HUNT cohort were not part of the dedicated ESCAPE LUR modelling campaign in 2008-11, therefore the ESCAPE-LUR models were not able to be applied to participants in HUNT.

### *ESCAPE-LUR models for EPIC-Oxford and UK Biobank*

The ESCAPE air pollution measurement campaign was conducted in the London/Oxfordshire area during 26 January 2010 and 18 January 2011(Eeftens et al., 2012b; Cyrys et al., 2012). The London/Oxfordshire area, one of the 36 study areas across Europe for measuring $NO_2$, NOx, PM during the ESCAPE measurement campaigns, is mainly along the river Thames, which stretches east to the Greater London area and west to Oxfordshire (Figure 3.1). Overall, 41 sites were selected in this area, 27 of which were located in Greater London (11 street sites, 15 residential background sites and one reference site), 13 (5 street sites, 8 residential background sites) in the small-to-medium-sized towns west of London, and one regional background reference site in the countryside of Oxfordshire(Eeftens et al., 2012b; Cyrys et al., 2012). Measured air pollution levels and accompanying variables extracted from GIS (ArcGIS10) at each measurement site (see below tables) were used to develop the LUR models for this study area. This LUR model was then applied to the EPIC-Oxford/UK Biobank cohort addresses to obtain annual air pollution estimates for the year 2010.

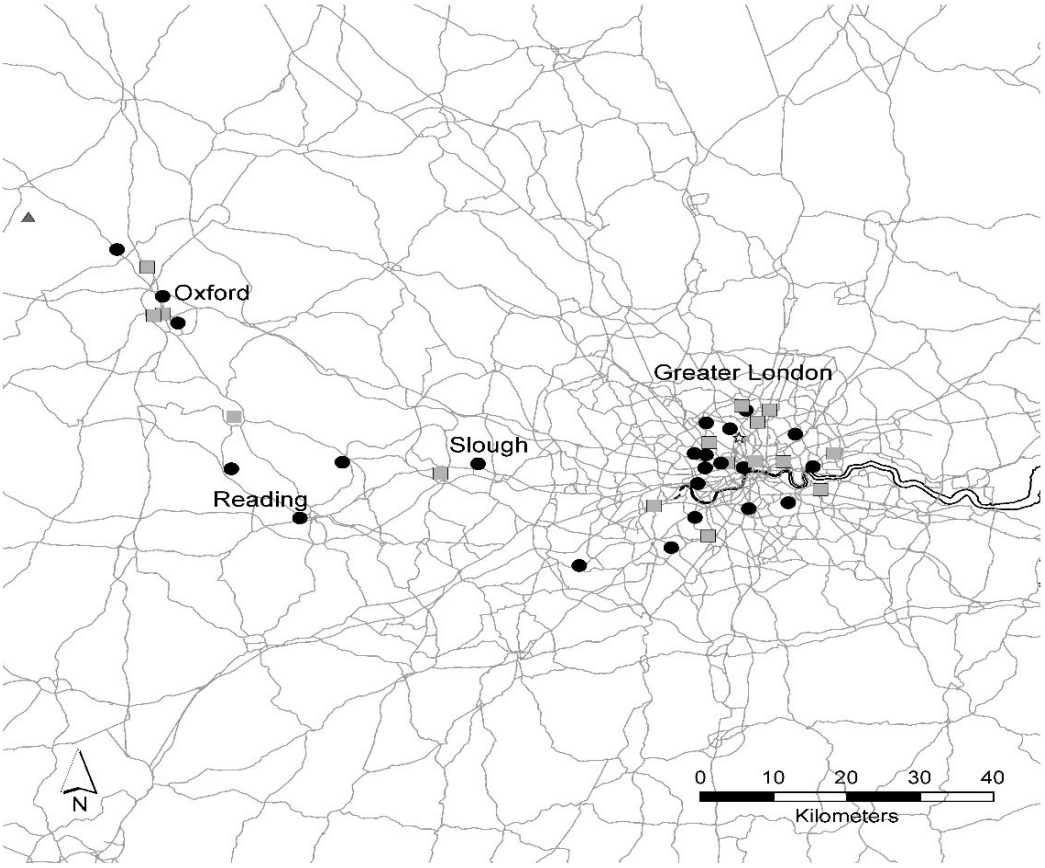



Figure 3.1 London/Oxford study area in the ESCAPE-LUR air pollution measurement campaign, as published by the ESCAPE project(Eeftens et al., 2012b; Cyrys et al., 2012) (grey-square: street site; black-dot: residential background; triangle: reference site)

$NO_2$ and NOx were measured using the Ogawa diffusion badge. At each monitoring site, $NO_2$ and NOx were measured for a consecutive two week period for each of the three seasons (cold, warm and intermediate)(Beelen et al., 2013). All badges were sent to the central laboratory of the ESCAPE project - the Institute for Risk Assessment Sciences (IRAS) in the Netherlands- for analysis, using ESCAPE standard operating procedures (SOPs). To adjust for temporal variation of air pollution, all measured estimates were adjusted using values from reference sites (either the urban site in London or rural background site in Oxfordshire) which operated continuously through the year.

Particulate matter (PM) was measured using Harvard impactors. As for $NO_2$ and NOx, PM was collected during three two-week periods in different seasons of the year(Eeftens et al., 2012a). In this London/Oxfordshire study area, $NO_2$, NOx and PM were all measured simultaneously. Sampling

volumes were calculated at the central laboratory at IRAS, following the established weighing and reflectance protocols developed for ESCAPE.

It should be noted that the ESCAPE estimates for particulate matter were considered valid up to 400km from the Greater London/Oxfordshire areas(Eeftens et al, 2012), but it is unclear how good the estimates are outside this area. All EPIC-Oxford/UK Biobank addresses which are more than 400km away from this Greater London/Oxfordshire area were therefore not assigned ESCAPE estimates for particulate matter.

Table 3.1 lists the GIS variables included in the final ESCAPE-LUR models in the London/Oxfordshire area for $NO_2$ and NOx. These variables were obtained from three main sources. The *central road network* consists of high resolution road data (mainly length and road classification) and were obtained from the Eurostreets V3.1 for the year 2008. Also centrally available was the *CORINE* (Coordination and Information on the Environmental programme) land use data, which was overseen by the European Commission. The *local road network* data (spatial resolution of 100m) were specified for each study area. For the London/Oxfordshire study area, traffic intensity data (vehicles/24hrs) for the year 2009 were obtained from the Road Traffic Statistics Branch at the Department for Transport in the UK.

Model performance ($R^2$ representing the percent of exposure variation that could be explained by the LUR model) was assessed by the leave-one-out cross-validation method. Briefly, the model was developed based on N-1 measurement sites, and this model used to predict measurements at the left-out site and this predicted value compared to the actual measured value. The procedure was repeated N times to calculate the overall $R^2$ across all measurement sites.

The model performance in the London/Oxfordshire areas (indicated by $R^2$) was 87% for the $NO_2$ model and 88% for the NOx model, which meant 87% and 88% of the exposure variations could be explained

by the variables included in the LUR models respectively. The exact LUR equations are listed in the reference paper(Beelen et al., 2013).

Table 3.1 GIS variables included in the final ESCAPE-LUR models for $NO_2$ and NOx for EPIC-Oxford and UK Biobank

| GIS variable name | Sources | Description |
|---|---|---|
| **$NO_2$** | | |
| **TRAFMAJORLOAD_500** | Local road network | Total traffic load of major roads † in a buffer of 500 metre (sum of (traffic intensity*length of all segments)) |
| **ROADLENGTH-500** | Central Road Network | Road length of all roads in a buffer of 500 metre. |
| **HLDRES_5000** | CORINE | Sum of high density and low density residential land in a buffer of 5000 metre |
| **NOx** | | |
| **TRAFLOAD_50** | Local road network | Total traffic load of all roads in a buffer of 50 metre (sum of (traffic intensity*length of all segments)). |
| **ROADLENGTH_300** | Central Road Network | Road length of all roads in a buffer of 300 metre. |
| **HLDRES_5000** | CORINE | Sum of high density and low density residential land in a buffer of 5000 metre |

† Definition of major roads in local road network: road with traffic intensity of >5000 motor vehicles (mvh)/24

Similarly, GIS variables included in the final ESCAPE-LUR models in the London/Oxfordshire area for each PM indicator are listed in Table 3.2. Model performance was particularly good for $PM_{2.5}$ absorbance, $PM_{10}$ and $PM_{2.5}$, with $R^2$ of 92%, 88% and 77% respectively. The LUR model for PM coarse only contained local traffic variables for the London/Oxfordshire area, with a $R^2$ of 57%. The exact LUR equations are listed in the reference paper(Eeftens et al., 2012a).

Table 3.2 GIS variables included in the final ESCAPE-LUR models for Particulate Matter for EPIC-Oxford and UK Biobank

| GIS variable name | Sources | Description |
|---|---|---|
| **$PM_{2.5}$** | | |
| **INTMAJORINVDIST** | Local road network | Product of traffic intensity (per 24 hour) on the nearest major road† (INTMAJOR) and inverse of distance to the nearest major road (INVDIST). |
| **ROADLENGTH_500** | Central Road Network | Road length of all roads in a buffer of 500 metre. |
| **$PM_{2.5}$ Absorbance** | | |
| **HEAVYTRAFLOAD_500** | Local road network | Total heavy-duty traffic load of all roads in a buffer of 500 metre (sum of (heavy-duty traffic intensity*length of all segments)). |

| | | |
|---|---|---|
| **HLDRES_5000** | CORINE | Sum of high density and low density residential land in a buffer of 5000 metre |
| **DISTINVMAJORC2** | Central Road Network | Distance to the nearest major road# |
| **PM coarse (as $PM_{10}$ minus $PM_{2.5}$)** | | |
| **DISTINVMAJOR1** | Local road network | Distance to the nearest major road † |
| **HEAVYTRAFMAJOR** | Local road network | Heavy-duty traffic intensity on nearest major road † |
| **$PM_{10}$** | | |
| **HEAVYTRAFMAJOR** | Local road network | Heavy-duty traffic intensity on nearest major road † |
| **HLDRES_300** | CORINE | Sum of high density and low density residential land in a buffer of 300 metre |
| **DISTINVMAJORC1** | Central Road Network | Distance to the nearest major road# |

† Definition of major roads in local road network: road with traffic intensity of >5000 motor vehicles (mvh)/24 hours. # Definition of major roads in central road network: classes 0, 1, and 2.

### *ESCAPE-LUR models for LifeLines*

Eight major cities (larger cities: Amsterdam, Rotterdam, Utrecht and Antwerp; smaller cities: Amersfoort, Groningen, Doetinchem and Maastricht) in the Netherlands and Belgium comprised one large study area (Dutch-Belgian study area) in the ESCAPE project (Figure 3.2). This study area as a

whole is a flat area with a high population density. In this study area, twenty regional background monitoring sites were selected in small villages and the countryside; ten sites (both urban background and street sites) were selected in each of the larger cities; in smaller cities, only six or four sites were selected per city(Eeftens et al., 2012b; Cyrys et al., 2012).

The ESCAPE air pollution measurement campaign was conducted in the Dutch-Belgian area during 17 February 2009 and 19 February 2010. The procedures for measurement and laboratory analysis of $NO_2$, NOx and PM were exactly the same as those for the London/Oxfordshire area as described earlier.

Measured estimates and variables extracted from GIS at each site were used to develop the LUR models for this study area. This LUR model was then applied to the LifeLines cohort addresses (mainly in the north of the Netherlands) to obtain annual air pollution estimates for the year 2009.

Figure 3.2 Dutch-Belgian study area in the ESCAPE-LUR air pollution measurement campaign, as published by the ESCAPE project(Eeftens et al., 2012b; Cyrys et al., 2012) (grey-square: street site; black-dot: residential background; triangle: reference site)

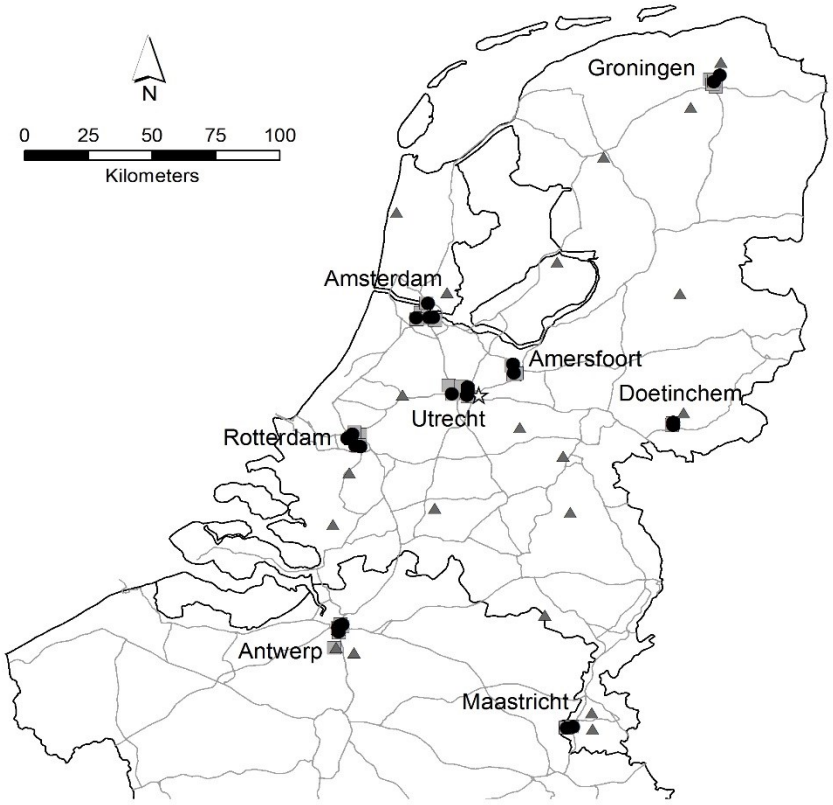



Table 3.3 displays the GIS variables included in the final ESCAPE-LUR models for $NO_2$ and NOx for the LifeLines cohort. Compared to the LUR models for the London/Oxfordshire area (Table 3.1), the final LUR models for $NO_2$ and NOx in the Netherlands included more local road network and population density GIS-derived variables. In the Netherlands, *local road network traffic intensities* data (mvh/24h) were obtained from PBL Netherlands Environmental Assessment Agency (PlanBureau voor de Leefomgeving) for the year of 2008. This extensive local road network data covers all roads across the Netherlands. *Local population density data* were obtained from the same source but for the year of 2009. As for the model for London/Oxfordshire area, the central road network data were obtained from Eurostreets V3.1 for the year 2008.

Model performance, as assessed using the leave-one-out cross validation method, was generally good for both models (81% for the $NO_2$ model and 82% for the NOx model). The exact LUR equations are displayed in the reference paper(Beelen et al., 2013).

Table 3.3 GIS variables included in the final ESCAPE-LUR models for $NO_2$ and NOx for LifeLines

| GIS variable name | Sources | Description |
|---|---|---|
| **$NO_2$** | | |
| **REGIONALESTIMATE** | Regional background monitoring sites | A regional background concentration estimate for each site location, based on inverse distance weighted interpolation of regional background sites. |
| **POP_5000** | Local population density data | Number of inhabitants within a buffer of 5000 metre. |
| **TRAFLOAD_50** | Local road network | Total traffic load of all roads in a buffer of 50 metre. |
| **ROADLENGTH_1000** | Central road network | Road length of all roads in a buffer of 1000 metre. |
| **HEAVYTRAFLOAD_25** | Local road network | Total heavy-duty traffic load of all roads in a buffer of 25 metre. |
| **DISTINVNEARC1** | Central road network | Distance to the nearest road |
| **HEAVYTRAFLOAD_25_500** | Local road network | Total heavy-duty traffic load of all roads within a buffer of 25-500 metre. |

| NOx | | |
|---|---|---|
| **REGIONALESTIMATE** | Regional background monitoring sites | A regional background concentration estimate for each site location, based on inverse distance weighted interpolation of regional background sites. |
| **TRAFLOAD_50** | Local road network | Total traffic load of all roads in a buffer of 50 metre. |
| **POP_1000** | Local population density data | Number of inhabitants within a buffer of 1000 metre. |
| **HEAVYTRAFLOAD_500** | Local road network | Total heavy-duty traffic load of all roads in a buffer of 500 metre. |
| **DISTINVMAJOR1** | Local road network | Distance to the nearest major road † |
| **MAJORROADLENGTH_25** | Central road network | Road length of major roads in a buffer of 25 metre. |

†: Definition of major roads in local road network: roads with traffic intensity of >5000 motor vehicles (mvh)/24 hours.

Table 3.4 lists the GIS variables included in the final ESCAPE-LUR model for each PM indicator for the LifeLines cohort. Model performance was particularly good for $PM_{2.5}$ absorbance with a $R^2$ of 89%, followed by $PM_{10}$ and $PM_{2.5}$, with $R^2$ of 60% and 61% respectively. As in the London/Oxfordshire area, the LUR model for PM coarse contained only local traffic variables and additionally the percent surface area of the local port in the Dutch-Belgian study area, with a relatively poor $R^2$ at 38%. The exact LUR equations are listed in the reference paper(Eeftens et al., 2012a).

Table 3.4 GIS variables included in the final ESCAPE-LUR models for Particulate Matter for LifeLines

| GIS variable name | Sources | Description |
|---|---|---|
| **$PM_{2.5}$** | | |
| **REGIONALESTIMATE** | Regional background monitoring sites | A regional background concentration estimate for each site location, based on inverse distance weighted interpolation of regional background sites. |
| **MAJORROADLENGTH_50** | Central Road Network | Road length of major roads in a buffer of 50 metre. # |
| **TRAFMAJORLOAD_1000** | Local road network | Total traffic load of major roads in a buffer of 1000 metre. † |
| **$PM_{2.5}$ Absorbance** | | |
| **TRAFLOAD_500** | Local road network | Total traffic load of all roads in a buffer of 500 metre. |
| **HLDRES_5000** | CORINE | Sum of high density and low density residential land in a buffer of 5000 metre |
| **MAJORROADLENGTH_50** | Central Road Network | Road length of major roads in a buffer of 50 metre. # |
| **REGIONALESTIMATE** | Regional background monitoring sites | A regional background concentration estimate for each site location, based on inverse |

| | | |
|---|---|---|
| | | distance weighted interpolation of regional background sites. |
| **HEAVYTRAFLOAD_50** | Local road network | Total heavy-duty traffic load of all roads in a buffer of 50 metre. |
| **PM coarse (as $PM_{10}$ minus $PM_{2.5}$)** | | |
| **TRAFLOAD_1000** | Local road network | Total traffic load of all roads in a buffer of 1000 metre. |
| **PORT_5000** | CORINE | Surface area ($m^2$) of port within 5000 metre |
| **TRAFNEAR** | Local road network | Traffic intensity (per 24 h) on the nearest road |
| **$PM_{10}$** | | |
| **TRAFMAJORLOAD_500** | Local road network | Total traffic load of major roads in a buffer of 500 metre. † |
| **POP_5000** | Local population density data | Number of inhabitants within a buffer of 5000 metre. |
| **MAJORROADLENGTH_50** | Central Road Network | Road length of major roads in a buffer of 50 metre. # |

†: Definition of major roads in local road network: road with traffic intensity of >5000 motor vehicles (mvh)/24 hours. # Definition of major roads in central road network: classes 0, 1, and 2.

### *Pan-European LUR models for all four cohorts*

For the HUNT cohort as well as for EPIC-Oxford, LifeLines and UK Biobank, a pan-European LUR model, using additional satellite-derived ground-level air pollution data, was developed by Dr. de Hoogh, and the resultant $NO_2$ and $PM_{10}$ estimates for 2005 to 2007 were also used in this PhD project.

This pan-European LUR model was enhanced by including satellite-based ground-level concentration of $NO_2$ and $PM_{2.5}$ (available on a 10km grid) as independent variables(Vienneau et al., 2013). The authors used satellite-derived $PM_{2.5}$ data in the $PM_{10}$ LUR model partly because the availability of data but also because that it was reported that $PM_{2.5}$ comprises up to 80% of $PM_{10}$ in Europe(Eeftens et al., 2012b). This LUR model was developed mainly for western European countries on a resolution of 100x100m. In brief, annual mean $NO_2$ and $PM_{10}$ data during 2005-2007 were obtained from over 1500 monitoring sites across Europe, which were centrally regulated and reported by EuroAirnet. Only those monitoring sites which captured over 75% of the total hours for $NO_2$ and over 75% of the days for $PM_{10}$ were included. These data served as dependent variables in the LUR modelling. An overview of the independent GIS-derived variables and satellite-derived air pollution data included in the final LUR models is presented below. Both $NO_2$ and $PM_{10}$ estimates were derived from these LUR models for year 2007 which were then applied to cohort participants in all four participating cohorts. Participant addresses were intersected with the modelled 100mx100m resolution exposure surface in GIS to obtain an exposure estimate.

Table 3.5 lists variables included in the final LUR model for $NO_2$ for the year 2007. GIS-derived variables included in the final LUR models for $NO_2$ in 2005, 2006 and 2007 across countries were highly consistent. In Great Britain, the model performance (assessed using the independent subset (20% of monitoring sites) reserved for this, indicated by $R^2$) for $NO_2$ for the year 2005 across all monitoring sites was 64%. In the Netherlands, this figure was 53%.

Table 3.5 GIS variables included in the final satellite-enhanced LUR model for $NO_2$ for year 2007

| GIS variable name | Sources | Description |
|---|---|---|
| **$NO_2$ (year 2007)** | | |
| **Minor roads 1500m** | Central Road Network: EuroStreets V3.1 | Lengths (metre) of all minor roads within 1500 metre |

| | | |
|---|---|---|
| **Major roads 100m** | Central Road Network: EuroStreets V3.1 | Lengths (metre) of all major roads within 100 metre |
| **Semi-natural land 600m** | CORINE | Semi-natural land (% of area) within a 600 metre buffer |
| **Minor roads 1500-10000m** | Central Road Network: EuroStreets V3.1 | Lengths (metre) of all minor roads within 1500-10,000 metre |
| **Total built up land 300m** | CORINE | Total built up, % of area (residential, industrial, port, airports, mines, dumps and construction sites) |
| **Satellite-derived surface $NO_2$ 2007** | OMI (Ozone Monitoring Instrument)derived from the Aura Satellite | Surface $NO_2$ concentration: OMI derived $NO_2$ (ppb) ~10km grid resolution |

Table 3.6 list variables included in the final LUR model for $PM_{10}$ for the year 2007. In general, final LUR model for $PM_{10}$ across all regions for the year 2007 yielded the best performance ($R^2$: 50%), compared with those in 2005 ($R^2$: 35%) and 2006 ($R^2$: 37%). In Great Britain, the model performance ($R^2$) for $PM_{10}$ of the year 2007 across all monitoring sites was 57%. In the Netherlands, this figure was 32%.

Table 3.6 GIS variables included in the final satellite-enhanced LUR model for $PM_{10}$ for year 2007

| GIS variable name | Sources | Description |
|---|---|---|
| **$PM_{10}$ (year 2007)** | | |
| **Minor roads 200-2500m** | Central Road Network: EuroStreets V3.1 | Lengths (metre) of all minor roads from 200-2500 metre |

| | | |
|---|---|---|
| **Minor roads 200m** | Central Road Network: EuroStreets V3.1 | Lengths (metre) of all minor roads within 200 metre |
| **Major roads*** | Central Road Network: EuroStreets V3.1 | Lengths (metre) of all major roads |
| **Altitude** | SRTM Digital Elevation Database V4.1 | Altitude of the geocoded address |
| **Tree canopy 100m** | Coarser Global land cover | % of area of tree canopy within 100 metre |
| **Y coordinate** | GIS database-ArcGIS10 | Y coordinates for 100m cell centroids |
| **Satellite-derived surface $PM_{2.5}$ 2001-2006** | Terra Satellite | Surface $PM_{2.5}$ concentration: Terra- derived and humidity-corrected $PM_{2.5}$ aggregated from 2001-2006 ~10km grid resolution |

*Buffer not stated in the reference paper

It should be noted that very few monitoring sites were available in Norway for the years 2005-2007 for model building/validation. In order to secure a harmonised air pollution dataset for all cohorts including HUNT, it was decided also to directly apply this pan-European LUR models for $NO_2$ and $PM_{10}$ to the addresses of HUNT participants at a 100mx100m resolution. This decision was justified because of the reasonable performance of the models in other Northern European countries including Sweden, Finland, Denmark and Latvia (personal communications from Dr Kees de Hoogh). As a result, there is no direct measure of performance of this pan-European LUR model in Norway as no Norwegian air pollution measurements were used in the model building. It is therefore likely that the models are less accurate in Norway.

### *Air pollution estimates used in this project*

As described above, two sets of air pollution data from each LUR model (ESCAPE-LUR and pan-European LUR) were developed for EPIC-Oxford, LifeLines and UK Biobank. While for the HUNT cohort, only estimates obtained from pan-European LUR model were available.

Each LUR model has its strengths and limitations. The ESCAPE-LUR was developed in many study areas across Europe, following standardised procedures in each study area in terms of site selection (e.g. street sites over-represented), model developments (e.g. forward stepwise) and validations (e.g. leave-one out cross-validation). Predictive variables included in the ESCAPE-LUR models came from both European-wide centrally available GIS-derived data as well as local GIS-derived data in each study area, as can be seen in Table 3.1 to 3.4. However, quantity and quality of local GIS-derived data differ across countries, e.g. data were not directly comparable in the UK and the Netherlands, and this will affect the model performance. These differences however were allowed in the data harmonisation process for this PhD project. Nevertheless, compared to earlier cross-country LUR studies such as SAVIAH, the ESCAPE-LUR had fine-scale GIS-derived data to model the spatial variations of air pollution across more study areas.

Given the fact that LUR models are in general spatially confined, Vienneau et al provided another approach in the developments of LUR over a very large area (e.g. at the continental level), using the centrally available GIS variables across the study areas and satellite-derived ground-level air pollution data(Vienneau et al., 2013). One advance of this pan-European LUR model is that it built on a very dense monitoring network across Western Europe to develop the LUR model at a fine spatial scale of a 100m grid. However, this was not necessarily reflected in terms of the model performance as measured by $R^2$ (30%-60% for both $NO_2$ and $PM_{10}$). Nevertheless, this pan-European LUR model in this specific context enables the assignments of harmonised air pollution exposures to participants of all four cohorts included in this PhD project.

Performance was generally better for ESCAPE-LUR than the pan-European LUR models, partly because more local GIS-derived data (e.g. traffic load) were included in the ESCAPE-LUR model to capture the local traffic effect. Whilst for the pan-European standardised LUR models, some lower resolution inputs were used, and no region specific modification of the model permitted by this approach.

Table 3.7 lists the standardised air pollution variables from both LUR models. These variables were then used in the health impact assessments (Chapter 5 & 6) in this PhD project. Averaged annual air pollution estimates from the ESCAPE-LUR model were for the year 2010 for EPIC-Oxford and UK Biobank, and for the year 2009 for LifeLines cohort.

Table 3.7 Standardised air pollution variables used in this project

| Harmonised name | Description |
|---|---|
| **NO2_ESCAPE** | Nitrogen dioxide; ESCAPE-LUR annual average estimate (μg/$m^3$) |
| **NOx_ESCAPE** | Nitrogen oxides; ESCAPE-LUR annual average estimate (μg/$m^3$) |
| **PM10_ESCAPE** | PM10 (particulate matter with diameter ≤10μm); ESCAPE-LUR annual average estimate (μg/$m^3$) |
| **PM25_ESCAPE** | PM2.5 (particulate matter with diameter ≤2.5μm); ESCAPE-LUR annual average estimate (μg/$m^3$) |
| **PM25abs_ESCAPE** | PM2.5 absorbance (measurement of the blackness of PM2.5 filters; a proxy for elemental carbon, which is the dominant light absorbing substance); ESCAPE-LUR annual average estimate ($m^{-1}$) |
| **PMcoarse_ESCAPE** | PM coarse (particulate matter 2.5-10μm); ESCAPE-LUR annual average estimate (μg/$m^3$) |
| **NO2_05** | Nitrogen dioxide; Satellite-enhanced LUR annual average estimate of the year 2005 (μg/$m^3$) |

| | |
|---|---|
| **NO2_06** | Nitrogen dioxide; Satellite-enhanced LUR annual average estimate of the year 2006 (µg/m$^3$) |
| **NO2_07** | Nitrogen dioxide; Satellite-enhanced LUR annual average estimate of the year 2007 (µg/m$^3$) |
| **PM10_07** | PM10 (particulate matter with diameter ≤10m); Satellite-enhanced LUR annual average estimate of the year 2007 (µg/m$^3$) |

## 3.3. Road traffic noise model applied in this project

The other exposure for this PhD project is the address-level road traffic noise. Various road traffic noise models have been proposed or updated in recent years and now are being employed in different research settings around the world. As a result of technological advancements, new models based on numerical methods are preferable to earlier non-numeric models, as highlighted in an article which critically reviewed various "classical" models from the late 1990s until more recent developments(Garg & Maji, 2014). The review commented that noise source and noise propagation ideally should be treated as two independent parts in the model which also allows subsequent separate updates. More importantly, the authors suggested that a harmonised approach for the noise propagation modelling is critical in future developments.

The Common Noise Assessment Methods in Europe (**CNOSSOS-EU**)(Kephalopoulos et al., 2014), initiated by European Commission, is working towards such a harmonised approach to noise mapping across the EU member states. CNOSSOS-EU aims to provide EU stakeholders in noise policy making with a set of consistent, comparable and reliable noise estimates within each and across the member states.

This PhD project is among the first to use an adaptation of the CNOSSOS-EU model to estimate standardised noise estimates across several countries, and subsequently to apply these noise estimates in health impact assessments (Chapter 5 & 6). Dr John Gulliver, of the MRC-PHE Centre for

Environment and Health at Imperial College, led the road traffic noise modelling work for BioSHaRE. I briefly summarise the general aspects of this modelling work below.

The CNOSSOS-EU model for road traffic noise assessment is an extremely detailed method (Figure 3.3), which generally assumes that high-resolution data inputs (e.g. traffic flow, land cover) are available for modelling. While such high-resolution data may be available for city-wide noise modelling, at a national or a wider geographical scale, such detailed data are not always available. In order to predict and harmonise noise estimates over a broader geographical area (e.g. across several countries as for this PhD project), Gulliver and colleagues made a simplification of this CNOSSOS-EU model, using some lower resolution data inputs(Morley et al., 2015). They tested the feasibility of this simplified version of the CNOSSOS-EU model, by comparing its performance with that of the model which had high-resolution data inputs, in the southwest part of the city of Leicester, UK. They concluded that, in epidemiological contexts, this simplified CNOSSOS-EU model is able to provide good performance for exposure ranking (Spearman rank: 0.75). However, the authors did acknowledge that, its ability to predict noise estimate may be relatively poor, indicated by a large RMSE (root mean square error) of 4.5 dB(A).

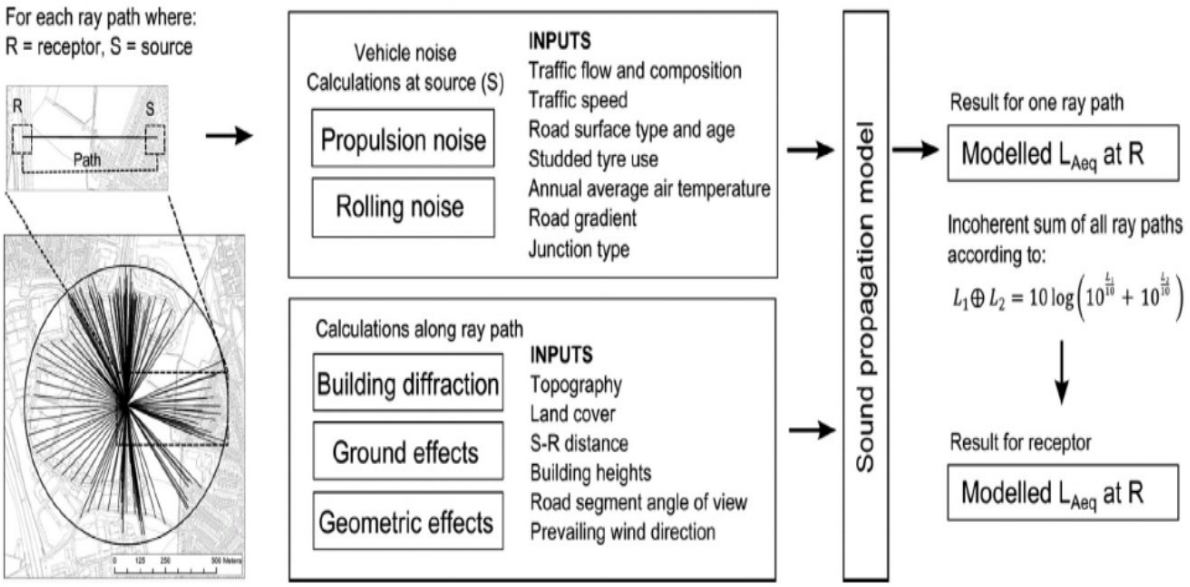



Figure 3.3 Working flows of the CNOSSOS-EU noise model, as published by Morley et al(Morley et al., 2015)

Based on this simplified version of the CNOSSOS-EU model, Gulliver and colleagues assigned road traffic noise estimates at address level to participants of all the four cohorts participating in this project. In brief, noise sound pressure level was estimated from all roads within 500 meters of home address at recruitment. Noise propagation due to refraction and diffraction, absorption from buildings, distance and angle of view were considered in the modelling. Road network geography, hourly vehicle flows, building heights, land cover and meteorological data were obtained for the respective study areas in the UK, the Netherlands and Norway (Table 3.8). To account for participants living on minor roads that were not captured in the national level traffic datasets, a fixed low-level baseline flow was assigned. Traffic data were for the year 2009 and land cover data were for the year 2006.

Five A-weighted noise indicators were estimated for each cohort participant:

- **Lday** (day-time noise sound level from 07:00-19:00),
- **Lnight** (night-time noise sound level from 23:00-7:00),
- **Leve** (evening noise sound level from 19:00-23:00),
- **Laeq16h** (noise sound level from 07:00-23:00)
- **Lden** (noise sound level over a 24h period, with a penalty of 5dB added for the evening hours and a penalty of 10dB added for night-time hours).

Since any two of these noise indicators are highly correlated (r=0.99), I only used two noise indicators, Lday and Lnight, in the main analyses, with an aim to investigate whether there exists a differential effect of each on the studied outcomes in this project.

Table 3.8 Input variables for road traffic noise model in each participating cohort

| **EPIC-Oxford and UK Biobank, UK** | |
|---|---|
| Variable | **Source** |
| Road network geography | ESCAPE project |
| Hourly vehicle flows | ESCAPE project modelled light and heavy vehicle flows in 2009 |
| Land cover | Corine land cover 2006 (v16) at 100m resolution |
| Building heights | Landmap (major urban areas available only) |

| | |
|---|---|
| Air temperature | UK Met office (annual average 2001-2010) |
| Prevailing wind direction | UK Met office (annual average 2001-2010) |
| **LifeLines, the Netherlands** | |
| Variable | **Source** |
| Road network geography | ESCAPE project |
| Hourly vehicle flows | ESCAPE project modelled light and heavy vehicle flows in 2009 |
| Land cover | CORINE land cover 2006 (v16) at 100m resolution |
| Building heights | Based on CORINE urban areas with constant height of 9.5m* |
| Air temperature | Wikipedia (annual average 2001-2010) |
| Prevailing wind direction | Wikipedia (annual average 2001-2010) |
| **HUNT, Norway** | |
| Variable | **Source** |
| Road network geography | Speed limits and 1-week average traffic counts were provided by road authority and polygons are provided by local municipality for spatial accuracy. |
| Hourly vehicle flows | Daily traffic flow data from road authority divided by 24 |
| Land cover | CORINE land cover 2006 (v16) at 100m resolution |
| Building heights | Based on CORINE urban areas with constant height of 9.5m* |
| Air temperature | Wikipedia (annual average 2001-2010) |
| Prevailing wind direction | Wikipedia (annual average 2001-2010) |

# Chapter 4 Statistical methods

Statistical analyses in this project were supported by the novel tools developed by BioSHaRE. These new novel tools enable virtual individual-level data pooling from various sources located in different centres, permitting the formation of a large and integrated database for research. In this PhD project, I applied one of these novel tools to the epidemiological analyses, demonstrating its effective use in data sharing and data analyses across cohorts, with data from these cohorts hosted in different countries. In this Chapter, I first introduce this new novel tool, DataSHIELD, and then detailed statistical methods are described with respects to each research aim.

## 4.1. DataSHIELD

As discussed in Chapter 2, there are generally two ways in combining research data from several studies. One is the study-level meta-analysis, in which summary statistics were provided by each study to contribute to the meta-analyses. This method is convenient, effective and usually used as a preferable option in international research consortia (e.g. the ESCAPE project), as data owners retain governance control over their data, allowing them to meet their binding ethical and/or legal guidelines. However, this method of combining data may not be as statistically flexible as combining individual-level data across studies. For example, in the study-level meta-analyses, all analyses conducted by participating studies must strictly follow the pre-defined analysis protocol before the effective pooling of the study-specific estimates. Any additional/*post hoc* analyses (subgroup or interactions) need to be requested at each participating study, which could be time-consuming and requires coordination of efforts across studies. Combining individual-level data also has some challenges for researchers. The major concern is the ethico-legal issues associated with physical sharing of data across studies nationally or internationally(Wolfson et al., 2010). The actual research work cannot even be started without sorting out those legally binding guidelines first, for instance, to get participants re-consent about this physical sharing of data with external partners.

To overcome this ethico-legal issues of combining individual-level data, DataSHIELD (Data Aggregation Through Anonymous Summary-statistics from Harmonised Individual levEL Databases) was developed

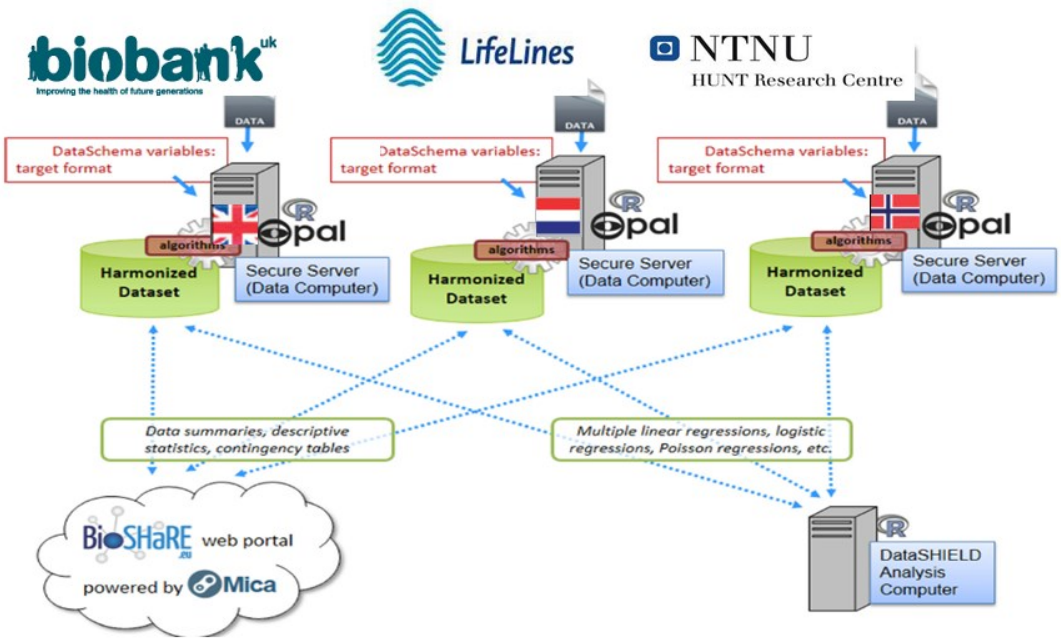
as a novel and practical tool to pool harmonised data from various cohorts for individual-level analyses(Gaye et al., 2014). One prerequisite for using DataSHIELD, is that data must have been harmonised across cohorts. I have described in Chapter 2 how the data harmonisation work across the four participating cohorts was conducted in this PhD project.

DataSHIELD has some important strengths(Gaye et al., 2014). First, this tool allows virtual, but not physical, pooling of harmonised data across studies, meaning that data are never physically shared between parties but rather remain in the local research sites. Data owners still retain control over their data and ethico-legal guidelines are therefore able to be met. This strength is further clarified by two papers. Wallace et al concluded that application of DataSHIELD meets the standards of sharing biomedical data in the UK(Wallace et al., 2014). Another paper has described how DataSHIELD could technically solve those general ethics-related issues with respect to data sharing, for example, the protection of a participant's privacy, confidentiality and right of the data (e.g. withdrawn from participation) when data were being shared and after data had been shared(Wolfson et al., 2010). Second, pooling of these harmonised data across studies forms a large, integrated and high-quality individual-level database for analyses. This will not only be statistically flexible but also a cost-effective approach in maximising the use of existing data resources for research discovery and advancements. Third, DataSHIELD aims to provide most of the analysis package available within other conventional statistical tools (R, Stata etc.), and results generated by DataSHIELD should therefore comparable to those using the conventional statistical tools.

Workflows of the virtual data pooling for this PhD project are shown in Figure 4.1. First of all, each obtained dataset was deposited to a local secure server called "OPAL" (http://www.p3g.org/biobank-toolkit/opal accessed May 2016), an open-source software which could be installed inside the 'firewall' within each local research site in the respective countries. Second, as described in chapter 2, a dataschema with a harmonised definition for each variable was administered, by running programming scripts, to each of the original cohorts to derive a harmonised dataset (as shown in the

green boxes in the figure). Each harmonised dataset could be checked and updated via the BioSHaRE web portal. Third, researchers could then log on to the DataSHIELD platform (equipped with Rstudio) from a working computer to pool harmonised datasets on each local OPAL server to form a virtual database for analyses. Log-on credentials for DataSHIELD are generated specifically for the named researchers only, neither DataSHIELD developers nor other unnamed researchers will have access to this designated analysis account. DataSHIELD is based on the conventional R statistical packages, with DataSHIELD-specific R statistical codes (e.g. codes that prevent the inadvertent disclosure of participant specific details, prevent any attempt to re-construct the dataset, etc.) written to run the required statistical analyses. These slightly revised R codes remain similar to those in the normal R statistical packages, making it easy for researchers easily to adapt to the DataSHIELD statistical environment.

Figure 4.1 Workflows of virtual data pooling using various BioSHaRE tools for this PhD project




2

This PhD work was made feasible by the application of DataSHIELD for data pooling across cohorts to permit individual-level pooled data analyses. In fact, this project is one of the first to test the DataSHIELD application in a real-life epidemiological project. Earlier work conducted by DataSHIELD developers and BioSHaRE colleagues demonstrated its application in a project studying healthy obesity among European populations in 10 cohorts from seven countries(van Vliet-Ostaptchouk et al., 2014). This PhD project, with fewer cohorts but a far larger population, required the development of additional statistical functionality in DataSHIELD to permit the incident analyses for cardiovascular disease. This project therefore provided an informative exemplar for future studies in which DataSHIELD will be used as a statistical tool.

## 4.2. Statistical analyses: blood biochemistry markers

At the time of conducting this PhD project, blood biochemistry data were only available in HUNT3 and LifeLines (Table 2.1). Serum concentrations of hsCRP (mg/L), total cholesterol (mmol/L), triglycerides (mmol/L), and high-density lipoprotein (HDL) cholesterol (mmol/L) were measured in both cohorts. Four outcomes were subsequently harmonised as, "**LAB_TSC**" (total serum cholesterol), "**LAB_TRIG**" (triglycerides), "**LAB_HDL**" (High-density lipoprotein cholesterol) and "**LAB_HsCRP**" (high-sensitivity C-reactive protein). Harmonised data from both cohorts were then virtually pooled via DataSHIELD. Each outcome was analysed on a continuous scale. Based on the distributions, the natural logarithmic value of hsCRP was used to achieve an approximate normal distribution.

In addition, for LifeLines, fasting blood glucose (mmol/L) and glycated haemoglobin (HbA1c) concentrations (mmol/mol) were also measured. Analyses on these two outcomes were conducted in LifeLines only.

Spearman correlations between metrics of ambient air pollution and road traffic noise were calculated for each cohort. Cross-sectional associations between the pan-European LUR modelled $PM_{10}$, $NO_2$ (both for year 2007) or noise (for year 2009) and each biochemical parameter were analysed using multivariate linear regression. Both air pollution and noise metrics were analysed on a continuous

scale, assuming a linear effect. Additionally, noise was categorised for Lday (<55, 55-60, ≥60 dB(A)) and Lnight (<45, 45-50, ≥50 dB(A)).

The covariates were chosen *a priori* based on current knowledge, and harmonised as described in Chapter 2. The sequence of models was as follows:

**Model 1**: adjusted for cohort (in pooled analyses on hsCRP and blood lipids) or unadjusted (in analyses on blood glucose and HbA1c)

**Model 2**: adjusted for cohort, age and sex

**Model 3**: adjusted for cohort, age and sex, and further adjusted for season of blood draw, smoking status and pack-years, education level, paid employment and weekly alcohol consumption (**main model**)

Based on Model 3, metrics of traffic noise (or air pollution) were additionally adjusted for in the air pollution (or noise) models.

Sensitivity analyses were conducted based on model 3: a) further adjusting for BMI; b) further adjusting for ever-had hypertension or diabetes; c) restricting analyses to those living at the same address ≥10 years; d) excluding those with an hsCRP over 10 mg/ L from the analysis as levels above this may indicate a current infection as with a previous study(Lanki et al., 2015).

Based on model 3, potential effect modifications by: i) sex, ii) age <60 or ≥60 years, iii) BMI <25, 25-30, ≥30 kg/$m^2$, iv) diabetes, and v) hypertension were examined by inclusions of respective interaction terms in the models.

I also conducted study-specific analyses for hsCRP and blood lipids and then pooled estimates via meta-analysis in R v3.2.2. Further, I examined associations between air pollutants from the ESCAPE model and each biochemical parameter in the LifeLines cohort only, following the analysis steps as above.

## 4.3. Statistical analyses: incident CVD outcomes

Three specific incident CVD outcomes, total incident cardiovascular disease ('FAILURE_CVD', ICD-10: I00-I99), incident ischaemic heart disease ('FAILRUE_MI', ICD10:I20-I25) and incident cerebrovascular diseases ('FAILURE_STROKE', ICD10: I60-I69), were analysed in relation to ambient air pollution and/or road traffic noise exposures. Ascertainment and harmonisation of these outcomes across the three participating cohorts (EPIC-Oxford, HUNT2 and UK Biobank) were described in Chapter 2 and summarised in Table 2.6.

Data were analysed in parallel using three different statistical procedures. First, I physically pooled the harmonised data from all three cohorts and undertook a pooled individual-level analyses via Stata using Cox proportional hazards regression models.

Second, I virtually pooled and analysed the harmonised data via DataSHIELD using a newly developed function -piecewise exponential regression- through a Poisson regression model. This new function permits a non-disclosive co-analysis of sensitive individual-level data. The survival time-span of an incident event is divided through DataSHIELD into pre-specified sub-intervals where the baseline hazard in each sub-interval is assumed to be constant but can vary across different sub-intervals. The choice of widths of sub-intervals is therefore crucial to allow baseline hazards to be approximately constant within each of them. Nevertheless, even a division of survival time-span to an arbitrary set of sub-intervals gives qualitatively similar results to Cox's regression analysis (personal communications from DataSHIELD developers-Demetris Avraam and Paul Burton).

Third, I analysed each cohort separately using Cox models and then pooled the cohort-specific estimates via meta-analysis methods. Effect estimates obtained from these three statistical approaches were then compared in this study.

In both the pooled individual-level analyses (conducted in both Stata and DataSHIELD) and the cohort-specific analyses (conducted in Stata), adjusted regression models were defined *a priori* based on

current knowledge from air pollution/noise-related CVD studies. Three models with increasing level of adjustments were specified:

**Model1**: adjusted for cohort;

**Model2**: adjusted for cohort, age, sex and calendar year of recruitment;

**Model3**: adjusted for cohort, age, sex and calendar year of recruitment, education level, employment status, smoking status and weekly alcohol consumption (**main model**).

As for the biochemistry analyses described above, in the main model, metrics of traffic noise (or air pollution) were additionally adjusted for in the air pollution (or noise) models.

For all the analyses, the pan-European LUR modelled $PM_{10}$, $NO_2$ (both for year 2007) and Lday were used as the main exposure estimates in the regression models. Additionally, noise was categorised for Lday (<50, 50-55, ≥55 dB(A)) and Lnight (<40, 40-45, ≥45 dB(A)).

BMI and ever-had diabetes may be on the causal pathway between air pollution/noise and CVD outcomes. BMI and ever-had diabetes were therefore further included in the main model in sensitivity analyses. To compare my findings with previous studies, in particularly those from the ESCAPE studies(Cesaroni et al., 2014; Stafoggia et al., 2014), I also undertook additional sensitivity analyses by restricting the studied outcomes to incident acute coronary events (ICD10: I20.0, I21, I23, I24), and to different subtypes of cerebrovascular diseases, including ischaemic stroke (ICD10: I63), haemorrhagic stroke (ICD10: I60, I61, I62) and unspecified stroke (ICD10: I64).

Effect modifications were investigated by adding an interaction term with exposure to the main model. Potential effect modifiers identified *a priori* were sex, age (<60 and ≥60 years), BMI (<25, 25-30, ≥30 kg/$m^2$), smoking (never-, ex- and current-), education (low, medium, high) and ever-had diabetes (yes, no).

Since air pollutants from the ESCAPE models were also available for both UK cohorts (EPIC-Oxford and UK Biobank), I repeated all the above analyses using the ESCAPE metrics, by physically pooling these two cohorts in Stata.

All the analyses were done using DataSHIELD v4.1.2 and Stata v12.1, Texas, USA.

# Chapter 5 Findings and discussion: blood biochemistry markers

## 5.1. The findings

### 5.1.1. Descriptive statistics

Altogether, pooling data from both HUNT3 and LifeLines, there was a total of 144,082 participants who had questionnaire data as shown in Table 5.1. The mean age was 47.6 years; 56% of the total population were women; the mean BMI was 26.5 kg/m$^2$, and nearly 60% of the participants were classified as overweight or obese (BMI≥25 kg/m$^2$); 24% of the pooled population were current smokers; Approximately 90% of the participants had received at least some secondary education and 75% of the pooled population were currently in paid employment. Prevalence of pre-existing health conditions was generally higher in HUNT3 than in LifeLines except for ever-had hypertension.

Table 5.1 Baseline characteristics: pooling data from both HUNT3 and LifeLines

| | Pooled data | HUNT3 | LifeLines |
|---|---|---|---|
| Total N | 144,082 | 50,805 | 93,277 |
| Age, years, mean(SD) | 47.6 (13.7) | 52.7 (16.7) | 44.9(12.3) |
| Sex, women, [n (%)] | 82,574 (56%) | 27,756 (55%) | 54,818 (59%) |
| Waist size, cm, mean(SD) | 91.5 (12.4) | 93.6 (12.3) | 90.3 (12.4) |
| BMI, kg/m$^2$, mean(SD) | 26.5 (4.4) | 27.2 (4.4) | 26.1 (4.3) |
| BMI categories, [n (%)] | | | |
| **<25 kg/m$^2$** | 58,629 (41%) | 16,481 (33%) | 42,148 (45%) |
| **25-30 kg/m$^2$** | 58,982 (41%) | 22,356 (44%) | 36,626 (39%) |
| **≥30 kg/m$^2$** | 26,049 (18%) | 11,575 (23%) | 14,474 (16%) |
| Smoking status, [n (%)] | | | |
| **Never-smoker** | 60,370 (44%) | 21,053 (43%) | 39,317 (45%) |
| **Ex-smoker** | 43,730 (32%) | 16,114 (32%) | 27,616 (32%) |
| **Current-smoker** | 32,144 (24%) | 12,208 (25%) | 19,936 (23%) |
| Current working status, [n (%)] | | | |
| **Not in paid employment** | 35,559 (25%) | 18,096 (36%) | 17,463 (19%) |
| **In paid employment** | 106,914 (75%) | 32,440 (64%) | 74,474 (81%) |
| Education level, [n (%)] | | | |
| **primary education or less** | 13,474(11%) | 11,221 (31%) | 2,252 (2%) |
| **Secondary education** | 78,565 (62%) | 16,826 (47%) | 61,739 (68%) |
| **Post-secondary school or above** | 35,287(28%) | 8,154 (22%) | 27,133 (30%) |
| Alcohol consumption, gram per week, mean(SD) | 31.6 (88.8) | 28.0 (40.0) | 56.9 (71.5) |
| Ever-had hypertension, [n (%)] | 31,070 (24%) | 11,099 (22%) | 19,971 (25%) |
| Ever-had diabetes, [n (%)] | 4,383 (3.0%) | 2,264 (4.5%) | 2,119 (2.3%) |
| Ever-had stroke, [n (%)] | 2,035 (1.4%) | 1,391 (2.7%) | 644 (0.7%) |
| Ever-had myocardial infarction, [n (%)] | 2,593 (1.8%) | 1,702 (3.4%) | 891 (1.0%) |

For both total cholesterol and triglycerides, participants in HUNT3 had a higher median than participants in LifeLines (Table 5.2). The median hsCRP, total cholesterol, triglycerides, HDL cholesterol was 1.2 mg/l, 5.1 mmol/l, 1.1 mmol/l and 1.4 mmol/l respectively in the pooled dataset. In LifeLines, the median (IQR) for fasting blood glucose was 4.9 (0.6) mmol/L while for HbA1c it was 37.0 (4.0) mmol/mol.

Table 5.2 Descriptive statistics (Median, IQR) for each blood biochemical parameter by cohort and in the pooled dataset

| | **Pooled** | | **HUNT3 (n=50,805)** | | **LifeLines (N=93,277)** | |
|---|---|---|---|---|---|---|
| | N* | Median (IQR) | N | Median (IQR) | N | Median (IQR) |
| **hsCRP, mg/L** | 104,580 | 1.2 (2.1) | 50,059 | 1.2 (2.1) | 54,521 | 1.2 (2.2) |
| **Total cholesterol, mmol/L** | 142,251 | 5.1 (1.5) | 49,346 | 5.4 (1.5) | 92,905 | 5.0 (1.4) |
| **Triglycerides, mmol/L** | 142,965 | 1.1 (0.8) | 50,060 | 1.4 (1.0) | 92,905 | 0.9 (0.7) |
| **HDL cholesterol, mmol/L** | 142,249 | 1.4 (0.4) | 49,345 | 1.3 (0.4) | 92,904 | 1.4 (0.5) |
| **Fasting blood glucose, mmol/L** | - | - | - | - | 90,260 | 4.9 (0.6) |
| **HbA1c, mmol/mol** | - | - | - | - | 69,970 | 37.0(4.0) |

*N: number of participants who had the measure of respective biochemical parameter

The pooled median pan-European $PM_{10}$ and $NO_2$ exposure was 18.8 and 17.2 µg/$m^3$, with an inter-quartile range (IQR) of 2.0 and 7.4 µg/$m^3$ respectively (Table 5.3). Pooled median daytime noise (Lday) and night-time noise (Lnight) was 51.6 and 43.4 dB(A), with an IQR of 5.1 and 4.4 dB(A) respectively.

Spearman correlations between $PM_{10}$ and Lday were r= 0.04 (HUNT3) and 0.38 (Lifelines), and between $NO_2$ and Lday were r= -0.05 (HUNT3) and 0.43 (Lifelines) (Table 5.4). The correlation between the pan-European LUR modelled $NO_2$ and $PM_{10}$ was r=0.8 and between Lday and Lnight r=0.99 in both cohorts. For all analyses using continuous noise scales, Lnight effects were similar to Lday effects, and are therefore not reported here.

Table 5.3 Distributions of exposures (pan-Europe LUR modelled $PM_{10}$, $NO_2$ and road traffic noise) by cohort and in the pooled dataset

| | N | 5% | 10% | 25% | 50% | 75% | 90% | 95% | Mean(SD) | IQR |
|---|---|---|---|---|---|---|---|---|---|---|
| **$PM_{10}$,µg/m$^3$** | | | | | | | | | | |
| **HUNT3** | 50,567 | 9.7 | 10.0 | 10.4 | 11.2 | 12.0 | 12.6 | 12.9 | 11.3 (1.1) | 1.6 |
| **LifeLines** | 61,927 | 21.0 | 21.4 | 22.3 | 23.6 | 24.7 | 25.7 | 26.5 | 23.6 (1.7) | 2.4 |
| **Pooled** | 112,494 | 16.6 | 17.0 | 17.7 | 18.8 | 19.7 | 20.6 | 21.2 | 18.8 (1.5) | 2.0 |
| **$NO_2$, µg/m$^3$** | | | | | | | | | | |
| **HUNT3** | 50,628 | 8.2 | 8.8 | 10.1 | 11.9 | 15.4 | 18.6 | 19.5 | 13.0 (3.9) | 5.3 |
| **LifeLines** | 62,212 | 13.6 | 14.1 | 16.6 | 20.6 | 25.4 | 28.9 | 31.1 | 21.2 (5.7) | 8.8 |
| **Pooled** | 112,840 | 11.5 | 12.1 | 14.1 | 17.2 | 21.5 | 24.9 | 26.6 | 18.0 (5.1) | 7.4 |
| **Daytime noise(Lday,dB(A))** | | | | | | | | | | |
| **HUNT3** | 45,644 | 39.1 | 39.5 | 43.6 | 47.4 | 50.3 | 52.9 | 54.6 | 47.0 (4.9) | 6.7 |
| **LifeLines** | 74,744 | 51.3 | 51.7 | 52.4 | 53.9 | 56.6 | 60.4 | 63.9 | 55.2 (4.0) | 4.2 |
| **Pooled** | 120,388 | 47.0 | 47.3 | 49.3 | 51.6 | 54.4 | 57.7 | 60.6 | 52.3 (4.3) | 5.1 |
| **Night-time noise(Lnight, dB(A)** | | | | | | | | | | |
| **HUNT3** | 45,644 | 35.1 | 35.2 | 37.5 | 40.2 | 42.5 | 44.8 | 46.4 | 40.3 (3.7) | 5.0 |
| **LifeLines** | 74,744 | 42.5 | 42.8 | 43.6 | 45.1 | 47.8 | 51.6 | 55.1 | 46.4 (4.0) | 4.2 |
| **Pooled** | 120,388 | 39.8 | 40.1 | 41.5 | 43.4 | 45.9 | 49.2 | 52.0 | 44.2 (3.9) | 4.4 |

Table 5.4 Spearman correlations between air pollutants and road traffic noise exposures

| **HUNT3 (N=45,581)** | **$NO_2$** | **$PM_{10}$** | **Lday** | | |
|---|---|---|---|---|---|
| $NO_2$ | - | | | | |
| $PM_{10}$ | 0.80 | - | | | |
| Lday | -0.05 | 0.04 | - | | |
| Lnight | -0.04 | 0.06 | 0.99 | | |
| **LifeLines (N=62,653)** | $NO_2$ | $PM_{10}$ | $NO_2$_ESCAPE | $PM_{10}$_ESCAPE | Lday |
| $NO_2$ | - | | | | |
| $PM_{10}$ | 0.78 | - | | | |
| $NO_2$_ESCAPE | 0.86 | 0.78 | - | | |
| $PM_{10}$_ESCAPE | 0.67 | 0.54 | 0.73 | - | |
| Lday | 0.43 | 0.38 | 0.56 | 0.57 | - |
| Lnight | 0.46 | 0.40 | 0.61 | 0.64 | 0.99 |

$NO_2$, $PM_{10}$: pan-European LUR modelled metrics for year 2007, available for both cohorts; $NO_2$_ESCAPE, $PM_{10}$_ESCAPE: ESCAPE LUR modelled metrics for 2010, available for LifeLines only. Lday: daytime (07:00-19:00) noise, Lnight: night-time (23:00-07:00) noise.

### 5.1.2. Ambient air pollution effects

Analytical results for the main model (Model3) were presented in this section; results for Model1 and Model2 are available in Appendix-5.1.

- **hsCRP**

In Model3, after adjusting for age, sex, education, employment, smoking status, smoking pack-years, alcohol consumption, season of blood draw and cohort, increasing pan-European LUR modelled $PM_{10}$ and $NO_2$ were both significantly associated with increased levels of hsCRP (Table 5.5). Each IQR increase in $PM_{10}$ or $NO_2$ was positively associated with hsCRP levels (1.4%, 95%CI: 0.1% to 2.7%) and (1.9%, 95%CI: 0.5% to 3.3%) respectively. Statistical significance remained for $NO_2$ after further adjustment for Lday.

- **Blood lipids**

No associations were seen between any of the exposure measures and total cholesterol levels (Table 5.5). In contrast, positive significant associations with HDL cholesterol were observed for both air pollutants, though the association with $PM_{10}$ became null after further adjustment for Lday. Each IQR higher pan-European LUR modelled $NO_2$ was significantly associated with a 0.007mmol/l higher HDL cholesterol level (95%CI: 0.003 to 0.011), independent of adjustment for Lday.

For each IQR increase in pan-European LUR modelled $PM_{10}$, triglycerides increased by 0.021mmol/L (95%CI: 0.013 to 0.028); a similar association was also seen for $NO_2$ (Table 5.5). Both associations were robust to adjustment for Lday.

- **Fasting glucose and HbA1c**

Higher ambient $PM_{10}$ and $NO_2$ exposure was significantly associated with higher fasting glucose levels but not with HbA1c (Table 5.5). For each IQR increase in the pan-European LUR modelled $PM_{10}$ and $NO_2$, fasting glucose increased by 0.029 mmol/L (95%CI: 0.020 to 0.038) and 0.033 mmol/L (95%CI: 0.022 to 0.042) respectively; significant associations remained after adjusting for Lday.

Table 5.5 Associations between per IQR increase of pan-European LUR modelled $PM_{10}$, $NO_2$ and each blood biochemistry marker

| | $PM_{10}$, µg/m³ | | | | $NO_2$, µg/m³ | | | |
|---|---|---|---|---|---|---|---|---|
| **Pooled analysis (HUNT3 and LifeLines)** | N | IQR | Model3 | Model3+Lday | N | IQR | Model3 | Model3+Lday |
| HsCRP, % diff | 51,238 | 2 | **1.4% (0.1% to 2.7%)** | 0.9% (-0.4% to 2.3%) | 51,459 | 7.4 | **1.9% (0.5% to 3.3%)** | **1.7% (0.2% to 3.2%)** |
| Total cholesterol, mmol/l | 72,551 | 2 | -0.005 (-0.014 to 0.004) | -0.005 (-0.015 to 0.004) | 72,783 | 7.4 | -0.001 (-0.011 to 0.009) | -0.001 (-0.011 to 0.010) |
| HDL cholesterol, mmol/l | 72,551 | 2 | **0.003 (0.0002 to 0.007)** | 0.002 (-0.002 to 0.005) | 72,783 | 7.4 | **0.008 (0.005 to 0.012)** | **0.007 (0.003 to 0.011)** |
| Triglycerides, mmol/l | 72,794 | 2 | **0.021 (0.013 to 0.028)** | **0.020 (0.012 to 0.028)** | 73,026 | 7.4 | **0.021 (0.013 to 0.030)** | **0.021 (0.012 to 0.030)** |
| **LifeLines only** | | | | | | | | |
| Fasting glucose mmol/l | 52,234 | 2.4 | **0.029 (0.020 to 0.038)** | **0.024 (0.014 to 0.034)** | 52,543 | 8.8 | **0.033 (0.022 to 0.042)** | **0.027 (0.016 to 0.038)** |
| HbA1c mmol/mol | 43,481 | 2.4 | 0.008 (-0.049 to 0.064) | 0.012 (-0.051 to 0.073) | 43,537 | 8.8 | -0.020 (-0.083 to 0.044) | -0.022 (-0.093 to 0.049) |

Model3: adjusted for age, sex, education, employment, smoking status, smoking pack-years, alcohol consumption, season of blood draw and study (for pooled analysis only).
Lday: day-time road traffic noise (07:00-19:00).
Bold indicates where significance level<0.05

- Sensitivity analyses

Results from most sensitivity analyses were similar to the main findings presented in Table5.5, except the significant associations between both air pollutants and HDL cholesterol were lost after further adjustment for BMI. All sensitivity analyses are shown in Appendix-5.2.

- Subgroup analyses

Effects of both air pollutants on hsCRP, total cholesterol, HDL cholesterol and fasting glucose were modified by sex (all p-values for interaction<0.05), with stronger associations consistently found for women (Table 5.6). For total cholesterol, increased air pollution exposure was significantly associated with higher levels among women but lower levels among men. No statistically significant effect modifications by sex were found for triglycerides and HbA1c (Appendix-5.3). Inconsistent patterns of effect modification by BMI were found among the three lipids markers. A stronger association with total cholesterol or HDL cholesterol was observed among those with a BMI<25 kg/$m^2$ whilst for triglycerides a stronger association was observed among those with a BMI≥30 kg/$m^2$ (Appendix-5.3). No significant effect modifications effects were found based on age, ever-had diabetes or hypertension.

Table 5.6 Associations between pan-European LUR modelled $PM_{10}$, $NO_2$ and blood biochemistry marker by sex

| | | N | $PM_{10}$, µg/m$^3$, per IQR | | $P_{interaction}$ | N | $NO_2$, µg/m$^3$, per IQR | | $P_{interaction}$ |
|---|---|---|---|---|---|---|---|---|---|
| HsCRP, %diff | men | 22,116 | 1.03% | -0.84% to 2.94% | 0.00 | 22,221 | 1.05% | -0.93% to 3.15% | 0.00 |
| | women | 29,122 | **1.93%** | **0.19% to 3.70%** | | 29,238 | **2.82%** | **0.92% to 4.77%** | |
| Total cholesterol, mmol/l | men | 31,170 | **-0.026** | **-0.040 to -0.011** | 0.00 | 31,278 | **-0.029** | **-0.045 to -0.013** | 0.00 |
| | women | 41,381 | **0.012** | **0.001 to 0.024** | | 41,505 | **0.020** | **0.008 to 0.033** | |
| HDL cholesterol, mmol/l | men | 31,170 | -0.003 | -0.007 to 0.002 | 0.00 | 31,278 | -0.0002 | -0.005 to 0.005 | 0.00 |
| | women | 41,381 | **0.006** | **0.002 to 0.011** | | 41,505 | **0.013** | **0.008 to 0.018** | |
| Fasting glucose, mmol/l | men | 21,908 | 0.020 | 0.004 to 0.036 | 0.04 | 22,009 | 0.020 | 0.002 to 0.038 | 0.02 |
| | women | 30,326 | **0.037** | **0.026 to 0.048** | | 30,444 | **0.042** | **0.030 to 0.054** | |

Model3: adjusted for age, education, employment, smoking status, smoking pack-years, alcohol consumption, season of blood draw and study (for pooled analysis only).
IQR for $PM_{10}$ and $NO_2$ was 2 and 7.4 µg/m$^3$ for analyses of hsCRP and blood lipids; IQR for $PM_{10}$ and $NO_2$ was 2.4 and 8.8 µg/m$^3$ for analyses of fasting glucose.
BOLD indicates where significance level<0.05.

### 5.1.3. Road traffic noise effects

As with air pollution, analytical results for the main model (Model3) are presented in this section; results for Model1 and Model2 are available in the Appendix-5.1.

- hsCRP

Each IQR increase in Lday was significantly associated with hsCRP levels (1.1%, 95%CI: 0.02% to 2.2%), but association became null when air pollution was further adjusted for (Table 5.7).

- Blood lipids

No association between Lday and total cholesterol level was seen(Table 5.7). In contrast, each IQR higher Lday was significantly associated with a 0.007 mmol/l higher HDL cholesterol level (95%CI: 0.004 to 0.010). Effect estimate was slightly reduced after adjustment for $PM_{10}$ or $NO_2$, but the association was still statistically significant. Higher Lday exposure was also significantly associated with higher triglycerides levels (0.008mmol/l per IQR higher Lday, 95%CI: 0.001 to 0.015), however, this association was no longer statistically significant after adjustments for air pollution effects.

- Fasting glucose and HbA1c

As with air pollution, a statistically significant positive association was observed between Lday and fasting glucose, but no association was seen with HbA1c (Table5.7). Each IQR higher Lday exposure was significantly associated with a 0.013 mmol/l higher fasting glucose (95%CI: 0.006 to 0.019). This significant association was robust to adjustments for air pollution.

- Sensitivity analyses and subgroup analyses

The previously observed significant positive association between Lday and triglycerides or HDL became null when analyses were restricted to those living at the same address ≥10 years, but significance remained for fasting glucose. Other sensitivity analyses only resulted in minor changes to the main findings (Appendix-5.2). Stronger associations between Lday and each biochemistry marker

were generally seen among women or those with a BMI<25 kg/m$^2$, although not all interactions were statistically significant (Appendix-5.3).

Table 5.7 Associations between day-time (07:00-19:00) noise (Lday) and each blood biochemistry marker

| | **Lday, dB(A)** | | | | |
|---|---|---|---|---|---|
| | N | per IQR | Model3 | Model3+$PM_{10}$ | Model3+$NO_2$ |
| **Pooled analysis** | | | | | |
| hsCRP, %diff | 55,930 | 5.1 | **1.1% (0.02% to 2.2%)** | 1.0% (-0.1% to 2.2%) | 0.9% (-0.3% to 2.0%) |
| Total cholesterol, mmol/l | 81,590 | 5.1 | 0.001 (-0.007 to 0.009) | 0.003 (-0.006 to 0.012) | 0.002 (-0.007 to 0.011) |
| HDL cholesterol, mmol/l | 81,590 | 5.1 | **0.007 (0.004 to 0.010)** | **0.006 (0.002 to 0.009)** | **0.004 (0.001 to 0.007)** |
| Triglycerides, mmol/l | 81,799 | 5.1 | **0.008 (0.001 to 0.015)** | 0.003 (-0.005 to 0.010) | 0.003 (-0.005 to 0.010) |
| **LifeLines only** | | | | | |
| Fasting glucose, mmol/l | 62,765 | 4.2 | **0.013 (0.006 to 0.019)** | **0.010 (0.002 to 0.017)** | **0.008 (0.001 to 0.016)** |
| HbA1c, mmol/mol | 50,194 | 4.2 | -0.010 (-0.050 to 0.030) | -0.007 (-0.056 to 0.042) | 0.004 (-0.047 to 0.055) |

Model3: adjusted for age, sex, education, employment, smoking status, smoking pack-years, alcohol consumption, season of blood draw and study (for pooled analysis only).
Bold indicates where significance level<0.05

### 5.1.4. Results from cohort-specific meta-analysis

In general, cohort-specific meta-analyses for hsCRP and blood lipids yielded very similar results as those from the individual-level analyses on DataSHIELD (Table 5.8). However, as expected, DataSHIELD tend to give smaller 95% confidence intervals due to the stronger statistical power associated with the individual level analyses. The meta-analysis graphs of cohort-specific effect estimates for each biochemistry marker are shown in Appendix-5.4.

Table 5.8 Comparison of pooled estimates (Effect estimate, 95%CI, per IQR* higher exposure) based on Model3 from both cohort-specific meta-analysis and pooled analysis on DataSHIELD

| **$PM_{10}$, µg/m³** | **cohort-specific meta-analysis** | **pooled analysis on DataSHIELD** |
|---|---|---|
| hsCRP, %diff | 1.4% (-0.1% to 2.8%) | **1.4% (0.1% to 2.7%)** |
| Total cholesterol | 0.004 (-0.006 to 0.014) | -0.005 (-0.014 to 0.004) |
| HDL cholesterol | **0.004 (0.001 to 0.008)** | **0.003 (0.0002 to 0.007)** |
| Triglycerides | **0.026 (0.018 to 0.035)** | **0.021 (0.013 to 0.028)** |
| | | |
| **$NO_2$, µg/m³** | | |
| hsCRP, %diff | 1.4% (-0.3% to 3.7%) | **1.9% (0.5% to 3.3%)** |
| Total cholesterol | 0.010 (-0.001 to 0.021) | -0.001 (-0.011 to 0.009) |
| HDL cholesterol | **0.010 (0.006 to 0.013)** | **0.008 (0.005 to 0.012)** |
| Triglycerides | **0.027 (0.018 to 0.036)** | **0.021 (0.013 to 0.030)** |
| | | |
| **Lday, dB(A)** | | |
| hsCRP, %diff | 0.7% (-0.3% to 1.7%) | **1.1% (0.02% to 2.2%)** |
| Total cholesterol | 0.002 (-0.005 to 0.009) | 0.001 (-0.007 to 0.009) |
| HDL cholesterol | **0.007 (0.004 to 0.010)** | **0.007 (0.004 to 0.010)** |
| Triglycerides | **0.007 (0.001 to 0.013)** | **0.008 (0.001 to 0.015)** |

Model3: adjusted for age, sex, education, employment, smoking status, smoking pack-years, alcohol consumption, season of blood draw and study (in the pooled analyses only). BOLD indicates where significance level<0.05.
*IQR: for cohort-specific IQR and IQR for pooled dataset see Table5.3.

### 5.1.5. ESCAPE-LUR air pollution metrics: LifeLines only

The analyses were repeated in LifeLines only, using the ESCAPE-LUR modelled air pollutants. In general, ESCAPE-LUR modelled $PM_{10}$ or $NO_2$ effects on HDL cholesterol, triglycerides and fasting glucose levels (Table 5.9) were comparable to those based on the pan-European LUR modelled exposures (presented in Table 5.5).

No significant associations were observed between any of the ESCAPE air pollutants and hsCRP (Table 5.9). Only $PM_{2.5}$ absorbance was associated with total cholesterol level, independent of adjustment for noise. Sensitivity analyses based on Model3 are presented in Appendix-5.5. After further adjusting for BMI, the ESCAPE-LUR modelled $NO_2$ effects on hsCRP and total cholesterol became statistically significant.

A significant association was found between increased ESCAPE-LUR modelled $PM_{2.5}$ and increased HbA1c level.

Table 5.9 Associations between each ESCAPE air pollutant metric (per IQR higher) and each biochemistry marker in LifeLines only

| | Model3 | Model3+Lday |
|---|---|---|
| hsCRP (N=32266), % difference | | |
| **$NO_2$** | 1.0% (-1.4% to 3.4%) | 0.5% (-2.6% to 3.6%) |
| **$PM_{10}$** | -1.1% (-2.8% to 0.6%) | -2.7% (-4.8% to -0.5%) |
| **$PM_{2.5}$** | 0.03% (-0.8% to 0.8%) | -0.2% (-1.1% to 0.7%) |
| **$PM_{coarse}$** | -1.7% (-3.3% to 0%) | -3.1% (-5.0% to -1.1%) |
| **$PM_{2.5absorbance}$** | -0.6% (-2.6% to 1.4%) | -2.2% (-4.8% to 0.4%) |
| total cholesterol (N=53807), mmol/L | | |
| **$NO_2$** | 0.012 (-0.004 to 0.028) | 0.010 (-0.010 to 0.030) |
| **$PM_{10}$** | 0.011 (-0.001 to 0.022) | 0.010 (-0.004 to 0.025) |
| **$PM_{2.5}$** | **0.006 (0.001 to 0.011)** | 0.006 (-0.0002 to 0.012) |
| **$PM_{coarse}$** | 0.005 (-0.006 to 0.017) | 0.002 (-0.012 to 0.016) |
| **$PM_{2.5absorbance}$** | **0.017 (0.004 to 0.030)** | **0.021 (0.003 to 0.039)** |
| HDL cholesterol (N=53807), mmol/L | | |
| **$NO_2$** | **0.019 (0.013 to 0.025)** | **0.015 (0.007 to 0.023)** |
| **$PM_{10}$** | **0.019 (0.015 to 0.024)** | **0.021 (0.015 to 0.026)** |
| **$PM_{2.5}$** | **0.005 (0.003 to 0.007)** | **0.003 (0.0005 to 0.005)** |
| **$PM_{coarse}$** | **0.021 (0.016 to 0.025)** | **0.021 (0.016 to 0.026)** |
| **$PM_{2.5absorbance}$** | **0.021 (0.016 to 0.026)** | **0.023 (0.016 to 0.030)** |
| triglycerides (N=53807), mmol/L | | |
| **$NO_2$** | **0.035 (0.022 to 0.047)** | **0.055 (0.039 to 0.070)** |
| **$PM_{10}$** | 0.007 (-0.002 to 0.016) | **0.013 (0.001 to 0.024)** |
| **$PM_{2.5}$** | 0.001 (-0.003 to 0.006) | -0.001 (-0.006 to 0.003) |
| **$PM_{coarse}$** | 0.008 (-0.001 to 0.017) | **0.019 (0.008 to 0.029)** |
| **$PM_{2.5absorbance}$** | **0.011 (0.001 to 0.022)** | **0.020 (0.006 to 0.033)** |
| fasting blood glucose (N=52453), mmol/L | | |
| **$NO_2$** | **0.041 (0.028 to 0.054)** | **0.034 (0.017 to 0.051)** |
| **$PM_{10}$** | **0.023 (0.013 to 0.032)** | **0.014 (0.001 to 0.026)** |
| **$PM_{2.5}$** | **0.006 (0.001 to 0.010)** | 0.001(-0.004 to 0.006) |
| **$PM_{coarse}$** | **0.024 (0.015 to 0.033)** | **0.016 (0.005 to 0.028)** |
| **$PM_{2.5absorbance}$** | **0.024 (0.013 to 0.035)** | 0.011 (-0.004 to 0.026) |
| HbA1c (N=43537), mmol/mol | | |
| **$NO_2$** | -0.060 (-0.142 to 0.021) | -0.091 (-0.194 to 0.012) |
| **$PM_{10}$** | 0.014 (-0.044 to 0.072) | 0.029 (-0.048 to 0.105) |
| **$PM_{2.5}$** | **0.028 (0.001 to 0.055)** | **0.037 (0.006 to 0.068)** |
| **$PM_{coarse}$** | -0.027 (-0.084 to 0.031) | -0.037 (-0.107 to 0.034) |
| **$PM_{2.5absorbance}$** | 0.006 (-0.062 to 0.074) | 0.018 (-0.075 to 0.111) |

Effect estimates (95%CI) were calculated for each IQR increase of $NO_2$ (7.42 μg/$m^3$), $PM_{10}$ (0.95 μg/$m^3$), $PM_{2.5}$ (0.24μg/$m^3$) and PMcoarse (0.63μg/$m^3$) and PM2.5absorbance (0.22 $10^{-5}$/m).
Model 3: adjusted for age, sex, education, employment status, smoking status, smoking pack-years, alcohol consumption. Lday: day-time noise (07:00-19:00). BOLD indicates where significance level<0.05.

### 5.1.6. Summary of findings

A brief summary of main findings for the effect of air pollution and noise on blood biochemistry markers are presented in Table 5.10.

Table 5.10 Summary of the main findings for each biochemistry outcome

| | hsCRP | Total Cholesterol | HDL Cholesterol | Triglycerides | Fasting Glucose | HbA1c |
|---|---|---|---|---|---|---|
| $NO_2$ (pan-European LUR) | ++ | - | ++ | ++ | ++ | - |
| $NO_2$ (ESCAPE LUR) [LifeLines only] | - | - | ++ | ++ | ++ | - |
| $PM_{10}$ (pan-European LUR) | + | - | + | ++ | ++ | - |
| $PM_{10}$ (ESCAPE LUR) [LifeLines only] | - | - | ++ | ++ | ++ | - |
| Lday | + | - | ++ | + | ++ | - |

++: significant positive associations in main Model 3, also robust to noise (or air pollution) adjustment;
+: significant positive associations in main Model 3, but not robust to noise (or air pollution) adjustment;
-: no significant associations observed.
Model3: adjusted for age, sex, education, employment, smoking status, smoking pack-years, alcohol consumption, season of blood draw and study (for pooled analysis only).

#### *Joint effects of air pollution and road traffic noise*

In general, association between $NO_2$ and each of hsCRP, HDL cholesterol, triglycerides or fasting glucose was not confounded by further adjustment for road traffic noise (Table 5.10). In contrast, the significant association observed between $PM_{10}$ and hsCRP or HDL cholesterol was confounded by road traffic noise.

The significant positive association between daytime road traffic noise and hsCRP was lost after adjustment for $PM_{10}$, suggesting confounding by air pollution. This was also seen for triglycerides.

## 5.2. Discussion

### 5.2.1. Ambient air pollution effects

#### *hsCRP*

Inflammatory response to ambient air pollution has been perceived as one of the key mechanisms that explains the observed air pollution effects on a range of health outcomes, including cardiovascular diseases, diabetes, cancer and Alzheimer's disease(Lanki et al., 2015). hsCRP is one of the many possible biomarkers for systemic inflammation and is usually used as a clinical indicator for increased CVD risks(Li et al., 2012).

In this PhD project, pooling individual-level data from 51,459 middle-aged participants from two large European cohorts, I found that traffic-related air pollution, as indicated by $NO_2$ but not PM, was significantly associated with higher levels of hsCRP. This analysis is one of the largest conducted to date with regards to its sample size and provides further evidence of air pollution effects on systemic inflammation among general populations.

It was reported that a higher BMI was associated with a higher concentration of hsCRP(Visser et al., 1999). In this PhD study, BMI was not on the causal pathway between air pollution and hsCRP as adjustment of BMI did not attenuate the observed association in the main model.

Short-term air pollution effects on systemic inflammation have been well documented(Li et al., 2012), however only a few prior studies investigated long-term (e.g. annual or longer) air pollution effects on hsCRP level in general adult populations, with mixed results reported.

A longitudinal analysis of the Heinz Nixdorf Recall Study in Germany reported that a 2.4µg/$m^3$ increase of $PM_{2.5}$ was associated with a 5.4% (95%CI: 0.6% to 10.5%) change of hsCRP in both sexes (i.e. a combined population), independent of short-term air pollution exposure(Viehmann et al., 2015). ESCAPE-LUR modelled $PM_{2.5}$ estimates were only available for LifeLines participants in this PhD project. In contrast to the findings of the Heinz Nixdorf Recall Study, I did not find an association between $PM_{2.5}$ and hsCRP. Although the sample size is much larger in LifeLines (N=32,266), the ESCAPE

$PM_{2.5}$ exposure contrast in LifeLines was smaller by a factor of 10 (IQR: 0.24 μg/$m^3$) compared to that reported in the Heinz Nixdorf Recall Study, limiting my ability to detect any association in LifeLines if present.

The ESCAPE analysis of 22,561 adults from six European cohorts did not observe significant associations between any PM metrics or $NO_2$ and hsCRP level(Lanki et al., 2015). However, they found that increased traffic intensity on the nearest road was significantly associated with higher hsCRP levels. In their study, each 10 μg/$m^3$ increase in $NO_2$ was associated with 2.1% (95%CI: -0.8% to 5.1%) increase in hsCRP. This is compatible with my findings from 32,266 LifeLines participants in a similar main model also adjusted for BMI, giving a significant association with ESCAPE-LUR modelled $NO_2$ (5.2%, 95%CI: 2.0% to 8.3%, per 10 μg/$m^3$). One study in England reported no associations between $PM_{10}$, $NO_2$ and hsCRP in each of the three cross-sectional surveys in 1994, 1998 and 2003 respectively(Forbes et al., 2009b) while another study in Sweden reported that 30-year average exposure of $NO_2$ was not associated with hsCRP(Panasevich et al., 2009).

As with the ESCAPE analysis, I found that the effects of $NO_2$ on hsCRP level was more consistent, compared to the PM effects, in both pooled analyses using the pan-European LUR modelled estimates and in the LifeLines only analysis using the ESCAPE-LUR modelled estimates. PM could come from sources other than traffic(Kelly & Fussell, 2012b) and is generally regarded as an indicator for regional air pollution as it can travel over long distances(Kunzli, 2014). In contrast, $NO_2$ is more related to local traffic and its concentration reduces substantially within 50-500 metre of a busy street (reduced to only 10-30% of the street concentration level)(Kunzli, 2014). This has resulted in a higher spatial contrast (indicated by IQR) for $NO_2$ than for PM, which then increases the potential to detect effects for $NO_2$.

It is plausible that participants who had been constantly exposed to a higher level of $NO_2$ (i.e. those living near a busy road) may have an increasing level of inflammatory response. Correspondingly, I also found that, only the ESCAPE-LUR modelled $PM_{2.5}$ absorbance, also known as soot (mostly

associated with vehicle exhaust), was also significantly associated with hsCRP levels in the LifeLines participants, after further adjustment for BMI. This further suggests that near-road traffic-related air pollution may have more adverse effects on systemic inflammation, as measured by hsCRP level. Yet, whether increasing $NO_2$ is causally related to increased hsCRP level is unclear. There are still debates as to whether $NO_2$ on its own acts as an air pollutant which has direct effects on health, or merely a marker of traffic-related air pollution, or indeed a marker of PM metrics that were not modelled for this PhD project such as ultrafine particles or one of the constituents of $PM_{10}$.

It should be noted though, since both gaseous and particulate air pollution could induce an inflammatory response, it may be difficult to disentangle the effect of each on hsCRP level. The role of road traffic noise adds even more complexity to the associations between air pollution and hsCRP. In the pooled analyses using the pan-European LUR modelled $PM_{10}$, the significant association between $PM_{10}$ and hsCRP was attenuated to null by adjusting for road traffic day-time noise, suggesting a substantial confounding effect. An analysis of the Heinz Nixdorf Recall Study suggested that effect of $PM_{2.5}$ on hsCRP was independent of adjustment for traffic noise(Hennig et al., 2014). More studies are needed to better clarify the role of road traffic noise on this possible PM-hsCRP link.

Contrary to the findings of the Heinz Nixdorf Recall Study(Viehmann et al., 2015), I found stronger associations between air pollution and hsCRP levels in women, with sex showing a statistically significant effect modification. The ESCAPE study however did not suggest that the associations were modified by sex(Lanki et al., 2015). The biological explanations for modification of effect by sex are unclear, but my analyses benefited from a sufficiently large sample size to investigate such an interaction.

I found a stronger association with $NO_2$ among those aged less than 60 years, but this was not seen for $PM_{10}$. In fact, I found a stronger association with $PM_{10}$ among those aged 60 years and above. The Heinz Nixdorf Recall Study reported stronger associations among older participants (≥60 years) with $PM_{2.5}$ exposures, although the interaction effects were not statistically significant(Viehmann et al.,

2015). Noteworthy, in my study, stronger associations between air pollution and hsCRP were found among those who had ever had diabetes, suggesting diabetes patients are more vulnerable to the adverse air pollution effects.

What this PhD project suggests is that near-road traffic-related air pollution indicated by $NO_2$ is related to higher levels of hsCRP. More studies are needed in this area, specifically to further investigate the roles of $NO_2$ and PM at the individual level on hsCRP and other biomarkers of systemic inflammation in different populations. This PhD is one of very few to study the effects of traffic noise on the link between air pollution and hsCRP, suggesting a likely confounding effect on the PM but not $NO_2$ effects. This needs to be confirmed in more studies.

### *Blood lipids*

Growing evidence has pointed to a link between air pollution exposure and atherosclerosis(Kunzli et al., 2010), for which dyslipidaemia is one of the major established risk factors. But the epidemiological links between air pollution and blood lipid measures are rarely studied. Ambient air pollution, especially particulate air pollution, has the potential to alter blood lipid metabolism via activation of the autonomic nervous system (ANS) imbalance, elevated systemic inflammation and oxidative stress, impaired endothelial function and even a direct transfer of fine particle (i.e. $PM_{2.5}$) from the lung to blood circulation(Brook et al., 2010). These possible pathways may ultimately lead to manifest CVD.

This PhD project, in line with only two recently published studies in 2015-2016(Shanley et al., 2016; Sorensen et al., 2015), suggest associations between long-term ambient air pollution exposure and raised levels of blood lipids in general populations. I found that, higher levels of both $PM_{10}$ and $NO_2$, modelled by either the pan-European LUR or ESCAPE-LUR model, were consistently associated with higher levels of triglycerides. These associations were independent from a range of covariates including noise and remained in sensitivity analyses. Similar findings were also observed for HDL cholesterol.

However, I did not observe significant associations between any pan-European LUR or ESCAPE-LUR modelled air pollutant and total cholesterol levels in the analyses except for the ESCAPE-LUR modelled $PM_{2.5\ absorbance}$ in the model adjusted for noise. For the pooled analyses using the Pan-European LUR modelled $PM_{10}$ or $NO_2$, the association between each air pollutant and total cholesterol level was null, even after adjustment for BMI; whilst for the analyses using the ESCAPE-LUR modelled $PM_{10}$ or $NO_2$ in LifeLines, the null associations based on Model3 became statistically significant after adjustment for BMI.

A recent study consisting of nearly 40,000 adults aged 50-64 years in Copenhagen, reported a statistically significant association between no-fasting total cholesterol and both ESCAPE LUR-modelled $PM_{2.5}$ (0.101 mmol/L, 95%CI: 0.028 to 0.173, per 5 μg/$m^3$) and dispersion modelled $NO_2$ (0.026mmol/L, 95%CI: 0.008 to 0.045, per 10 μg/$m^3$), independent of adjustment for BMI(Sorensen et al., 2015). Other lipids measures were not available. My analyses of 53,807 relatively younger LifeLines participants (mean age: 47.6 years) showed comparable results for total cholesterol in the model also adjusted for BMI, using ESCAPE LUR-modelled $PM_{2.5}$ (0.146 mmol/L, 95%CI: 0.021 to 0.25, per 5 μg/$m^3$) and $NO_2$ (0.023 mmol/L, 95%CI: 0.001 to 0.044, per 10 μg/$m^3$).

Another interesting finding is that, in LifeLines, the ESCAPE LUR-modelled $PM_{2.5}$ was not associated with triglyceride levels in the main model adjusted for noise, but $PM_{10}$ or PMcoarse was. Since this is the first report on a null association between $PM_{2.5}$ and triglycerides, whether PM-induced effects on triglyceride levels were mainly related to larger particles rather than fine particle would need to be further investigated in more studies. A nationwide USA study of 11,623 participants (median age: 41 years) reported significant positive associations between nearest monitor-based annual $PM_{10}$ exposure and both triglycerides and total cholesterol levels(Shanley et al., 2016). Association with HDL cholesterol was positive but not significant. $PM_{2.5}$ data were not available for analyses in this USA study.

I found both higher pan-European and ESCAPE LUR modelled $PM_{10}$ and $NO_2$ were significantly associated with HDL cholesterol. Although still generally regarded as a 'good' cholesterol which can reduce CVD risk, HDL is in fact functionally diverse and the simple link between elevated HDL levels and reduced CVD risks was recently questioned by some clinical trials and genetic studies(Rader & Tall, 2012; Heinecke, 2012). In both review papers of Rader(Rader & Tall, 2012) and Heinecke(Heinecke, 2012), it was suggested that HDL's ability to protect against CVD may not simply rely on its quantity but more importantly on its functionality.

A recent experimental study first reported that diesel exhaust could reduce the anti-inflammatory and anti-oxidant effects of HDL in mice, two important pathways for HDL cholesterol to act against CVD(Yin et al., 2013). The authors came to an interesting conclusion that air pollution exposure may potentially change HDL from a 'good' cholesterol to just a 'normal' or even a 'bad' one by altering functionality. Although this is an experimental study in mice, the finding has an important implication for what I report here. In my analysis, air pollution exposure was associated with higher levels of HDL cholesterol. However, if anti-CVD properties of these raised HDL cholesterol are impaired by exposure to air pollution, a link between air pollution and CVD may still be plausible via an increasing level of 'dysfunctional' HDL cholesterol. Future human studies on how air pollution affects anti-CVD properties of HDL cholesterol would provide novel insights into this possible pathway.

It should be noted that this PhD analysis did not give a clear suggestion as to whether BMI may be on the causal pathway between air pollution and raised total cholesterol or HDL cholesterol - the significant associations between pan-European LUR air pollution metrics and total or HDL cholesterol became null after adjustment for BMI in main models, but not in models using alternative ESCAPE air pollution metrics in LifeLines only.

Several smaller studies conducted in asthmatics(Yeatts et al., 2007), traffic police personnel(Tomao et al., 2002), adolescences(Poursafa et al., 2014b) and elderly participants(Chuang et al., 2011), provide varied results. For example, one reported positive effects of air pollution on triglycerides only(Yeatts

et al., 2007), and one reported such effects on both HDL and triglycerides(Tomao et al., 2002) and in contrast to these, the other reported positive effects of air pollution on total cholesterol only but not on levels of triglycerides or HDL cholesterol(Chuang et al., 2011). These results however cannot be generalized to other populations due to their specific study designs in the chosen populations.

Taken together, from this PhD analysis and previous reports, it appears air pollution (either PM or $NO_2$) has a role in raising blood lipids levels. Dyslipidaemia, usually associated with lifestyle factors such as smoking, unhealthy diet and infrequent exercise, is one of the most common risk factors for CVD(Kannel, 1990). If the link between air pollution and dyslipidaemia is casual related, it will not only help identify yet another modifiable environmental risk factor, but also help explain, in part, the links between air pollution and CVD outcomes. Given the current literature in this field is very limited, there is a need for more epidemiological studies investigating this potentially important link.

### *Fasting glucose and HbA1c*

Accumulating evidence shows possible associations between long-term exposure to air pollution and diabetes in several European populations(Eze et al., 2015). Fasting blood glucose and HbA1c level (average of blood glucose levels in the previous 30-120 days prior to blood draw) are the two important indicators used to identify impaired glucose tolerance, which is a risk factor itself for CVD. However, the direct link between long-term air pollution exposure and blood glucose has not been studied extensively. Biological mechanisms which may be involved in linking air pollution and blood glucose levels include air pollution exerting a direct effect on insulin resistance, and/or inducing oxidative stress and adipose tissue inflammation(Brook et al., 2013).

Among 52,453 LifeLines participants included in the main model in this PhD study, address-level $PM_{10}$ and $NO_2$ assigned from either the pan-European or ESCAPE LUR model, were significantly associated with increased fasting glucose; associations which were robust to adjustment for traffic noise. Additionally, ESCAPE LUR-modelled $PM_{2.5}$, $PM_{2.5absorbance}$ and PMcoarse were all significantly associated with increased fasting glucose, although only the association with PMcoarse was independent of

traffic noise. The results from this PhD analysis are consistent with only a few prior studies which examined either short-term(Chen et al., 2016; Sade et al., 2015; Chuang, Yan & Cheng, 2010) or long-term(Tamayo et al., 2014; Teichert et al., 2013; Chuang et al., 2011) air pollution effects on blood glucose.

In contrast to the findings for blood glucose, I did not observe an association between any air pollutant and HbA1c except for the ESCAPE-LUR modelled $PM_{2.5}$. This however may be a chance finding due to the multiple testing, and more studies are needed to replicate this association between $PM_{2.5}$ and HbA1c. Only two earlier studies investigated long-term air pollution effects on HbA1c levels. Annual average $PM_{10}$ from fixed monitoring stations was significantly associated with increased HbA1c among elderly participants in Taiwan(Chuang et al., 2011), while in Germany higher regional $PM_{10}$ levels were significantly associated with higher HbA1c levels among diabetics(Tamayo et al., 2014). Unlike these two previous studies, this PhD study was conducted in a general population with modelled air pollution estimates assigned at address-level, which may partly explain the inconsistent findings. Another German study of 363 elderly women reported that long-term exposure to $NO_2$ may contribute to impaired glucose metabolism, although significance was lost after adjustment for multiple testing(Teichert et al., 2013).

A recent review in 2015 suggested that no firm conclusions could be drawn on the associations between air pollution and diabetes-related traits at this stage due to the small number of studies assessing this association(Thiering & Heinrich, 2015). This PhD analysis, one of the largest to date, suggests a positive significant association between long-term air pollution and fasting glucose level but not with HbA1c. This finding may provide some mechanistic insights into the emerging link between air pollution and diabetes.

### 5.2.2. Road traffic noise effects

#### *hsCRP*

To my best knowledge, there are no other studies exploring the association between long-term road traffic noise exposure and hsCRP level. Chronic exposure to excess environmental noise is generally regarded as a psychological stressor which may lead to increased level of stress hormones and ultimately manifest CVD(Babisch, 2014). A recent study suggested the link between long-term psychological stress and CVD may operate via a chronic low-grade systemic inflammation(Rohleder, 2014). This provides a biological rationale to explore the direct link between road traffic noise and hsCRP level. Results from this PhD analysis provided some initial evidence to support this hypothesis. Every 5.1 dB(A) increase in day-time road traffic noise was associated with a small increase of 1.1% in hsCRP (95%CI: 0.02 to 2.2), however this significant association was lost after adjustment for $PM_{10}$ or $NO_2$ exposure. This result suggested that long-term exposure to road traffic noise may have a positive effect on hsCRP in its own, but this effect may not be seen after air pollution was accounted for. Neither the ESCAPE nor Heinz Nixdorf Recall Study published the effect estimates of traffic noise on hsCRP in their analyses(Lanki et al., 2015; Viehmann et al., 2015). My findings need to be replicated in more studies before a possible conclusion could be drawn.

#### *Blood lipids*

With respects to noise exposure, a statistically significant association with increased HDL and a moderate non-significant increase in total cholesterol and triglycerides were observed after adjustment for air pollution, in line with only two previous population-based studies(Sorensen et al., 2015; Babisch et al., 1993b) and one occupational study(Melamed et al., 1997). The underlying mechanisms, in particular for increased HDL cholesterol, are not very clear. An earlier review found that, among the reviewed studies, none had documented a decrease in HDL in response to stress(Niaura, Stoney & Herbert, 1992).

Only one recent study conducted in Copenhagen assessed the dose-response relationship between continuous noise exposure estimates and blood lipid levels(Sorensen et al., 2015). They found that,

for every 5 dB(A) increase of $L_{den}$, total cholesterol level increased by 0.008 mmol/L (95%CI: -0.001 to 0.017). This is similar to what I reported here in the pooled analysis. Whilst the Copenhagen study assessed only total cholesterol level in relation to road traffic noise exposure, two earlier studies, conducted in the 1990s, assessed not only total cholesterol but also triglycerides and HDL cholesterol levels. Both studies reported a positive significant trend (p-value<0.05) between higher noise exposures and higher total cholesterol and triglycerides level(Melamed et al., 1997; Babisch et al., 1993b); the study of blue-collar workers additionally reported that noise annoyance was positively correlated with increased level of HDL in both men and women(Melamed et al., 1997).

One possible explanation may be based on the general stress model. In the Whitehall study of 199 middle-aged men and women in the UK, it was found that after 3-year of performing moderately stressful behavioural tasks, both total cholesterol and HDL cholesterol had raised for all participants after adjusting for lifestyle factors, with individuals who had a larger stress response having a greater rise(Steptoe & Brydon, 2005). It has been suggested that stress may increase release of lipids into the circulation or reduce plasma volume which then leads to a more concentrated level of circulating proteins. In fact, in some but not all studies, after adjusting for plasma volume, the association between stress and increased serum lipids were no longer significant(Steptoe & Brydon, 2005). It is not possible to obtain plasma volume data for this PhD project, but this is possibly a new angle to explore when assessing the link between traffic noise and blood lipids further.

Cortisol, a stress hormone, may also play a crucial role on the link between noise and lipids levels. Following stressful episodes, cortisol is released to restore homeostasis mainly via metabolic pathways: increasing the supply of energy in the forms of glucose and fatty acids(Brindley et al., 1993). Catecholamines is an important stimulator in this increasing supply of energy by breaking down triacylglycerols(Brindley et al., 1993). All these complex responses then lead to an increased hepatic production of the very low-density lipoprotein, which will be converted to low-density lipoprotein as a major carrier of cholesterol in the circulation(Brindley et al., 1993).

Taken together, it is biologically plausible that chronic exposure to noise, particularly if at a consistently high level, may lead to adverse changes in blood lipids (dyslipidaemia) long term. Since only two studies, including the present one, investigated this topic in an epidemiological context, more studies are needed, in particular studies on the underlying metabolic or molecular mechanisms to provide insights into this potential pathway.

### *Fasting glucose and HbA1c*

In this PhD analysis, I also found that road traffic noise, independent of air pollution effects, was significantly associated with increased fasting glucose levels. However, I did not see a clear corresponding association between noise exposure and HbA1c levels. Possible mechanisms underlying a link between noise exposure and glucose dysregulation have been proposed, such as stress hormone (e.g. cortisol) secretion leading to increased supply of glucose(Brindley et al., 1993), as well as noise-induced sleep disturbance(Tasali, Leproult & Spiegel, 2009), but epidemiological evidence for the effects of traffic noise on blood glucose levels is extremely limited.

There is some emerging evidence linking noise exposures to diabetes morbidity and mortality. In a study of 57,053 participants in Denmark, a 10-dB higher road traffic noise was associated with a 8% (95%CI: 2% to 14%) increase of risk for incident diabetes after nearly 10-years of follow-up(Sorensen et al., 2013). In the city of Madrid, short-term night-time noise exposure was significantly associated with mortality from diabetes(Tobias et al., 2015). Findings from these studies, however, would have been further supported if the association between noise and glucose levels was directly analysed in the same study.

Only one study of 2,348 men conducted in the early 1990s investigated the link between road traffic noise exposure and blood glucose levels, in which it was suggested that higher day-time noise levels were associated with increased glucose levels (P-value for trend < 0.05)(Babisch et al., 1993b). Another study in Stockholm found no associations between aircraft noise exposure and impaired fasting glucose(Eriksson et al., 2014). The findings from this PhD analysis, which comprised a large population

study (N=62,765) with individual noise estimates assigned to each participant at address level, provide one of the first suggestive evidence of a dose-response relationship between road traffic noise and blood glucose level.

### 5.2.3. Chapter summary

A suggestive conclusion revealed from this analysis on CVD biochemical parameters is that long-term exposure to air pollution or traffic noise is significantly associated with adverse blood biochemistry including systemic inflammation, blood lipids (except for HDL) and blood glucose. These associations were observed in a large sample pooled from two general population cohorts, and were independent of a range of covariates. These findings are in support of a link between air pollution, road traffic noise exposures and CVD outcomes, the results for which I presented in the next Chapter.

# Chapter 6 Findings and discussions: incident cardiovascular disease

## 6.1. The findings

### 6.1.1. Descriptive statistics

Individual-level data from 625,327 participants, pooled from EPIC-Oxford, HUNT2 and UK Biobank, were available for this analysis (Table 6.1). The mean age of the pooled population was 54.8 years; 56% were women; 63% were overweight or obese (BMI≥25 kg/m$^2$) and 13% were current smokers. More than half of the participants in both UK cohorts had received education at post-secondary or degree level, compared to 20% in HUNT2. Participants of both UK cohorts also had higher levels of alcohol consumption per week than did those of the HUNT2.

Table 6.1 Baseline characteristics of EPIC-Oxford, HUNT2 and UK Biobank

| | Pooled data | EPIC-Oxford | HUNT2 | UK Biobank |
|---|---|---|---|---|
| Total N | 625,327 | 57,446 | 65,232 | 502,649 |
| Baseline years | 1993-2010 | 1993-1999 | 1995-1997 | 2006-2010 |
| Age, years, mean(SD) | 54.8 (10.7) | 45.5 (14.2) | 50.1 (17.2) | 56.5 (8.1) |
| Sex, women, [n (%)] | 352,359 (56%) | 44,227 (77%) | 34,665 (53%) | 273,467 (54.4%) |
| BMI, kg/m$^2$, mean(SD) | 27.0 (4.8) | 24.0 (3.9) | 26.4 (4.1) | 27.4 (4.8) |
| BMI categories, [n (%)] | | | | |
| **<25 kg/m$^2$** | 230,142 (37%) | 39,297 (68%) | 25,762 (40%) | 165,083 (33%) |
| **25-30 kg/m$^2$** | 254,232 (41%) | 14,092 (25%) | 27,972 (43%) | 212,168 (42%) |
| **≥30 kg/m$^2$** | 137,053 (22%) | 4,039 (7%) | 10,722 (17%) | 122,292 (25%) |
| Smoking status, [n (%)] | | | | |
| **Never-smoker** | 335,409 (54%) | 34,044 (60%) | 27,759 (44%) | 273,606 (55%) |
| **Ex-smoker** | 207,287 (33%) | 16,675 (29%) | 17,510 (27%) | 173,102 (35%) |
| **Current-smoker** | 78,015 (13%) | 6,493 (11%) | 18,533 (29%) | 52,989 (10%) |
| Current working status, [n (%)] | | | | |
| **Not in paid employment** | 252,685 (41%) | 19,000 (34%) | 24,036 (37%) | 209,649 (42%) |
| **In paid employment** | 365,270 (59%) | 37,658 (66%) | 40,374 (63%) | 287,238 (58%) |
| Education level, [n (%)] | | | | |
| **primary education or less** | 107,977 (18%) | 0 | 22,685 (37%) | 85,292 (17%) |
| **Secondary education** | 158,495 (26%) | 21,174 (47%) | 26,770 (43%) | 110,551 (23%) |

| | | | | |
|---|---|---|---|---|
| **Post-secondary school or above** | 333,012 (56%) | 24,039 (53%) | 12,309 (20%) | 296,664 (60%) |
| Alcohol consumption, gram per week, mean(SD) | 150 (164) | 65 (84) | 22 (36) | 175 (170) |
| Ever-had hypertension, [n (%)] | 149,297 (24%) | 6,265 (12%) | 7,245 (11%) | 135,787 (27%) |
| Ever-had diabetes, [n (%)] | 29,219 (5%) | 783 (2%) | 2,028 (3%) | 26,408 (5%) |

Table 6.2 Incident cardiovascular outcomes in each cohort and in the pooled data

| | **Pooled data** | **EPIC-Oxford** | **HUNT2** | **UK Biobank** |
|---|---|---|---|---|
| **linkages with medical records** | 430,728 | 40,384 | 53,121 | 337,223 |
| **Percentage of the original cohort** | 69% | 70% | 81% | 67% |
| **Averaged (SD) person-years follow-ups** | 4.4 (5.9) | 13.2 (3.1) | 14.6 (5.9) | 1.4 (0.8) |
| **Total person-years follow-ups** | 1,719,692 | 534,192 | 776,110 | 409,390 |
| **Incident CVD ,n (ICD10, I00-I99)** | 30,428 | 6,421 | 20,294 | 3,713 |
| **Incident ischaemic heart disease, n (ICD10, I20-I25)** | 5,259 | 807 | 3,712 | 740 |
| **Incident cerebrovascular disease, n (ICD10, I60-I69)** | 2,871 | 423 | 2,208 | 240 |

After excluding prevalent CVD cases at recruitment, linkages to medical records were possible for 430,728 (69%) participants to permit analysis of incident CVD since baseline recruitment in each cohort. The average person-years of follow-up for incident cardiovascular disease was 13.2, 14.6 and 1.4 for EPIC-Oxford, HUNT2 and UK Biobank respectively (Table 6.2). In total, 1,719,692 person-years were recorded. Over the follow-up period, 30,428 incident CVD (ICD10 I00-I99) cases from all three cohorts were registered, among these, 5,259 cases of incident ischaemic heart disease (ICD10 I20-I25) and 2,871 of incident cerebrovascular disease (ICD10 I60-I69).

The length of follow-up period was similar in both EPIC-Oxford and HUNT2, however the incidence rate for CVD was higher in HUNT2. Participants in HUNT2 were recruited during 1995-1997, and at recruitment nearly 60% of them were over the age of 45 years (mean age: 50 years). After nearly 20-

years follow-up to 31-March-2015, most of these participants were in early old age and above (mean age: 62 years). According to Statistics Norway (http://www.ssb.no/ accessed March 2016), the prevalence of CVD in the whole Trøndelag region, the northern part of which was the base for HUNT2, was 51% for those aged 67 years above and 22% for those aged 45-66 years in year 2012. In addition, a higher smoking prevalence and lower education attainment were found for HUNT2 participants than for EPIC-Oxford participants, which may have partly contributed to the high incidence rate of CVD. EPIC-Oxford had a slightly younger age structure at baseline recruitment, compared to HUNT2, but as described earlier in Chapter 2, most participants of EPIC-Oxford were health-conscious and were more likely to maintain a healthy lifestyle throughout the follow-up period, which might contribute to the lower incident rate in this cohort compared to HUNT2.

The pooled median pan-European $PM_{10}$ and $NO_2$ was 21.4 and 27.3 µg/$m^3$, with an inter-quartile range (IQR) of 4.1 and 13.2 µg/$m^3$ respectively (Figure 6.1). $PM_{10}$ estimates for HUNT2 participants were lower than those in both EPIC-Oxford and UK Biobank. Pooled median (IQR) of Lday and Lnight was 54 (3.9) and 45 (3.8) dB(A) respectively (Figure 6.2). In each cohort, for participants living on minor roads that were not captured in the national level traffic datasets, a fixed low-level baseline traffic flow was assigned during modelling, which resulted in a high density of participants in the lower end of noise exposure ranges as seen from Figure 6.2.

The Spearman correlation between pan-European $PM_{10}$ and Lday was moderate (r= 0.35); a similar correlation was seen for $NO_2$ (r=0.46). Correlations were almost unity between Lday and Lnight road traffic noise (r=0.99). For all analyses investigating noise on a continuous scale, Lnight effects were similar to those of the Lday effects, and therefore are not reported here.

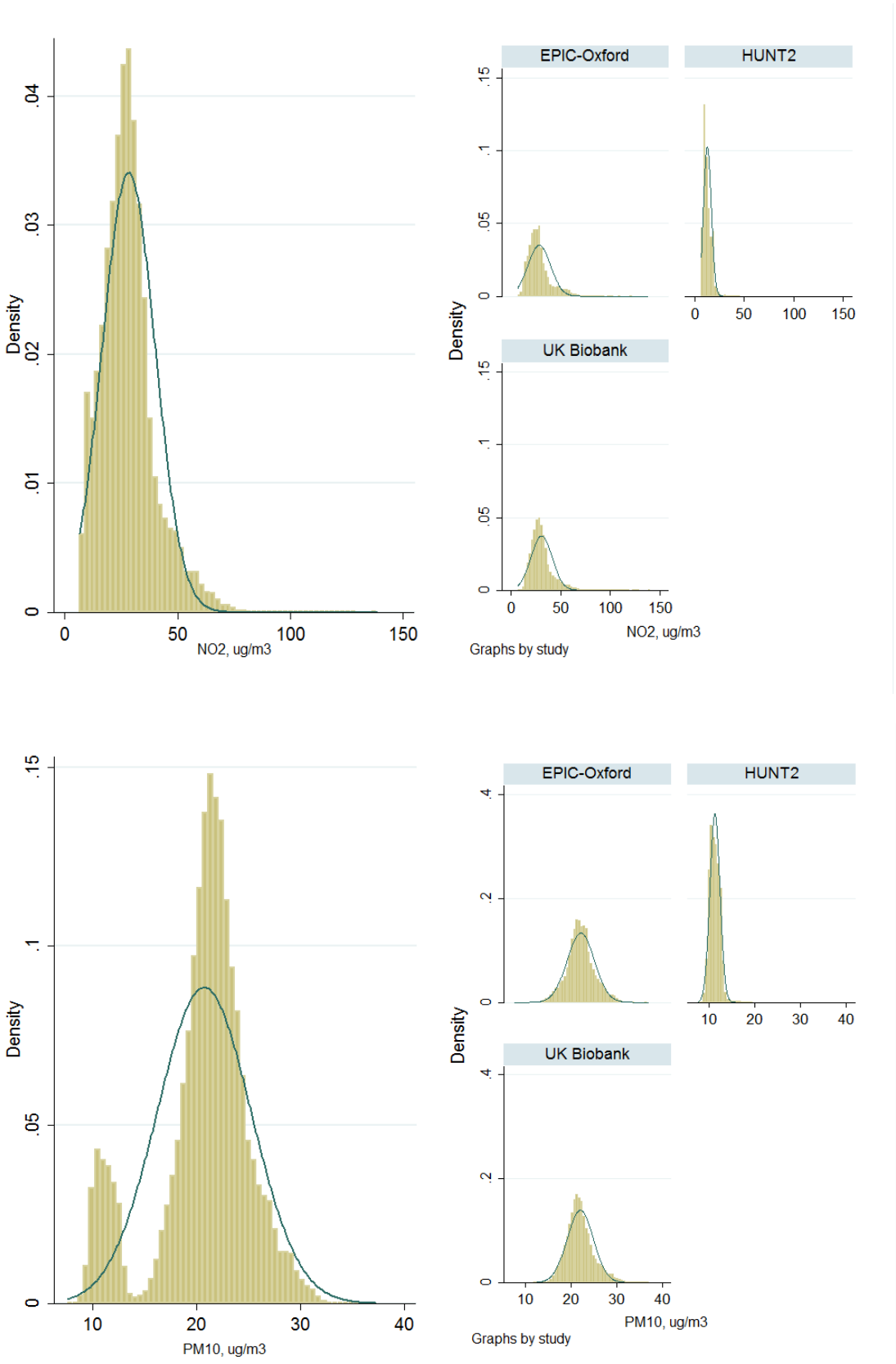



Figure 6.1 distributions of pan-European $NO_2$ (above) and $PM_{10}$ (below) and in the pooled data and in each cohort.

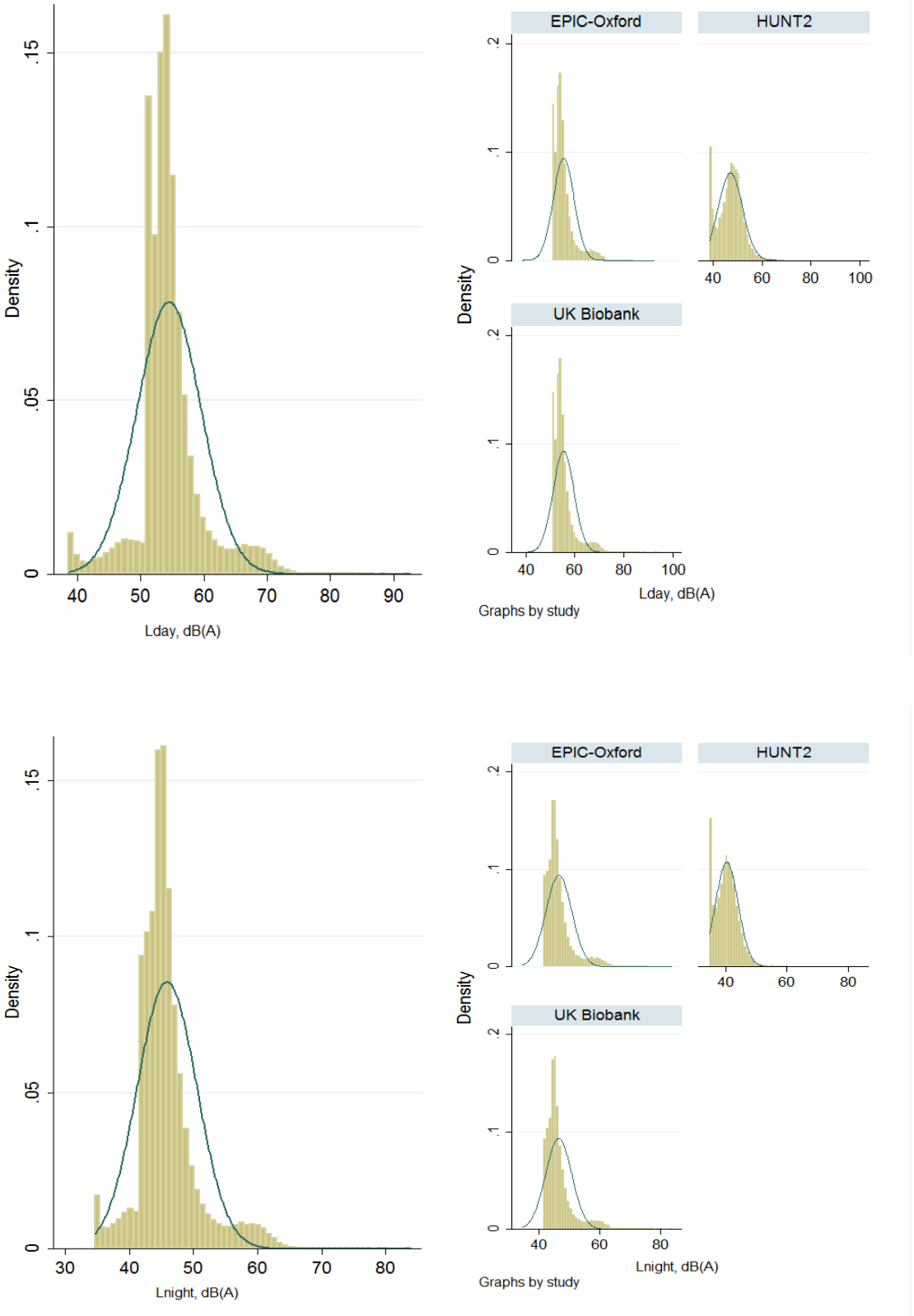



Figure 6.2 distributions of Lday (07:00-19:00) (above) and Lnight (23:00-07:00) (below) road traffic noise in the pooled data and in each cohort.

### 6.1.2. Ambient air pollution and incident CVD outcomes

#### *Pooled analysis 1: virtual pooling of data in DataSHIELD*

Data from all three cohorts were virtually pooled on DataSHIELD and analysed using piecewise regression methods. Overall, the main model (Model3), indicated no statistically significant associations between any of the three CVD outcomes and the pan-European LUR modelled $PM_{10}$ or $NO_2$ (Table 6.3). Each IQR (4.1 µg/m$^3$) increase of $PM_{10}$, incident cerebrovascular disease was associated with a non-significant increase in risk of 8% (RR: 1.08, 95%CI: 0.96 to 1.22). Each IQR (13.2 µg/m$^3$) increase of $NO_2$, was associated with a non-significant increased risk of incident cerebrovascular disease of 5% (RR: 1.05, 95%CI: 0.95 to 1.17). Further adjustments for Lday did not change these associations except for $PM_{10}$ and incident CVD where after adding Lday to the model3, a significant association was seen with RR: 1.04 (95%CI: 1.00 to 1.07*, P-value*: 0.028).

Further adjusting for BMI or ever-had diabetes did not alter these associations in Model3 (data not shown).

Table 6.3 Associations between pan-European LUR modelled $PM_{10}$, $NO_2$ and incident cardiovascular disease: a virtually pooled analysis from EPIC-Oxford, HUNT2 and UK Biobank

| $PM_{10}$, per each IQR (4.1 µg/m$^3$) | | | $NO_2$, per each IQR (13.2 µg/m$^3$) | | |
|---|---|---|---|---|---|
| Incident cerebrovascular disease | RR | 95%CI | Incident cerebrovascular disease | RR | 95%CI |
| Model 1 | 1.03 | 0.93 to 1.13 | Model 1 | 0.98 | 0.90 to 1.07 |
| Model 2 | 1.10 | 0.99 to 1.21 | Model 2 | 1.07 | 0.98 to 1.16 |
| Model 3 | 1.08 | 0.96 to 1.22 | Model 3 | 1.05 | 0.95 to 1.17 |
| Model 3+Lday | 1.07 | 0.94 to 1.21 | Model 3+Lday | 1.04 | 0.93 to 1.16 |
| | | | | | |
| Incident ischaemic heart disease | RR | 95%CI | Incident ischaemic heart disease | RR | 95%CI |
| Model 1 | 0.93 | 0.87 to 0.99 | Model 1 | 0.88 | 0.83 to 0.93 |
| Model 2 | 1.00 | 0.93 to 1.07 | Model 2 | 0.96 | 0.90 to 1.01 |
| Model 3 | 1.00 | 0.92 to 1.08 | Model 3 | 0.97 | 0.90 to 1.04 |
| Model 3+Lday | 1.02 | 0.94 to 1.11 | Model 3+Lday | 0.99 | 0.91 to 1.06 |
| | | | | | |
| Incident CVD | RR | 95%CI | Incident CVD | RR | 95%CI |
| Model 1 | 0.98 | 0.95 to 1.00 | Model 1 | 0.94 | 0.92 to 0.96 |
| Model 2 | 1.03 | 1.01 to 1.06 | Model 2 | 1.00 | 0.98 to 1.03 |
| Model 3 | 1.02 | 0.99 to 1.05 | Model 3 | 0.99 | 0.97 to 1.02 |
| Model 3+Lday | **1.04** | **1.00 to 1.07** | Model 3+Lday | 1.01 | 0.98 to 1.03 |

Model 1: adjusted for cohort, time segments of follow-up period;

Model 2: further adjusted for age, sex;
Model 3: further adjusted for education, employment status, smoking status and alcohol consumption (**main model**).
RR: Rate ratio.
Incident cerebrovascular disease (ICD10 I60-69); Incident ischaemic heart disease (ICD10 I20-I25); Incident CVD (ICD10 I00-I99). Lday: daytime (07:00-19:00) road traffic noise at sound level, dB(A).
BOLD indicates where significance level<0.05.

Most of the subgroup analyses were generally not statistically significant (Appendix-6.1). However, for total incident CVD, larger and significant associations with $PM_{10}$ were seen among men (RR: 1.06, 95%CI: 1.01 to 1.12), those aged≥60 years (RR: 1.10, 95%CI: 1.04 to 1.16) and those with a BMI≥30kg/$m^2$ (RR: 1.09, 95%CI: 1.00 to 1.19). Larger associations were also seen for diabetics than for non-diabetics for all three CVD outcomes although none of these associations were significant.

Both $PM_{10}$ and $NO_2$ effects on incident cerebrovascular diseases were higher in men and in those less educated, but opposite patterns were seen for incident ischaemic heart diseases. None of these associations were statistically significant.

### *Pooled analysis 2: physical pooling of data in Stata*

To validate the results from DataSHIELD, subsequent analyses on Stata, using Cox regression methods, were applied to the physically pooled data. Overall, the main results in Table 6.4 were very similar to those from the piecewise regression in DataSHIELD presented in Table 6.3.

Table 6.4 Associations between pan-European LUR modelled $PM_{10}$, $NO_2$ and incident cardiovascular disease: a physically pooled analysis from EPIC-Oxford, HUNT2 and UK Biobank

| **$PM_{10}$, per each IQR (4.1 µg/$m^3$)** | | | | **$NO_2$, per each IQR (13.2 µg/$m^3$)** | | | |
|---|---|---|---|---|---|---|---|
| Incident cerebrovascular diseases | N | HR | 95%CI | Incident cerebrovascular diseases | N | HR | 95%CI |
| Model 1 | 391,584 | 1.03 | 0.94 to 1.13 | Model 1 | 392,517 | 0.99 | 0.91 to 1.07 |
| Model 2 | 391,584 | 1.08 | 0.98 to 1.19 | Model 2 | 392,517 | 1.07 | 0.99 to 1.17 |
| Model 3 | 282,240 | 1.04 | 0.92 to 1.18 | Model 3 | 282,893 | 1.03 | 0.93 to 1.15 |
| Model 3+Lday | 278,253 | 1.03 | 0.90 to 1.17 | Model 3+Lday | 278,878 | 1.03 | 0.92 to 1.15 |
| | | | | | | | |
| Incident ischaemic heart diseases | N | HR | 95%CI | Incident ischaemic heart diseases | N | HR | 95%CI |
| Model 1 | 391,584 | 0.93 | 0.87 to 1.00 | Model 1 | 392,517 | 0.88 | 0.83 to 0.93 |
| Model 2 | 391,584 | 0.97 | 0.91 to 1.04 | Model 2 | 392,517 | 0.95 | 0.89 to 1.01 |
| Model 3 | 282,240 | 0.97 | 0.89 to 1.06 | Model 3 | 282,893 | 0.97 | 0.90 to 1.05 |
| Model 3+Lday | 278,253 | 0.99 | 0.91 to 1.08 | Model 3+Lday | 278,878 | 0.98 | 0.91 to 1.06 |
| | | | | | | | |

| Incident CVD | N | HR | 95%CI | Incident CVD | N | HR | 95%CI |
|---|---|---|---|---|---|---|---|
| Model 1 | 391,584 | 0.98 | 0.95 to 1.00 | Model 1 | 392,517 | 0.94 | 0.92 to 0.96 |
| Model 2 | 391,584 | 1.02 | 0.99 to 1.05 | Model 2 | 392,517 | 1.00 | 0.98 to 1.03 |
| Model 3 | 282,240 | 1.00 | 0.97 to 1.04 | Model 3 | 282,893 | 0.99 | 0.97 to 1.02 |
| Model 3+Lday | 278,253 | 1.02 | 0.98 to 1.05 | Model 3+Lday | 278,878 | 1.00 | 0.97 to 1.03 |

Model 1: adjusted for cohort;
Model 2: further adjusted for age, sex, year of recruitment;
Model 3: further adjusted for education, employment status, smoking status and alcohol consumption (**main model**).
HR: hazard ratio.
Incident cerebrovascular diseases (ICD10 I60-69); Incident ischaemic heart diseases (ICD10 I20-I25); Incident CVD (ICD10 I00-I99).

### *Pooled analysis 3: cohort-specific meta-analysis on Stata*

Both fixed-effect and random-effect cohort-specific meta-analyses were conducted using Stata. Overall, the pooled estimates obtained from the cohort-specific meta-analyses were similar to those obtained from the individual-level pooled analyses, either in DataSHIELD (Table 6.3) or Stata (Table 6.4).

In Model3, for each IQR higher $PM_{10}$ and $NO_2$, incident cerebrovascular disease increased non-significantly by 4% (HR: 1.04, 95%CI: 0.96 to 1.12) and 3% (HR: 1.03, 95%CI: 0.96 to 1.10) (Figure 6.3) respectively. These estimates were similar to those obtained from the individual-level analyses either by virtually pooling of data in DataSHIELD (Table 6.3) or physically pooling of data in Stata (Table 6.4). However, as expected, the confidence intervals were slightly wider for estimates obtained from the cohort-specific meta-analysis, compared with those obtained from pooled individual-level analysis.

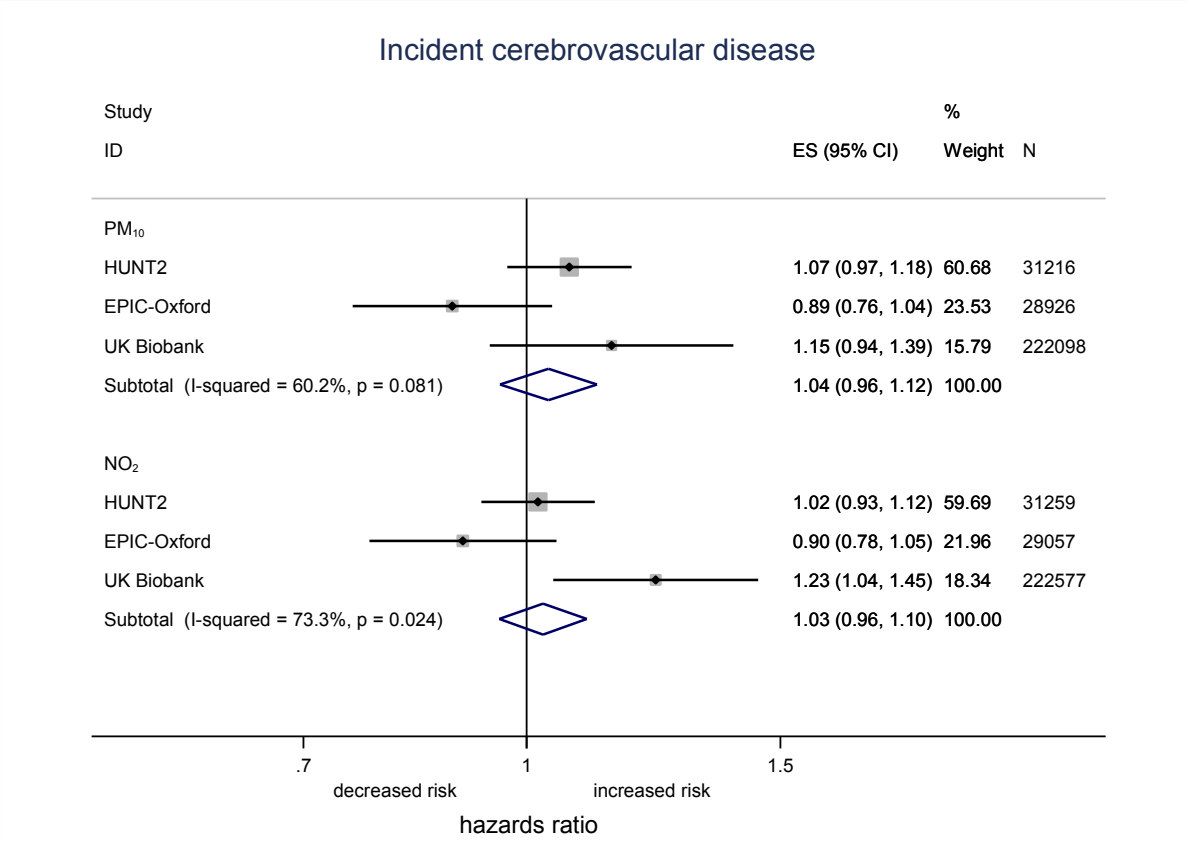
Heterogeneity existed across the three cohorts for the associations between $NO_2$ air pollution and incident cerebrovascular disease (I-square: 73%) or incident ischaemic heart disease (I-square: 70%).

A statistically significant association between $NO_2$ and incident cerebrovascular diseases was seen for UK Biobank only (Figure 6.3). Each IQR (11.3μg/$m^3$) increase of $NO_2$ was associated with a 23% higher incidence of cerebrovascular disease (HR: 1.23, 95%CI: 1.04 to 1.45) among UK Biobank participants. A non-significant association was observed for $PM_{10}$.

In HUNT2, statistically significant *negative* associations were observed for both air pollutants and incident ischaemic heart disease (Figure 6.4).

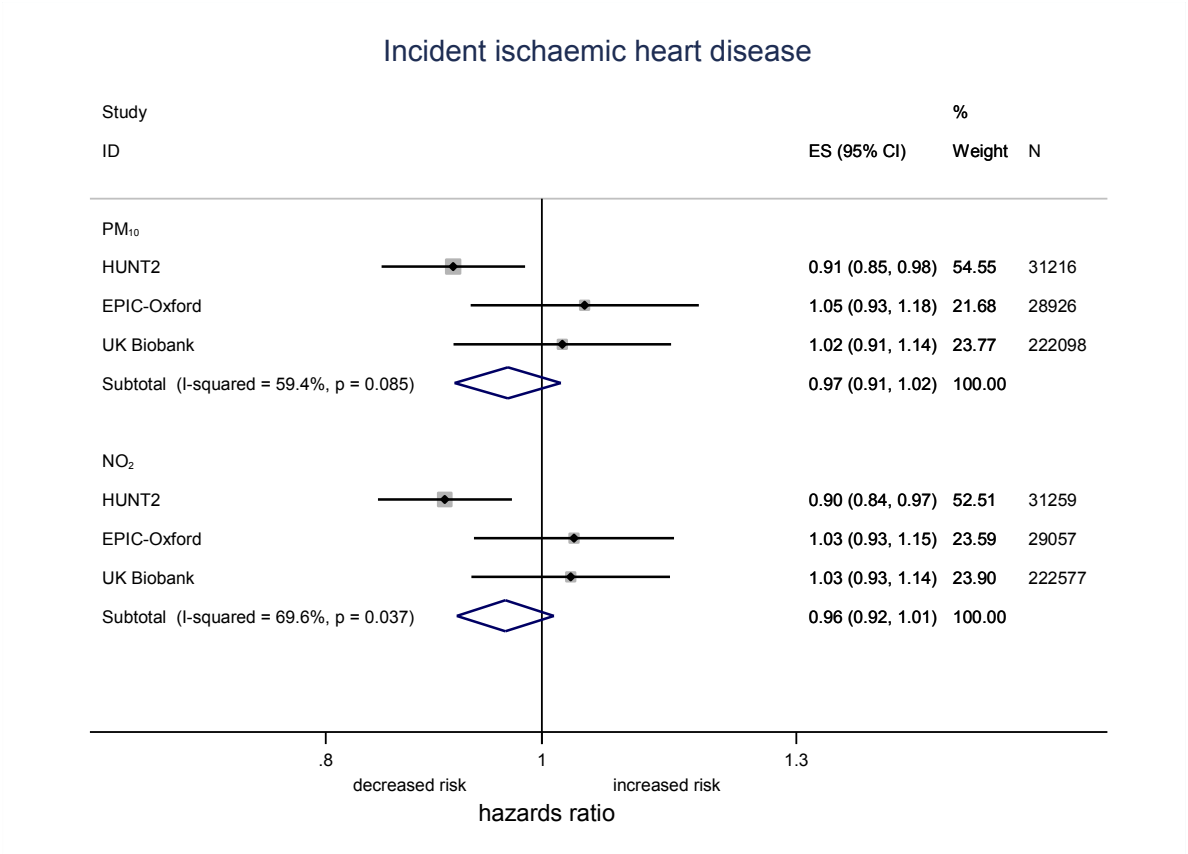
Significant positive associations between $PM_{10}$, $NO_2$ and total incident CVD were also observed in the UK Biobank (Appendix-6.2). Each IQR (11.3µg/$m^3$) higher $NO_2$ was associated with an 11% increase in total incident CVD (HR: 1.11, 95%CI: 1.07 to 1.16), whilst each IQR (3.4µg/$m^3$) higher $PM_{10}$ was associated with a 16% increase in total incident CVD (HR: 1.16, 95%CI: 1.10 to 1.22).

Figure 6.3 Effects of pan-European LUR modelled $PM_{10}$ and $NO_2$ on incident cerebrovascular disease (ICD10 I60-I69): cohort-specific meta-analysis on Model3. Hazard ratio expressed as per IQR increase of each air pollutant: IQR of $PM_{10}$ was 1.6, 3.4 and 3.4 µg/$m^3$, IQR of $NO_2$ was 5.1, 11.7 and 11.3 µg/$m^3$ for HUNT2, EPIC-Oxford and UK Biobank respectively.




I-squared, variation in estimated effect attributable to heterogeneity

Figure 6.4 Effects of pan-European LUR modelled $PM_{10}$ and $NO_2$ on incident ischaemic heart disease (ICD10 I20-I25): cohort-specific meta-analysis on Model3. Hazard ratio expressed as per IQR increase of each air pollutant: IQR of $PM_{10}$ was 1.6, 3.4 and 3.4 µg/m$^3$, IQR of $NO_2$ was 5.1, 11.7 and 11.3 µg/m$^3$ for HUNT2, EPIC-Oxford and UK Biobank respectively.




I-squared, variation in estimated effect attributable to heterogeneity

### 6.1.3. Road traffic noise and incident CVD outcomes

#### *Pooled analysis 1: virtual pooling of data in DataSHIELD*

In the main model (Model3), no significant associations were found between Lday and incident cerebrovascular disease (Table 6.5). There was a weak non-significant positive association between Lday and incident ischaemic heart disease (HR: 1.011, 95%CI: 0.981 to 1.041) per 3.9 dB(A) of Lday. A similar finding was also observed for total incident CVD (HR: 1.006, 95%CI: 0.994 to 1.018) per 3.9 dB(A) of Lday. Further adjusting for $PM_{10}$ or $NO_2$ did not change these associations, nor did further adjustment for BMI or ever-had diabetes (data not shown).

Table 6.5 Associations between Lday (07:00-19:00) road traffic noise and incident cardiovascular disease: a virtually pooled analysis across EPIC-Oxford, HUNT2 and UK Biobank

| **Lday, per each IQR (3.9 dB(A))** | | |
|---|---|---|
| **Incident cerebrovascular disease** | RR | 95%CI |
| Model 1 | **0.959** | **0.928 to 0.991** |
| Model 2 | 0.987 | 0.956 to 1.019 |
| Model 3 | 0.993 | 0.953 to 1.035 |
| Model 3+$PM_{10}$ | 0.992 | 0.951 to 1.034 |
| Model 3+$NO_2$ | 0.993 | 0.953 to 1.035 |
| | | |
| **Incident ischaemic heart disease** | RR | 95%CI |
| Model 1 | 0.998 | 0.974 to 1.022 |
| Model 2 | 1.018 | 0.994 to 1.042 |
| Model 3 | 1.011 | 0.981 to 1.041 |
| Model 3+$PM_{10}$ | 1.010 | 0.980 to 1.041 |
| Model 3+$NO_2$ | 1.011 | 0.981 to 1.041 |
| | | |
| **Incident CVD** | RR | 95%CI |
| Model 1 | **0.989** | **0.979 to 0.999** |
| Model 2 | 1.007 | 0.998 to 1.017 |
| Model 3 | 1.006 | 0.994 to 1.018 |
| Model 3+$PM_{10}$ | 1.005 | 0.993 to 1.017 |
| Model 3+$NO_2$ | 1.006 | 0.994 to 1.018 |

Model 1: adjusted for cohort, time segments of follow-up period;
Model 2: further adjusted for age, sex;
Model 3: further adjusted for education, employment status, smoking status and alcohol consumption (**main model**).
RR: Rate ratio.
Incident cerebrovascular disease (ICD10 I60-69); Incident ischaemic heart disease (ICD10 I20-I25); Incident CVD (ICD10 I00-I99). BOLD indicates where significance level<0.05.

In the subgroup analyses, significant associations were seen between Lday and incident ischaemic heart disease (RR: 1.072, 95%CI: 1.012 to 1.135) and total incident CVD (RR: 1.037, 95%CI: 1.010 to 1.064) among current-smokers. A borderline significant association was also seen among women for

the association with incident ischaemic heart diseases (HR: 1.052, 95%CI: 0.997 to 1.109). All other subgroup analyses were non-significant, with overlapping confidence intervals (Appendix-6.1).

### *Pooled analysis 2: physical pooling of data in Stata*

As with the air pollution analyses, the results from DataSHIELD were validated against analyses of the physically pooled data undertaken in Stata using Cox regression methods. Results from these Cox regression analyses on the physically pooled data are presented in Table 6.6, and are similar to those obtained from DataSHIELD in Table 6.5.

Table 6.6 Associations between Lday (07:00-19:00) road traffic noise and incident cardiovascular disease: a physically pooled analysis across EPIC-Oxford, HUNT2 and UK Biobank

| **Lday, per each IQR (3.9 dB(A))** | | | |
|---|---|---|---|
| **Incident cerebrovascular diseases** | N | HR | 95%CI |
| Model 1 | 384,683 | **0.959** | **0.928 to 0.991** |
| Model 2 | 384,653 | 0.982 | 0.951 to 1.014 |
| Model 3 | 278,867 | 0.997 | 0.953 to 1.045 |
| Model 3+$PM_{10}$ | 278,241 | 0.997 | 0.952 to 1.044 |
| Model 3+$NO_2$ | 278,866 | 0.997 | 0.952 to 1.044 |
| | | | |
| **Incident ischaemic heart diseases** | N | HR | 95%CI |
| Model 1 | 384,683 | 0.997 | 0.974 to 1.022 |
| Model 2 | 384,683 | 1.014 | 0.990 to 1.038 |
| Model 3 | 278,879 | 1.011 | 0.979 to 1.044 |
| Model 3+$PM_{10}$ | 278,253 | 1.011 | 0.979 to 1.044 |
| Model 3+$NO_2$ | 278,878 | 1.011 | 0.979 to 1.045 |
| | | | |
| **Incident CVD** | N | HR | 95%CI |
| Model 1 | 384,683 | **0.986** | **0.976 to 0.996** |
| Model 2 | 384,683 | 1.001 | 0.991 to 1.011 |
| Model 3 | 278,879 | 1.002 | 0.989 to 1.016 |
| Model 3+$PM_{10}$ | 278,253 | 1.002 | 0.989 to 1.015 |
| Model 3+$NO_2$ | 278,878 | 1.002 | 0.989 to 1.016 |

Model 1: adjusted for cohort;
Model 2: further adjusted for age, sex, year of recruitment;
Model 3: further adjusted for education, employment status, smoking status and alcohol consumption (**main model**).
HR: hazard ratio.
Incident cerebrovascular diseases (ICD10 I60-69); Incident ischaemic heart diseases (ICD10 I20-I25); Incident CVD (ICD10 I00-I99).
BOLD indicates where significance level<0.05.

Additional analyses were conducted in Stata to investigate the associations by noise categories (Table 6.7). Compared to those exposed to Lday levels of less than 50 dB(A), non-significantly increased

hazard ratios for incident cerebrovascular disease were found for those exposed to Lday levels between 50-55 dB(A) (HR: 1.024, 95%CI: 0.877 to 1.196) and those exposed to a level higher than 55 dB(A) (HR: 1.038, 95%CI: 0.843 to 1.278). However, this suggestive positive trend was not seen for incident ischaemic heart disease or total incident CVD.

Table 6.7 Associations between Lday (07:00-19:00), Lnight (23:00-07:00) road traffic noise and incident cardiovascular disease: a pooled analysis across EPIC-Oxford, HUNT2 and UK Biobank on Model 3*

| | N | Incident cerebrovascular disease | | Incident ischaemic heart disease | | Incident CVD | |
|---|---|---|---|---|---|---|---|
| **Lday, dB(A)** | | HR (95%CI) | | HR (95%CI) | | HR (95%CI) | |
| **<50** | 17,937 | 1 | | 1 | | 1 | |
| **50-55** | 160,158 | 1.024 | 0.877 to 1.196 | 1.038 | 0.929 to 1.161 | 1.042 | 0.994 to 1.092 |
| **>=55** | 100,577 | 1.038 | 0.843 to 1.278 | 0.949 | 0.818 to 1.100 | 0.996 | 0.938 to 1.058 |
| | | **Incident cerebrovascular disease** | | **Incident ischaemic heart disease** | | **Incident CVD** | |
| **Lnight, dB(A)** | | HR (95%CI) | | HR (95%CI) | | HR (95%CI) | |
| **<40** | 11,884 | 1 | | 1 | | 1 | |
| **40-45** | 114,841 | 0.989 | 0.860 to 1.138 | 1.061 | 0.958 to 1.174 | 1.022 | 0.979 to 1.067 |
| **>=45** | 151,947 | 0.934 | 0.776 to 1.125 | 1.037 | 0.910 to 1.181 | 0.989 | 0.936 to 1.043 |

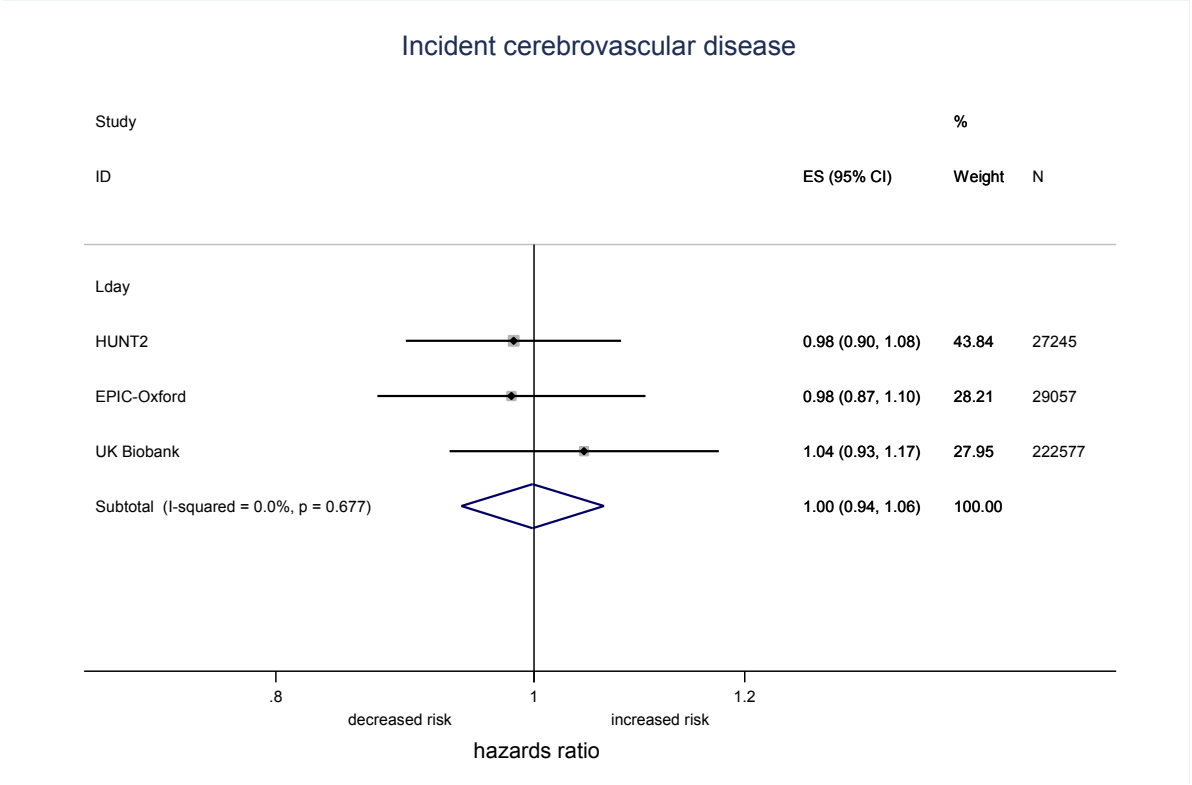
*adjusted for cohort, age, sex, year of recruitment, education, employment status, smoking status and alcohol consumption in grams per week.
HR: hazard ratio.
Incident cerebrovascular disease (ICD10 I60-69); Incident ischaemic heart disease (ICD10 I20-I25); Incident CVD (ICD10 I00-I99).

### *Pooled analysis 3: cohort-specific meta-analysis in Stata*

Overall, the pooled estimates obtained from cohort-specific meta-analyses were similar to those obtained from pooled individual-level analyses. There was no heterogeneity observed across cohorts for any of the studied associations. Null associations were seen for incident cerebrovascular disease (Figure 6.5).

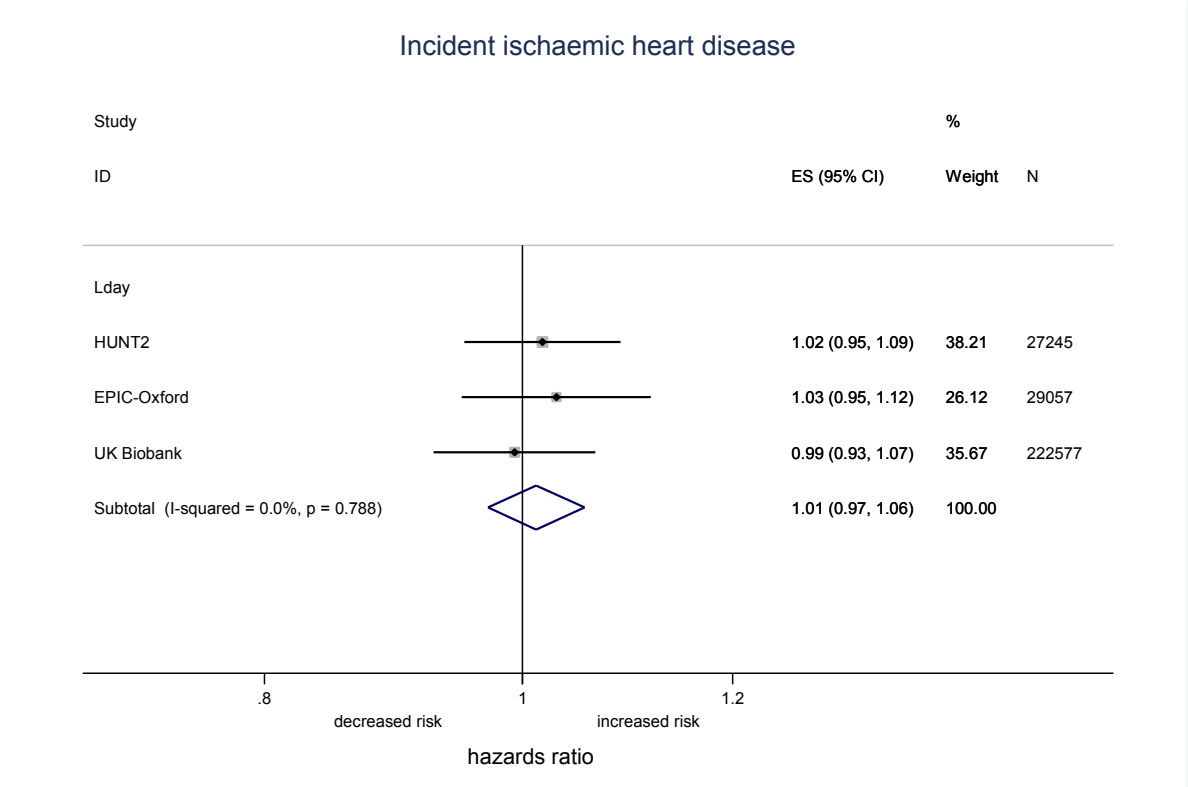
From the fixed effects (I-V) model, each IQR increase Lday was non-significantly associated with an increased risk of incident ischaemic heart disease (HR: 1.01, 95%CI: 0.97 to 1.06) (Figure 6.6), similar to the findings from the pooled individual-level analyses (Table 6.5 and 6.6). A similar pooled effect estimate was seen for total incident CVD (Appendix-6.3).

Figure 6.5 Effects of day-time (07:00-19:00) road traffic noise Lday on incident cerebrovascular disease (ICD10 I60-I69): cohort-specific meta-analysis on model3. Hazard ratio expressed per IQR increase in noise: IQR of Lday was 3.6, 6.8 and 3.5 dB(A) for EPIC-Oxford, HUNT2 and UK Biobank respectively.




I-squared, variation in estimated effect attributable to heterogeneity

Figure 6.6 Effects of day-time (07:00-19:00) road traffic noise Lday on incident ischaemic heart disease (ICD10 I20-I25): cohort-specific meta-analysis on model3. Hazard ratio expressed per IQR increase in noise: IQR of Lday was 3.6, 6.8 and 3.5 dB(A) for EPIC-Oxford,HUNT2 and UK Biobank respectively.




I-squared, variation in estimated effect attributable to heterogeneity;

### 6.1.4. ESCAPE LUR air pollutant metrics: EPIC-Oxford and UK Biobank only

ESCAPE air pollution data were available for both EPIC-Oxford and UK Biobank. Additional analyses were conducted by physically pooling data from both cohorts, using the ESCAPE LUR modelled air pollutant metrics.

ESCAPE LUR modelled $NO_2$ and $PM_{10}$ was positively but non-significantly associated with all three CVD outcomes (Table 6.8). The association between $NO_2$ and total incident CVD was borderline statistically significant (HR: 1.03, 95%CI: 1.00 to 1.06, P-value: 0.08) per 10 µg/m$^3$ of $NO_2$. There was a statistically significant association between $PM_{2.5}$ and total incident CVD (HR: 1.03, 95%CI: 1.00 to 1.07, P-value: 0.03) per 1.3 µg/m$^3$ of $PM_{2.5}$. Further adjusting for Lday did not change these observed associations.

Table 6.8 Associations between each ESCAPE LUR modelled air pollutant metric (per IQR higher) and each CVD outcome: physical pooling data from EPIC-Oxford and UK Biobank, based on Model3.

| | N | Incident cerebrovascular diseases | | Incident ischaemic heart diseases | | Incident CVD | |
|---|---|---|---|---|---|---|---|
| | | HR | 95%CI | HR | 95%CI | HR | 95%CI |
| **$NO_2$** | 251,629 | 1.07 | 0.94 to 1.21 | 1.03 | 0.94 to 1.12 | 1.03 | 1.00 to 1.06 |
| **$PM_{10}$** | 233,601 | 1.00 | 0.91 to 1.10 | 1.01 | 0.95 to 1.07 | 1.01 | 0.98 to 1.03 |
| **$PM_{2.5}$** | 233,601 | 1.05 | 0.93 to 1.19 | 1.04 | 0.96 to 1.12 | **1.03** | **1.00 to 1.07** |
| **PMcoarse** | 233,596 | 1.00 | 0.89 to 1.12 | 0.98 | 0.92 to 1.05 | 0.97 | 0.94 to 0.99 |
| **$PM_{2.5abs}$** | 233,601 | 1.04 | 0.91 to 1.15 | 0.96 | 0.88 to 1.04 | 1.01 | 0.98 to 1.04 |

Model3: adjusted for study, age, sex, year of recruitment, education, employment, smoking status and alcohol consumption.
Hazard ratio (95%CI) was calculated for each IQR increase of $NO_2$ (10 μg/$m^3$), $PM_{10}$ (1.8 μg/$m^3$), $PM_{2.5}$ (1.3μg/$m^3$), PMcoarse (1.4μg/$m^3$) and PM2.5absorbance (0.3 $10^{-5}$ /m).
HR: hazard ratio.
Incident cerebrovascular diseases (ICD10 I60-69); Incident ischaemic heart diseases (ICD10 I20-I25); Incident CVD (ICD10 I00-I99).
BOLD indicates where significance level<0.05.

## 6.2. Discussion

### 6.2.1. Ambient air pollution effects

#### *Air pollution and cerebrovascular disease*

In this PhD work, pooling harmonised data from EPIC-Oxford, HUNT2 and UK Biobank, I found a positive but non-significant association between air pollution and incident cerebrovascular disease (ICD-10: I60-I69). Effect sizes were similar for an IQR higher exposure to both pan-European LUR modelled $PM_{10}$ and $NO_2$.These estimates were adjusted for some major risk factors and independent of day-time road traffic noise effects.

To date, only a small number of European studies have investigated the long-term effects of air pollution on incident cerebrovascular disease. These studies were conducted in a range of geographical regions across Europe and in various study designs. As with the results I report here, most previous studies did not find a statistically significant association, although they did suggest that air pollution is a potential risk factor for cerebrovascular diseases.

A study of 20,070 participants in Stockholm investigated the effects of long-term exposure to road traffic air pollution on stroke (ICD-10: I61-64) incidence in four local cohorts, using cohort-specific meta-analysis(Korek et al., 2015). They found that a 20 µg/$m^3$ increase in nitrogen oxides (NOx) at recruitment address was associated with a non-significant 16% increase of incident stroke (HR: 1.16, 95%CI: 0.83 to 1.61). A similar hazard ratio was seen per 10 µg/$m^3$ increase in $PM_{10}$. The study was conducted in a region with relatively lower air pollution levels than other regions across Europe and annual exposure estimates were modelled at home address every year for each participant from the study entry at 1992-2004 up to the end of follow-up at 2011, taking into account the changes of residence. Of note, there was no marked difference in risk estimates regarding the air pollution effects on either ischaemic stroke or haemorrhagic stroke. Also, the observed association between air pollution and stroke was not affected by the use of different exposure windows. In fact, the authors found that the observed associations for NOx or $PM_{10}$ did not change materially when using either the time-weighted exposure 0-2 years or 6-10 years prior to a stroke event. Two other Scandinavian

studies have also investigated air pollution effects on stroke. In Oslo, a 10 μg/$m^3$ increase in $NO_2$ at recruitment address during 1974-1978 was non-significantly associated with a 4% increased risk of dying from cerebrovascular diseases up to 1998 (HR: 1.04, 95%CI: 0.94 to 1.15) in 16,209 middle-aged men(Nafstad et al., 2004). Again, the authors reported that exposure time window was not an important factor in modifying the observed association. In the Danish Diet, Cancer, and Health (DCH) Cohort, an IQR (6.2 μg/$m^3$) increase of $NO_2$ since 1971 to the end of follow-up in 2006 was non-significantly associated with a 5% increase in incident stroke (HR: 1.05, 95%CI: 0.99 to 1.11) and a 22% increase in fatal stroke (HR: 1.22, 95%CI: 0.99 to 1.49)(Andersen et al., 2012a). Unlike the Stockholm study, this Danish study found the association was stronger for ischaemic stroke than for haemorrhagic stroke. For all three Scandinavian studies, historical modelled individual-level air pollution data were available years if not decades prior to the incident stroke event.

Maheswaran and colleagues conducted several small-area studies in the UK on this topic(Maheswaran et al., 2012; Maheswaran et al., 2005; Maheswaran & Elliott, 2003). In a study in South London, incident stroke cases in a defined region were registered during 1995-2004. For each 10 μg/$m^3$ increase in $PM_{10}$ or $NO_2$ in 2002, the adjusted rate ratio for incident ischaemic stroke was non-significantly increased at 1.22 (95%CI: 0.77 to 1.93) and 1.11 (95%CI: 0.93 to 1.32) respectively in all age groups(Maheswaran et al., 2012). Although the overall association was non-significant, they found a significantly increased risk of incident ischaemic stroke with $PM_{10}$ exposure among those aged 65-79 years (adjusted rate ratio: 1.86, 95%CI: 1.10 to 3.13) but not those aged 80 years and above. In both this PhD study and the Danish DCH cohort study(Andersen et al., 2012a), associations were not materially different by different age groups. No evidence was found for incident haemorrhagic stroke. Two earlier small-area studies were also conducted by Maheswaran and colleagues(Maheswaran et al., 2005; Maheswaran & Elliott, 2003). One reported that living within 200m of a main road was significantly associated with an excess risk (5%, 95%CI: 4% to 7%) of dying from stroke in men and women in England and Wales(Maheswaran & Elliott, 2003). The other, using modelled air pollutant exposure ($NO_X$ and $PM_{10}$), further confirmed that a significant excess risk of stroke mortality was

observed among areas in the highest exposure group in Sheffield(Maheswaran et al., 2005). This Sheffield study also suggested that air pollution had stronger effects on stroke mortality (37%, 95%CI: 19% to 57%) than morbidity (13%, 95%CI: 1% to 27%), comparing highest exposure group of NOx to the lowest. Later the same research group found that outdoor air pollution increased the risk of death after first stroke among 3,323 stroke cases in London(Maheswaran et al., 2010). It should be noted that although very small geographical units were used in these studies, the possibility of ecological bias cannot be ruled out. Further, these small-area studies did not take into account individual-level risk factors such as smoking or alcohol consumption, which makes it difficult to compare my results with theirs.

The first report on long-term air pollution effects on incident CVD in a large UK population and taking into account individual risk factors was provided by Atkinson and colleagues(Atkinson et al., 2013). They investigated effects of annual $PM_{10}$ and $NO_2$ exposures on a range of incident CVD over a 5-year period (2003-2007), using data from 836,557 patients who registered in general practice across England. In their study, they found non-significantly increased hazard ratios for incident stroke of 1.01 (95%CI: 0.98 to 1.04) for each IQR (3.0 $\mu g/m^3$) increase of $PM_{10}$ and 1.03 (95%CI: 0.99 to 1.07) per each IQR (10.7 $\mu g/m^3$) increase of $NO_2$ respectively. These estimates were very similar to those I found in this PhD work using selected cohorts, although I have included all incident cases of cerebrovascular disease (ICD10: I60-I69) whilst Atkinson et al only investigated stroke (ICD10: I61, I63-I64) specifically.

There have been other studies conducted in other parts of Europe. In the Swiss National Cohort, the hazard ratio for stroke mortality was close to unity in relation to $PM_{10}$ exposure(Huss et al., 2010). In contrast, in a Dutch cohort, background $NO_2$ and $PM_{2.5}$ were significantly associated with cerebrovascular mortality (HR: 1.51, 95%CI: 1.07 to 2.12, per 30 $\mu g/m^3$ of $NO_2$, and HR: 1.62, 95%CI: 1.07 to 2.44, per 10 $\mu g/m^3$ of $PM_{2.5}$)(Beelen et al., 2009). In the German Heinz Nixdorf Recall (HNR) cohort of 4,433 participants in Ruhr area, $PM_{10}$ exposure was significantly associated with stroke incidence (HR: 2.61, 95%CI: 1.13 to 6.00)(Hoffmann et al., 2015). A slightly stronger significant

association was also observed for $PM_{2.5}$. In the Greek component of the EPIC cohort study, where air pollution levels were higher than those of other studies conducted in North-west Europe, a 10 μg/m$^3$ $PM_{10}$ was non-significantly associated with a hazard ratio of 1.17 (95%CI: 0.60 to 2.26) for incident stroke(Katsoulis et al., 2014). A census-based study in Rome also with a relatively higher level of air pollution reported that each 10 μg/m$^3$ increase of $PM_{2.5}$ was significantly associated with an excess risk of mortality from cerebrovascular disease (HR: 1.08, 95%CI: 1.04 to 1.13)(Cesaroni et al., 2013). However, an effect was not seen for $NO_2$ (HR: 1.01, 95%CI: 0.99 to 1.03).

Recently, two European-wide studies conducted by the ESCAPE consortium provided further evidence. A meta-analysis of 11 cohorts across Europe (Sweden, Finland, Denmark, Germany, Italy) reported no association between a 10 μg/m$^3$ increase of $PM_{10}$ and incident stroke (HR: 1.11, 95%CI: 0.90 to 1.36)(Stafoggia et al., 2014). The hazard ratio for $NO_2$ was close to unity. This is contrast to my findings in the pooled analyses across EPIC-Oxford and UK Biobank, using the same ESCAPE exposure metrics. I found that the hazard ratio was slightly stronger for $NO_2$ (HR: 1.07, 95%CI: 0.94 to 1.21) but unity for $PM_{10}$ (HR: 1.00, 95%CI: 0.91 to 1.10) per an IQR increase of 10 and 1.8 μg/m$^3$ respectively. Interestingly, in the ESCAPE study the authors found significantly increased risks for incident stroke with $PM_{2.5}$ exposure among those aged ≥60 years, never-smokers and those with an exposure level of $PM_{2.5}$ less than 25 μg/m$^3$.

Another ESCAPE study of 22 European cohorts found that most hazard ratios for the associations between air pollution and mortality from cardiovascular diseases were close to unity, except for between $PM_{10}$ and mortality from cerebrovascular diseases (HR: 1.22, 95%CI: 0.91 to 1.63, per 10 μg/m$^3$ of $PM_{10}$)(Beelen et al., 2014b). EPIC-Oxford was part of this ESCAPE mortality study. In fact, the hazard ratio for cerebrovascular mortality found for EPIC-Oxford was similar to what I found for cerebrovascular incidence in EPIC-Oxford using ESCAPE metrics.

I have summarised all the aforementioned European studies in Table 6.10. Including this PhD work, there have been 13 studies across Europe investigating this topic to date. These studies, most of which

were published in the last 3-4 years, have different study designs, follow-up periods, exposure assessment approaches, exposure time windows and outcome definitions. Nevertheless, the positive but non-significant associations with both $PM_{10}$ and $NO_2$ that I present here are fairly consistent across these European studies. In fact, only three studies detected a significant association(Hoffmann et al., 2015; Cesaroni et al., 2013; Beelen et al., 2009), two of which studied cerebrovascular mortality specifically(Cesaroni et al., 2013; Beelen et al., 2009).

A meta-analysis published in 2015 suggested a significantly positive association between short-term air pollution and stroke hospital admission or stroke mortality across different parts of the world(Shah et al., 2015). However, the long-term effects of air pollution on stroke or total cerebrovascular disease are less clear. More recently, studies on the long-term air pollution effects in Asia, North America and Europe were included in an analytical review published by Scheers et al in late 2015(Scheers et al., 2015). Overall, among 20 studies investigating stroke incidence and 12 studies investigating stroke mortality, a 10 $\mu g/m^3$ increase of $PM_{10}$ was associated with a hazard ratio of 1.061 (95%CI: 1.018 to 1.105) for incident events and a hazard ratio of 1.080 (95%CI: 0.992 to 1.177) for mortality events respectively. In Europe, the respective figures for incident and mortality events were 1.057 (95%CI: 0.973 to 1.148) based on eight studies and 1.213 (95%CI: 0.955 to 1.541) based on five studies.

Scheers et al further combined the data from Europe and North America. They observed that a 5 $\mu g/m^3$ increase of $PM_{2.5}$ was associated with a hazard ratio of 1.064 (95%CI: 1.021 to 1.109) for incident stroke events and a hazard ratio of 1.125 (95%CI: 1.007 to 1.256) for stroke mortality events. Associations with $PM_{10}$ were also positive but not statistically significant.

Since most of the European studies mentioned here had already been included in Scheers's review, I further complement these findings by pooling hazard ratios of $NO_2$ effects from these European studies (Table 6.9). To my knowledge, there is no published analytical review of the long-term effects of $NO_2$ on incident stroke or cerebrovascular disease in European studies. I excluded those five studies investigating stroke mortality, and focused on incident events. For incident studies, the ones in

Stockholm(Korek et al., 2015) and in Denmark(Andersen et al., 2012a) were already included as parts of the ESCAPE study(Stafoggia et al., 2014). Therefore, five estimates, including the one from this PhD work, were included, giving a hazard ratio of 1.031 (95%CI: 0.997 to 1.066) per 10 μg/$m^3$ increase of $NO_2$ (Figure 6.8). No heterogeneity across the studies was detected. Including the five mortality studies alone in the meta-analysis gave a hazard ratio of 1.02 (95%CI: 1.01 to 1.04) per 10 μg/$m^3$ increase of $NO_2$.

## General issues

Some issues need to be addressed. First, I have included all incident cerebrovascular cases (ICD10: I60-I69) in my main analyses and did not present results specifically for different subtypes of cerebrovascular diseases. Only three previous long-term air pollution studies in Europe have published the results for ischaemic stroke and haemorrhagic stroke separately(Korek et al., 2015; Andersen et al., 2012a; Maheswaran et al., 2012). One found that the observed associations were not much different between the two subtypes(Korek et al., 2015), which was probably due to the limited number of haemorrhagic stroke cases in that particular study, while the other two much larger studies showed stronger associations for ischaemic stroke(Andersen et al., 2012a; Maheswaran et al., 2012). In general, most other literature also suggests that air pollution may have stronger effects on ischaemic stroke(Scheers et al., 2015). Among the 2,871 incident cerebrovascular disease in this PhD study, 501 (17%) were classified as haemorrhagic stroke (ICD10: I60, I61, I62) and 1,334 (46%) were classified as ischaemic stroke (ICD10: I63). When I re-ran the main analyses for each subtype of stroke, I found a positive, but non-significant, association between $PM_{10}$ and $NO_2$ and ischaemic stroke, and null associations with haemorrhagic stroke (data not shown). It seems that air pollution effects may possibly be confined to ischaemic stroke only although more studies are warranted. This is possibly because haemorrhagic stroke is less common than ischaemic stroke and therefore studies usually lack power to detect effects if present(Scheers et al., 2015). However, and perhaps more importantly, ischaemic and haemorrhagic stroke have different pathogenesis(Scheers et al., 2015). For example, ischaemic stroke is more usually related to atherosclerosis, a general cardiovascular disease which has

been associated with long-term exposure to air pollution in the emerging epidemiological studies(Kunzli et al., 2010).

Second, as with most other European studies, I observed positive but non-significant associations between $NO_2$ and incident cerebrovascular events in the pooled analyses using both the pan-European and ESCAPE LUR modelled metrics. In UK Biobank only, the largest cohort of this PhD study, a significant association was observed. While pooling all the estimates from these European studies, I found a borderline significant excess risk of incident cerebrovascular event in relation to long-term $NO_2$ exposure, with a pooled effect size similar to that reported by Scheers et al for long-term $PM_{10}$ exposure among the reviewed European studies. This permits speculation that long-term exposure to $NO_2$ on stroke /cerebrovascular events is at least as important as that of PM.

Cardiovascular effects of long-term exposure to $NO_2$ are still under debate. It remains uncertain whether $NO_2$ on its own acts a casual pollutant or merely an indicator of other pollutants emitted from the same source. Road traffic is the shared source for $NO_2$, PM and noise, it is therefore likely that the effect of each pollutant on a specific CVD outcome may be confounded by these other exposures. I did not adjust for PM effects in the $NO_2$ models as correlations between PM and $NO_2$ were high ($r>0.6$) which precluded a co-pollutant analysis. According to an interim report published by the UK Committee on the Medical Effects of Air Pollutants (COMEAP) in December 2015, coefficients from a single-pollutant model of either $NO_2$ or $PM_{2.5}$ are likely to be overestimated as there may be some substantial overlaps between the two pollutants(UK Committee on the Medical Effects of Air Pollutants, December 2015). Nevertheless, the COMEAP suggested that although the possibility of $NO_2$ as an indicator of other pollutants cannot be entirely ruled out, increasing scientific evidence shows that $NO_2$ should be sensibly treated as a possible causal air pollutant which exerts its own health impacts.

This suggestion was recently supported by a meta-analysis of worldwide studies which reported a significant positive association between short-term $NO_2$ exposure and stroke(Shah et al., 2015). One

of the key strengths of this PhD study was that road traffic noise was adjusted for in the air pollution models. $NO_2$ or PM effects on incident cerebrovascular diseases did not change substantially after adjustment for day-time road traffic noise in my study, which suggested an independent effect of each. This is in line with two earlier ESCAPE studies(Beelen et al., 2014b; Stafoggia et al., 2014) and most other literature as reviewed in Chapter 1. However, findings from another study of the Danish DCH cohort suggested a possible confounding effect(Sorensen et al., 2014). In the Danish DCH cohort study of 57,053 participants, $NO_2$ and road traffic noise were each significantly associated with incident ischaemic stroke in a single-pollutant model. However, when mutually adjusted, only the effect of road traffic noise remained statistically significant. These findings should be interpreted with caution, as a moderately high correlation (r=0.66) was found between the road traffic noise and $NO_2$ in this Danish study, which suggested a possible collinearity issue in the model specifications.

Third, results from this current PhD work and those presented in Figure 6.7, show consistently stronger associations between long-term $PM_{2.5}$ exposure and stroke incidence than $PM_{10}$. This contrast was also observed in a recent review on short-term air pollution effects on hospitalisation for stroke(Shah et al., 2015). Possible explanations include the possibility that smaller particles are likely to cause additional systemic cardiovascular effects, and find it easier to enter the blood-brain barrier and impair the neural cells(Genc et al., 2012). However, since the long-term $PM_{2.5}$ effect on stroke incidence was only studied among a few studies to date, there is a need for more studies to confirm whether this pollutant is more hazardous to stroke than PMcoarse or $PM_{10}$.

Fourth, in this PhD, I observed that both $PM_{10}$ and $NO_2$ effects were stronger in men than in women (*P-value* for interaction<0.05), although most previous studies did not find significant effect modification by sex. Similar to the findings of the Danish DCH cohort study(Andersen et al., 2012a), I also found the effects of $NO_2$ on stroke were attenuated among those with more years of education (p-value for interaction<0.05). Similar findings were seen for $PM_{10}$ but the effect modification was not significant (p-value for interaction=0.06). There were some suggestions of possible effect

modifications by BMI in this PhD study but the interaction terms were not statistically significant. Contrary to the ESCAPE study(Stafoggia et al., 2014), I observed stronger associations between $PM_{10}$, but not $NO_2$, on stroke incidence among ex-smokers and current-smokers, which may suggest a synergic effect of both air pollution and tobacco smoking on stroke incidence. Reports of effect modifications on the associations between long-term air pollution and stroke incidence are rather mixed in the current literature and it is difficult to draw a clear conclusion based on current knowledge. Nevertheless, this PhD study benefited from its pooled individual-level analysis design, which provided a large study sample within which to detect significant interaction terms, if any.

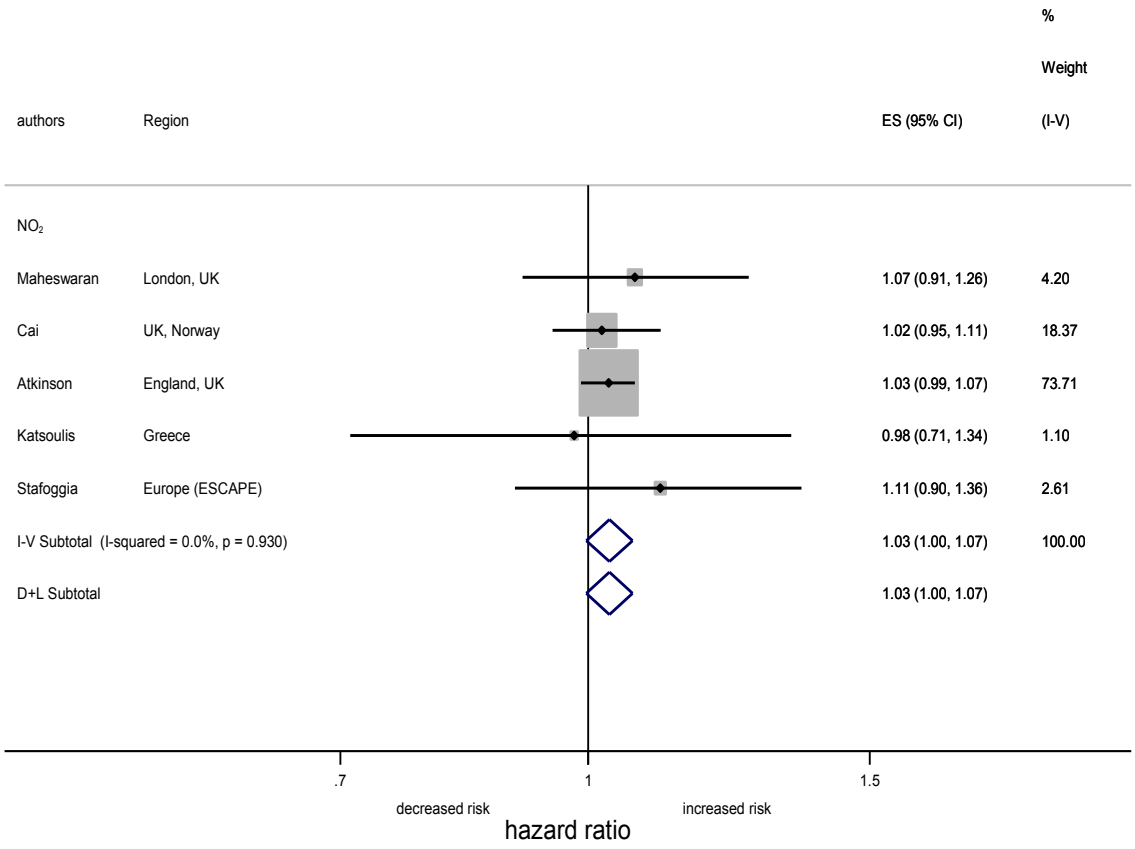
Considering both the previous studies and this PhD work through meta-analysis, there is suggestive evidence for an association between long-term PM or $NO_2$ air pollution exposure and cerebrovascular incidence among middle-aged adults. The results from this PhD work strengthen evidence for a possible dose-response relationship, by analysing harmonised data from three of the largest European cohorts. Also, pooled data were analysed on individual-level, with results were further validated, and heterogeneity explored using cohort-specific meta-analysis. Two sets of air pollution data (pan-European and ESCAPE) were used in the analyses and results were comparable. Only a few previous studies have been able to adjust for noise in the analyses, with one including aircraft noise(Huss et al., 2010), one night-time noise(Hoffmann et al., 2015) and three day-time road traffic noise(Beelen et al., 2014b; Stafoggia et al., 2014; Sorensen et al., 2014). All of these studies, except one(Sorensen et al., 2014), suggested that traffic noise was unlikely to have confounded the relationship between air pollution and stroke incidence or mortality. Results from my analyses are in support of an independent effect of air pollution and noise on incident cerebrovascular disease.

| Authors | Published year | Region | Study type | Study entry | Study end | Incidence rate (cases/person-year ) | ICD code | Endpoint | Mean age or age range, years | $NO_2$ | $PM_{10}$ | $PM_{2.5}$ | Independent of noise |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Korek** | 2015 | Stockholm, Sweden | cohort | 1992-2004 | 2011 | 3.6/1000 | ICD10: I61-I64 | incidence | 60 | 1.08 (0.91-1.27) | 1.14 (0.68-1.90) | n/a | n/a |
| **Nafstad** | 2004 | Oslo, Norway | cohort | 1972-1973 | 1998 | 0.7/1000 | ICD10:I60-I69 | mortality | 40-49 | 1.04 (0.94-1.15) | n/a | n/a | n/a |
| **Andersen** | 2012 | Denmark | cohort | 1993-1997 | 2006 | 3.9/1000 | ICD10: I60, I63, I64 | incidence /mortality | 56 | 1.08 (0.98-1.18) for incidence; 1.38 (0.98-1.90) for mortality | n/a | n/a | n/a |
| **Maheswaran** | 2012 | London, UK | small-area | 1995 | 2004 | 0.97/1000 | unkown | incidence | all | 1.07 (0.91-1.26) | 1.09 (0.72-1.65) | n/a | n/a |
| **Cai [this PhD]** | **2016** | **UK, Noway** | **cohort** | **1993-2010** | **2012-2015** | **1.7/1000** | **ICD10:I60-I69** | **incidence** | **55** | **1.02 (0.95-1.11)** | **1.10 (0.82-1.50)** | **1.46 (0.57-3.81)** | **yes** |
| **Atkinson** | 2013 | England, UK | census-based | 2003 | 2007 | 13012 cases /819370 subjects | unknown | incidence | 40-89 | 1.03 (0.99-1.07) | 1.03 (0.93-1.14) | n/a | n/a |
| **Beelen** | 2009 | Netherlands | cohort | 1986 | 1987-1996 | 1175 deaths/111391 subjects | ICD10:I60-I69 | mortality | 55-69 | 1.15 (1.02-1.28) | n/a | 1.62 (1.07-2.44) | n/a |
| **Huss** | 2010 | Switzerland | census-based | 2000 | 2005 | 1.1/1000 | ICD10: I60-64 | mortality | >=30 | n/a | 0.99 (0.98-1.00) | n/a | yes (aircraft) |
| **Hoffmann** | 2015 | Ruhr, Germany | cohort | 2000-2003 | 2008-2011 | 2.0/1000 | ICD10: I61, I63,I64 | incidence | 45-74 | n/a | 4.58 (1.21-17.2) | 27.75 (1.93-392.7) | yes |
| **Cesaroni** | 2013 | Rome, Italy | census-based | 2001 | 2010 | 0.13/1000 | ICD9: 430-438 | mortality | >=30 | 1.01 (0.99-1.03) | n/a | 1.08 (1.04-1.13) | n/a |
| **Katsoulis** | 2014 | Greece | cohort | 1997 | 2011 | 60 cases/2752 subjects | ICD10:I60-I69 | incidence | 47 | 0.98 (0.71-1.34) | 1.17 (0.60-2.26) | n/a | n/a |
| **Stafoggia** | 2014 | Europe (ESCAPE) | cohort | 1992 | 2010 | 2.7/1000 | ICD10:I61-I64 | incidence | 44-74 | 0.99 (0.89-1.11) | 1.11 (0.90-1.36) | 1.42 (0.77-2.62) | yes |
| **Beelen** | 2014 | Europe (ESCAPE) | cohort | 1985 | 2012 | 0.48/1000 | ICD10:I60-I69 | mortality | 51 | 1.01 (0.93-1.10) | 1.22 (0.91-1.63) | 1.46 (0.76-2.86) | yes |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ***Scheer*** | *2015* | *Europe (Scheers's review)* | | | | | | | | | *1.06 (0.97-1.15) for incidence; 1.21 (0.96-1.54) for mortality* | *1.54 (1.12-2.07)* | *n/a* |

Table 6.9 a summary of European (north to south) studies on long-term air pollution effects (scaled as per 10 µg/$m^3$ increase) on incident cerebrovascular events. Bold indicates the results from this PhD work.

Figure 6.7 Meta-analysis of European studies on long-term effects of per 10 µg/$m^3$ increase $NO_2$ (or NOx) on incident cerebrovascular disease




I-squared, variation in estimated effect attributable to heterogeneity; I-V, inverse-variance weighted fixed effects method; D-L, DerSimonian-Laird random effects method

### *Air pollution and Ischaemic heart disease*

Pooling data from EPIC-Oxford, HUNT2 and UK Biobank, I found null associations between the pan-European LUR modelled $PM_{10}$ or $NO_2$ and incident ischaemic heart disease (IHD). In the cohort-specific analyses, statistically significant negative associations were found in HUNT2 whilst positive non-significant associations were found in both EPIC-Oxford and UK Biobank.

The vast majority of previous studies have focused on air pollution effects on mortality from IHD or cardiovascular diseases in general. Only in recent years has evidence for an effect on incident cardiovascular events started to emerge. Results from this PhD work are consistent with previous studies in Europe and North America.

11 European cohorts (six from Scandinavian countries, two from Germany, and three from Italy) were included in an ESCAPE meta-analysis published in 2014 by Cesaroni et al(Cesaroni et al., 2014). In this ESCAPE study, annual mean $PM_{2.5}$ per 5 $\mu g/m^3$ was not associated with increased risk for incident acute coronary events (ICD10: I20.0, I21, I23, I24; HR: 1.13, 95%CI: 0.98 to 1.30). Although all incident cases of ischaemic heart disease (ICD10: I20-I25) were included in this PhD analysis, I observed a similar effect size but a wider 95% confidence interval for ESCAPE LUR modelled $PM_{2.5}$ (HR: 1.14, 95%CI: 0.84 to 1.55, per 5$\mu g/m^3$ increase) among the pooled analysis of both UK cohorts. I also restricted the studied outcome to acute coronary events in EPIC-Oxford and UK Biobank as defined by Cesaroni et al in the ESCAPE study, and observed a slightly higher effect estimate (HR: 1.18, 95%CI: 0.68 to 2.06, per 5$\mu g/m^3$ increase of $PM_{2.5}$).

In the ESCAPE study, while the pooled effect estimate for $PM_{2.5}$ on incident acute coronary events was not statistically significant, a significant pooled estimate was observed for $PM_{10}$ (HR: 1.12, 95%CI: 1.01 to 1.25, per 10$\mu g/m^3$ increase)(Cesaroni et al., 2014). Additionally for subgroup analyses, the ESCAPE study also reported statistically significant positive associations for both $PM_{2.5}$ and $PM_{10}$ when restricting the study sample to those with an air pollution exposure level under the current European

annual limits ($PM_{2.5}$: 25 μg/$m^3$, $PM_{10}$: 40μg/$m^3$)(European Union, ). The association with $NO_2$ was non-signficant (HR: 1.03, 95%CI: 0.97 to 1.08, per 10μg/$m^3$ increase) in this ESCAPE study.

In contrast to these findings on incident acute coronary events, another ESCAPE study looking at CVD mortality involving 22 cohorts, including all 11 cohorts of Cesaroni et al's incidence study, found no associations between either PM or $NO_2$ and mortality from ischaemic heart disease (ICD10: I20-I25) or myocardial infarction (ICD10: I21, I22)(Beelen et al., 2014b). The authors speculated that favourable changes in cardiovascular risk factors may have occurred in the past few decades (quit smoking, better medication etc.), which would result in a lower mortality rate for cardiovascular diseases over time in Europe, possibly masking the association between air pollution and cardiovascular mortality.

In Skåne county of southern Sweden, 13,512 participants of 18-80 years were followed to first incident myocardial infarction (ICD10: I20-I23) during 2000-2010(Bodin et al., 2016). The study reported a moderate non-significant association with NOx (Incidence rate ratio: 1.02, 95%CI: 0.86 to 1.21, per 10 μg/$m^3$ increase), which is similar to what I found for both UK cohorts. Further adjusting for road traffic noise in this Swedish study did not change this association. The statistical power of the Swedish study was in part limited by the fact that most participants were exposed to a relatively low level of $NO_X$ (mean annual NOx (5-95 percentile) was 13 (6-33) μg/$m^3$ at baseline in 2000, and 9 (5-21) μg/$m^3$ at the end of follow-up in 2010), compared to other parts of Europe, and lower than that in my UK cohorts. Another Swedish study of 7,494 men in the city of Gothenburg had a much longer follow-up period (1973-2007) for incident myocardial infarction (ICD10: I21). This study reported null associations between residential NOx exposure and incident myocardial infarction, with all hazard ratios close to unity(Stockfelt et al., 2015). The study did not observe substantial changes in the effect estimates when using different time windows (last year, last five years, since enrolment) for $NO_X$ exposure. However, this Gothenburg study reported a borderline significant association between per 10 μg/$m^3$ increase NOx since enrolment and mortality from ischaemic heart disease (HR: 1.02, 95%CI: 0.99 to 1.05).

Atkinson and colleagues previously investigated this topic in a very large sample of 836,557 patients in England(Atkinson et al., 2013) using health data from a primary care database. To assign exposure, postcode of residence was intersected with a 1km grid air pollution exposure surface. A non-significant 1% increased risk for myocardial infarction (ICD10: I21-I23) per 3 $\mu g/m^3$ increase of $PM_{10}$ (HR: 1.01, 95%CI: 0.98 to 1.05) and a similar effect size per 10.7 $\mu g/m^3$ increase of $NO_2$ (HR: 1.02, 95%CI: 0.98 to 1.07) were observed. These estimates are comparable to those I have observed for the two UK cohorts in this PhD work.

In Eindhoven of the Netherlands, de Kluizenaar et al reported a higher effect estimate per 5th to 95th percentile interval increase for $NO_2$ (HR: 1.12, 95%CI: 0.96 to 1.32, per 14.1$\mu g/m^3$) than for $PM_{10}$ (HR: 1.04, 95%CI: 0.90 to 1.21, per 3.1 $\mu g/m^3$) in relation to incident ischaemic heart disease or cerebrovascular disease among 18,213 participants(de Kluizenaar et al., 2013). These effect estimates however cannot be compared directly to others as they do not refer specifically to the incident ischaemic heart disease alone.

A study in Rome found that the association was stronger between each 10 $\mu g/m^3$ increase $NO_2$ and fatal coronary events (Rate ratio: 1.07, 95%CI: 1.02 to 1.12) than non-fatal events (Rate ratio: 1.01, 95%CI: 0.97 to 1.05)(Rosenlund et al., 2008). Combining data from fatal and non-fatal events, per 10 $\mu g/m^3$ increase $NO_2$ was significantly associated with 3% increased risk for incident coronary events (Rate ratio: 1.03, 95%CI: 1.00 to 1.07). It should be noted that participants in this Rome study were enrolled and followed using population registries. As a result, Information on some individual-level confounders such as smoking habit were not available for adjustments in the analyses.

Effect estimates that I observed in this PhD work fell within the range of estimates from other European studies, most of which did not find a statistically significant association. I summarise all the European studies on incident IHD including this PhD work, in Table 6.10. I did not include those European studies on IHD mortality because of the relatively large amount of literature available and

which has already been summarised in two recently published reviews(Faustini, Rapp & Forastiere, 2014; Hoek et al., 2013).

As with incident cerebrovascular disease, I summarised findings by pooling hazard ratios of $NO_2$ effects from published European studies on incident ischaemic heart disease only (Figure 6.8). Overall, among the six studies included, there was a borderline significant association between per 10 $\mu g/m^3$ higher $NO_2$ and incident IHD (HR: 1.01, 95%CI: 1.00 to 1.03).

There are relatively fewer studies investigating PM effects on incident IHD in Europe. The main evidence on this topic to date is based on three women-only cohorts in America.

In the California Teachers Study cohort, Lipsett et al found that there was no association between $PM_{10}$ and myocardial infarction incidence (HR: 0.98, 95%CI: 0.91 to 1.06, per 10 $\mu g/m^3$ increase) but a positive non-significant association with $NO_2$ (HR: 1.05, 95%CI: 0.90 to 1.24, per 10.27 $\mu g/m^3$ increase) among 100,340 women(Lipsett et al., 2011). In the Women's Health Initiative (WHI) Observational Study of 65,893 postmenopausal women, Miller et al reported that a 10 $\mu g/m^3$ increase of $PM_{2.5}$ was associated with 21% (HR: 1.21, 95%CI: 1.04 to 1.42) increased risk for first coronary heart disease(Miller et al., 2007). However, a significant association was not seen for first myocardial infarction (HR: 1.06, 95%CI: 0.85 to 1.34). Both these studies assigned air pollution estimates from the nearest fixed monitoring sites based on participant's residence.

In the Nurses' Health Study of 66,250 participants, Puett et al observed that a 10 $\mu g/m^3$ increase of modelled $PM_{2.5}$ was non-significantly associated with 11% (HR: 1.11, 95%CI: 0.79 to 1.55) increased risk for first coronary heart disease(Puett et al., 2008). Two recent analyses from this cohort were published. One suggested that moving to an area with a higher level of $NO_2$ may increase the incident risk for incident myocardial infarction(Hart et al., 2013). For each 1 ppb increase of $NO_2$, compared to levels at previous address, there was a 22% increased risk for incident myocardial infarction (HR: 1.22, 95%CI: 0.99 to 1.50). The other study found that while $PM_{10}$ effects on incident coronary heart disease

were non-significant in the whole study sample, statistically significant associations were seen among diabetics (HR: 1.12, 95%CI: 1.02 to 1.23, per 10 μg/$m^3$ increase)(Hart et al., 2015).

For these three women-only cohorts in America, there was a consistently stronger association between air pollution and mortality from IHD than that seen for incident IHD.

A study in Vancouver, using data from health insurance administrative database, reported that an IQR increase of black carbon was significantly associated with a 3% increased risk for coronary heart disease (HR: 1.03, 95%CI: 1.01 to 1.05)(Gan et al., 2011). This association was independent of adjustment for $PM_{2.5}$ and $NO_2$. However, neither $PM_{2.5}$ nor $NO_2$ was associated with incident coronary heart disease in the study.

### General issues

Some issues which are specifically related to this PhD analysis on the link between air pollution-incident IHD should be noted.

First, although my findings from these three large European cohorts were consistent with those previously reported, these associations were all null or not statistically significant, suggesting a general lack of evidence on the role of long-term air pollution in incident IHD. In this PhD work, I combined incident data from fatal and non-fatal IHD events and did not specifically study mortality from IHD. Most current evidence suggests a stronger association with IHD mortality than with morbidity. Two recent reviews reported a significant pooled effect estimate for both $PM_{2.5}$ (Hoek et al., 2013)and $NO_2$ (Faustini, Rapp & Forastiere, 2014) on cardiovascular mortality. This may be plausible as it was suggested that air pollution may not necessarily initiate the response but indeed affects the severity of the response(Faustini, Rapp & Forastiere, 2014; Hoek et al., 2013). For example, subjects who died from ischaemic heart disease may have suffered a severe response (e.g. severe ischemia) to a high level of air pollution exposure over a short time scale.

Second, in HUNT2 I found an unexpected beneficial direction of effect of long-term air pollution on incident IHD. I don't have a clear reason to explain this, and suggest this finding should be interpreted with caution.

Third, I found that for both $PM_{10}$ and $NO_2$, stronger suggestive (although still non-significant) associations were seen among women, those aged<60 years, never-smokers, diabetics, those with a BMI greater than 30 $kg/m^2$ and those with a higher level of education (all p-values for interaction<0.05). Unlike many previous studies limited by study sample size, I was able to investigate effect modifications in this very large sample by including interaction terms in the models. Current evidence is mixed on susceptible groups in relation to air pollution effects on IHD, and it remains difficult to compare different susceptibility factors across different study populations.

Perhaps more consistent with previous studies, stronger associations between air pollution and incident IHD were found among never-smokers and diabetics. In this PhD study, for never-smokers, there was a borderline significant association between $PM_{10}$ and incident IHD (HR: 1.13, 95%CI: 0.99 to 1.29, per 4.1 $\mu g/m^3$ increase) compared to ex-smokers (HR: 1.00, 95%CI: 0.87 to 1.15) and current-smokers (HR: 0.82, 95%CI: 0.68 to 0.98). It is possible that stronger associations with incident IHD were found as there may have been less 'noise' from tobacco smoking on the air pollution effects. This was not the case for incident cerebrovascular disease, as I found stronger associations in ex-smokers and current smokers in the same populations. For diabetics, it has been suggested that increased systemic inflammation found in diabetics may facilitate the role of air pollution on cardiovascular diseases(Hart et al., 2015). Indeed, I found a consistently stronger association among diabetics for both incident cerebrovascular and ischaemic heart disease.

Fourth, this PhD analysis provides further evidence that long-term air pollution effects on incident IHD were not affected by road traffic noise.

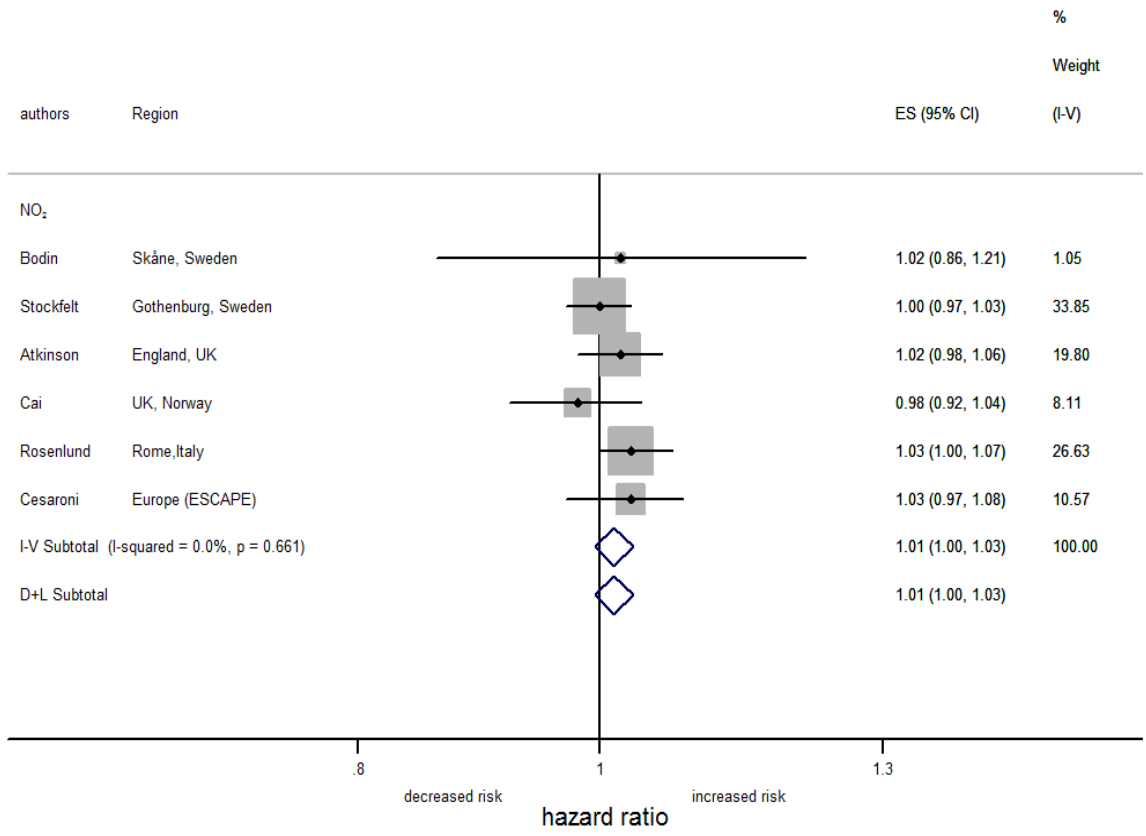
Taken together, in this pooled analysis of EPIC-Oxford, HUNT2 and UK Biobank, there is weak evidence for an association between long-term PM or $NO_2$ air pollution exposure and incident IHD. Nevertheless, effect estimates that I observed in this PhD work fell within the range of estimates from other European studies. Pooling hazard ratios for incident IHD and $NO_2$ from all European studies, including the one from this PhD, yielded a possible 1% increased risk for incident IHD (95%CI: 0 to 3%) among European populations. Fewer studies in Europe have investigated PM effects on incident IHD. Evidence from the ESCAPE study suggested the association was statistically significant for $PM_{10}$ but not for $PM_{2.5}$, but the effect estimate was higher for $PM_{2.5}$ for each 10 $\mu g/m^3$ increase. This again suggests that cardiovascular effects of $PM_{2.5}$ to humans may be more harmful than larger particles, presumably as these are able to penetrate deepest into the lung and potentially into the systemic circulation.

Some other general issues including possible biological mechanisms, exposure models, exposure time windows used in the analyses, loss-to-follow-up, covariate adjustments and statistical approaches are also shared by the traffic noise-CVD analyses, therefore I address these together at the end of the discussion (6.2.3 and 6.2.4).

Table 6.10 a summary of European (north to south) studies on long-term air pollution effects (scaled as per 10 µg/$m^3$ increase) on incident ischaemic heart disease.

| ***Authors*** | ***Published year*** | ***Region*** | ***Study type*** | ***Study entry*** | ***Study end*** | ***Incidence rate (cases/person-year )*** | ***ICD code*** | ***Mean age or age range, years*** | ***$NO_2$ (or NOx)*** | ***$PM_{10}$*** | ***$PM_{2.5}$*** | ***Independent of noise*** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Bodin** | *2016* | *Skåne, Sweden* | *cohort* | *2000* | *2010* | *4.7/1000* | *ICD10: I20-23* | *49* | *1.02 (0.86-1.21)* | *n/a* | *n/a* | *yes* |
| **Stockfelt** | *2015* | *Gothenburg Sweden* | *cohort* | *1973* | *2007* | *10.8/1000* | *ICD10:I21* | *53* | *1.00 (0.97-1.03)* | *n/a* | *n/a* | *n/a* |
| **Atkinson** | *2013* | *England, UK* | *census-based* | *2003* | *2007* | *13,956 cases/810,686 subjects* | *ICD10: I20-23* | *40-89* | *1.02 (0.98-1.06)* | *1.03 (0.93-1.18)* | *n/a* | *n/a* |
| **Cai [this PhD]** | ***2016*** | ***UK, Norway*** | ***cohort*** | ***1993-2010*** | ***2012-2015*** | ***3.2/1000*** | ***ICD10: I20-I25*** | ***55*** | ***0.98 (0.92-1.04)*** | ***0.93 (0.75-1.15)*** | ***1.35 (0.73-2.39)*** | ***yes*** |
| **Rosenlund** | *2008* | *Rome,Italy* | *census-based* | *1998* | *2000* | *11167 cases* | *ICD9: 410-414* | *35-84* | *1.03 (1.00-1.07)* | *n/a* | *n/a* | *n/a* |
| **Cesaroni** | *2014* | *Europe (ESCAPE)* | *cohort (11 cohorts)* | *1997* | *2007* | *4.5/1000* | *ICD10: I20.0, I21, I23, I24)* | *44-74* | *1.03 (0.97-1.08)* | *1.12 (1.01-1.25)* | *1.28 (0.96-1.29)* | *yes* |

Figure 6.8 Meta-analysis of European studies on long-term effects of per 10 µg/$m^3$ increase $NO_2$ (or NOx) on incident ischaemic heart diseases




I-squared, variation in estimated effect attributable to heterogeneity; I-V, inverse-variance weighted fixed effects method; D-L, DerSimonian-Laird random effects method

### 6.2.2. Road traffic noise effects

#### *Noise and Cerebrovascular disease*

In this PhD work, when assessing road traffic noise on a continuous scale, I did not observe an association between daytime road traffic noise and incident cerebrovascular disease in the pooled analysis of the three cohorts. In the cohort-specific analysis, a positive but non-significant association was observed in UK Biobank (HR: 1.04, 95%CI: 0.93 to 1.17, per 3.5 dB(A)) whilst null associations were observed in EPIC-Oxford and HUNT2.

However, when assessing daytime road traffic noise in categories, there was suggestive evidence that participants who were exposed to a higher level of daytime road traffic noise had a higher risk for incident cerebrovascular diseases. Compared to those exposed to a level of less than 50 dB(A) daytime road traffic noise, there was a 2% (95%CI: -12% to 20%) and a 4% (95%CI: -16% to 29%) increased risk for those exposed to a level of 50-55 dB(A) and greater than 55 dB(A) respectively. No such suggestive positive relationship was seen for night-time noise.

Findings from this PhD work contribute to the scant literature on this topic. To my best knowledge, only one prospective cohort study has investigated this association(Sorensen et al., 2011). This study, conducted by Sorensen et al, reported that a 10 dB(A) increase of road traffic noise (Lden) was significantly associated with a 14% (HR: 1.14, 95%CI: 1.03 to 1.25) increased risk for incident stroke among 57,053 participants from the Danish Diet, Cancer and Health (DCH) cohort study. In fact, I found a similar effect estimate per 10 dB(A) increase of daytime road traffic noise in the UK Biobank participants, although the association was not statistically significant (HR: 1.12, 95%CI: 0.81 to 1.57). Some important points were noted for this Danish study. First, the effect estimate for Lden in the study was robust to adjustments for noise from railway and aircraft, as well as road traffic air pollution indicated by NOx. In my PhD study, I did not adjust for railway and air traffic noise. Second, the effect estimate was also independent from adjustment for hypertension, one of the most important risk factors for stroke. I did not adjust for blood pressure directly in my analyses, however all participants with previous known cardiovascular conditions including hypertension were excluded from the

analyses. In another analysis involving both EPIC-Oxford and HUNT conducted by BioSHaRE colleagues (personal communications from Wilma Zijlema), there was no overall significant association between road traffic noise and blood pressure. Third, the Danish study reported that more recent noise exposure had a stronger association with incident stroke than did noise exposure in earlier years. Fourth, the study also reported stronger associations among those aged 64.5 years and above. The authors suggested that this may be due to the noise-induced sleep disturbance in elderly people, for whom the sleep structure becomes more fragmented(Jackson, Redline & Emmons, 2015a). Another possible reason may be that exposure misclassifications for noise were reduced among older people as this age group is likely to have been at the current residence for years if not decades, and is likely to spend most of their time at home as they retired. However, in my pooled analysis, I did not observe a stronger association among those aged 60 years and above.

In 2015, Halonen et al published a small-area study, which investigated long-term road traffic noise effects on hospital admissions and mortality from cardiovascular diseases in a population of 8.6 million Londoners(Halonen et al., 2015). This study, although ecological in study design, reported that associations between road traffic noise and hospital admissions for stroke were strongest among all the studied cardiovascular outcomes. Among adults (≥25 years), compared to those living in areas with a level of mean daytime road traffic noise <55 dB(A), a relative risk for hospital admissions for stroke was 1.05 (95%CI: 1.02 to 1.09) for those living in areas with a level >60 dB(A). The corresponding figure for the elderly (≥75 years) was 1.09 (95%CI: 1.04 to 1.14). A similar significant association was also found for night-time traffic noise and stroke hospital admission among the elderly. As with my analyses, this small-area study did not find a significant association between noise and hospital admissions for stroke when noise was assessed as a continuous variable.

Nevertheless, findings from Halonen et al were consistent with those reported in the Danish study. As with most other ecological studies, one major limitation of this London study is that most individual-level confounders and exposure modifiers were not available. The effect estimates from the London

study were adjusted for age, sex, area-level deprivation, area-level lung cancer mortality as a proxy for smoking, area-level ethnicity and area-level $PM_{2.5}$.

This PhD work is by far the largest cohort study with detailed individual-level information on some key confounders including smoking, alcohol consumption and education. As with Halonen et al, I did not observe associations between noise exposures and cerebrovascular disease in the linear model. Halonen et al suggested that this may in part be due to the fact that noise exposure misclassifications were greater in areas with a low noise level and in areas with minor heavily trafficked roads, as the noise model tended to over-estimate and underestimate exposures respectively. This would contribute to the uncertainty of continuous noise estimates and impact statistical power to detect a noise effect.

The categorical analysis provided a way to investigate a possible non-linear relationship and also a chance to identify a possible effect threshold. In my analyses, using categorical daytime, but not night-time, road traffic noise estimates, small non-significant increased risks for incident cerebrovascular disease was observed at a level above 50 dB(A). This is consistent with Halonen et al findings although we used different noise categories.

One of the suspected reasons that I did not see a corresponding relationship for night-time noise may be related to the noise model used in this PhD study, which was likely to over-estimate noise exposures for those at low exposure levels, and this misclassification may have been amplified particularly for night-time noise estimates. In Halonen et al's small-area study in London, effects of night-time noise on hospital admissions for stroke were not statistically significant but daytime noise was(Halonen et al., 2015). The Danish study however did not specifically study night-time noise effects(Sorensen et al., 2011).

Different noise categories were used in previous studies. In my analyses, I used 50 dB(A) as a reference for daytime noise because the noise model used in this PhD study predicted relatively less well below

this. More studies are needed to identify the possible effect threshold. Respectively for night-time noise, 40 dB(A) was chosen as the reference level according to the WHO night-time noise guidelines for Europe(World Health Organisation Europe, 2009), in which 40 dB(A) is recommended as equivalent to the lowest observed adverse effects level for night-time noise.

UK Biobank is the largest cohort included in my analyses. I found a positive non-significant association between daytime road traffic noise and incident cerebrovascular diseases in this large cohort. However, the average person-years of follow-up for UK Biobank was only 1.4 person-years. During the follow-up period, 240 incident cerebrovascular cases were identified in UK Biobank. Future studies in this cohort with more years of follow-up will be needed to replicate my findings in this PhD. The other two cohorts each with a follow-up period of nearly 20 years found null associations.

There are two other studies investigating stroke incidence in combination with other heart diseases. In a small cross-sectional study involving 4,712 participants from six European countries, Floud et al reported that 24hr average road traffic noise was associated with 'heart disease and stroke' (OR: 1.19, 95%CI: 1.00 to 1.41)(Floud et al., 2013). However, a subsample analysis suggested that this association may be confounded by air pollution. In the GLOBE study of 18,213 participants in the Netherlands, no association was found between Lden and 'IHD or cerebrovascular disease' either in the total sample or in the subgroup of those aged >65 years(de Kluizenaar et al., 2013).

A small-area study investigated relative risks of hospital admissions for stroke in the areas around London Heathrow Airport(Hansell et al., 2013). In that study, the relative risk of hospital admissions for stroke was 1.24 (95%CI: 1.08 to 1.43) in areas with the highest daytime aircraft noise levels (>63dB), compared to those in the lowest levels (≤51dB). This association was robust to adjustment for road traffic noise and particulate air pollution. In America, a study involving only elderly people (aged ≥65 years) who lived near one of the 89 airports across the country, also found a positive association between aircraft noise and hospital admission for stroke(Correia et al., 2013). A recent meta-analysis found an association between aircraft noise exposure and hypertension(Huang et al.,

2015), but heterogeneity between studies meant that definitive exposure-response relationships could not be established. Of note, aircraft noise is qualitatively different from road traffic noise and may have different associations with health outcomes.

In this PhD work, I found a small but non-significant increased risk for incident cerebrovascular diseases in men (HR: 1.004, 95%CI: 0.987 to 1.020) compared to women (HR: 0.995, 95%CI: 0.978 to 1.012), which was consistent with the findings in the Danish study(Sorensen et al., 2011). Additionally, I also found a stronger association among diabetics (HR: 1.010, 95%CI: 0.947 to 1.078) compared to non-diabetics (HR: 0.999, 95%CI: 0.987 to 1.011).

It should be noted that, in this PhD work, the effect estimate for continuous day-time noise on incident cerebrovascular disease may have been underestimated as I included all the incident cases that fall within ICD10: I60-I69. One study found that the effect estimate for incident ischaemic stroke was higher than that of all incident stroke(Sorensen et al., 2014).

I further repeated the main analyses separately for ischaemic stroke (ICD10: I63), haemorrhagic stroke (ICD10: I60, I61, I62) and unspecified stroke (ICD10: I64). For each 10 dB(A) increase of daytime road traffic noise, slightly positive associations were found for both ischaemic stroke (HR: 1.025, 95%CI: 0.862 to 1.219) and haemorrhagic stroke (HR: 1.060, 95%CI: 0.808 to 1.391). A null association was found for unspecified stroke. The higher effect estimate observed for haemorrhagic stroke was biologically plausible as a review published in 2012 showed that road traffic noise has been linked to hypertension(van Kempen & Babisch, 2012), a major risk factor for haemorrhagic stroke. In contrast to my findings, Sorensen et al reported a significant association between Lden and incident ischaemic stroke (HR: 1.16, 95%CI: 1.07 to 1.24) but not haemorrhagic stroke (HR: 0.99, 95%CI: 0.81 to 1.20) in the Danish DCH cohort(Sorensen et al., 2014).

In summary, results from this PhD work contribute to the scant literature on the possible role of road traffic noise on incident cerebrovascular disease. There were suggestive effects in those exposed to a

daytime road traffic noise level greater than 50 dB(A) in the pooled analysis of the three cohorts. I did not find a confounding effect by air pollution in my study. Since only two cohort studies including this current PhD work are available, more studies in other populations are warranted to better clarify the role of road traffic noise on incident cerebrovascular diseases.

### *Noise and Ischaemic heart disease*

In the pooled analysis of EPIC-Oxford, HUNT2 and UK Biobank, I found a weak, non-significant, increased risk for incident ischaemic heart disease (IHD) in relation to annual mean daytime road traffic noise exposure (HR: 1.01, 95%CI: 0.98 to 1.04, per 3.9 dB(A)) in the main model. Compared with those exposed to a level of daytime road traffic noise less than 50 dB(A), a higher non-significant risk for incident IHD was seen for those exposed to a level of 50-55 dB(A). Similarly for night-time noise, higher risks were seen for those exposed to levels of 40-45 dB(A) and greater than 45 dB(A), compared with those exposed to a level less than 40 dB(A).

A large number of studies had been undertaken in the last two decades to investigate this road traffic noise-IHD link in different populations. Most of these studies were summarised in three important reviews to date. A review published by van Kempen et al in 2002 was the first to systematically examine epidemiological evidence of road traffic noise on IHD(van Kempen et al., 2002). Among the identified 43 epidemiological studies published in 1970-1999, 28 were occupational studies and 15 were population studies. Only two of those 15 population studies had studied the outcomes of IHD: the Berlin case-control study and the Caerphilly-Speedwell study. A meta-analysis based on these two studies showed that road traffic noise was positively but not significantly associated with myocardial infarction prevalence. However, for total IHD, road traffic noise was significantly associated with prevalence (pooled estimate: 1.09, 95%CI: 1.05 to 1.13, per 5 dB(A)) but not with incidence. It should be noted that these earlier studies included in van Kempen's review were mainly cross-sectional, most of these had limited noise exposures assessments, and some did not adjust for important confounders such as smoking. Evidence of noise effects on incident IHD was very limited before year 2000. Indeed,

in the review the authors called for more large follow-up studies with better noise exposure assessment approaches to investigate this important link.

In the years after 2000, six studies from Germany, the Netherlands, Sweden, Denmark and Canada reported findings on the associations between road traffic noise and incident IHD. Two of these six studies investigated IHD mortality, three investigated incident myocardial infarction and one investigated self-reported CVD prevalence. All six studies, together with the Berlin and Caerphilly-Speedwell studies were included in an updated meta-analysis in 2014, published by Babisch(Babisch, 2014). In this meta-analysis, although different noise indicators and IHD outcomes were assessed across the included studies, a 10dB(A) increase of day-night noise level was associated with a 8% increased risk (95%CI: 4% to 13%) for IHD. One important aspect of this updated review was that this pooled effect estimate was based on studies with noise exposures ranging from 52-77 dB(A). Many earlier studies did not assess noise effects across such a wide exposure range. In addition, studies emerging after the year 2000 were generally population-based, had better noise exposures assessments and included adjustments for established risk factors for CVD. Only four studies in this review had adjusted for air pollution effects. A meta-analysis based on these four studies showed a similar pooled estimate (1.10, 95%CI: 1.02 to 1.09) for a 10 dB(A) increase of noise on IHD, suggesting a possible independent effect of noise.

Although the Babisch review in 2014 specifically focused on road traffic noise effects, it included studies on both prevalence and incidence. More recently in 2015, Vienneau et al published another review on incident studies only(Vienneau et al., 2015). This review, including 10 incident studies on both road and aircraft noise, reported that a 10 dB(A) increase of Lden was significantly associated with an increased risk for IHD (pooled estimate: 1.06, 95%CI: 1.03 to 1.09). It also suggested that this dose-response relationship started at 50 dB(A). The authors further considered the road traffic noise effects among the eight studies which measured this, and reported respective pooled estimate of 1.04

(95%CI: 1.00 to 1.10). My findings from the linear model (HR: 1.03, 95%CI: 0.95 to 1.12, per 10 dB(A)) are consistent with their pooled estimate.

Two additional studies have been published since Vienneau's review was published in 2015. In Sweden, a prospective cohort study conducted by Bodin et al found that current or medium-term exposure to Lden was not associated with incident myocardial infarction during the follow-up period of 2000-2010 among 13,512 participants (HR:0.99, 95%CI: 0.86 to 1.14, per 10 dB(A)). However, when Lden was categorised in the analyses, higher effect estimates were seen among those with higher noise exposure levels. For example, risk for incident myocardial infarction increased by 13% (95%CI: -13% to 47%) and 12% (95%CI: -16% to 51%) respectively for those exposed to a level of 45-55 dB(A) and 55-65 dB(A) Lden, compared to the reference level of less than 45 dB(A), although a null association was found in the highest noise exposure group (65-80 dB(A)). A small-area study in London generally found no associations between either daytime or night-time road traffic noise and hospital admissions for IHD(Halonen et al., 2015). However, small increased risks for mortality from IHD were found in areas with a level of daytime road traffic noise (55-60 dB(A)), compared to areas with a level of less than 55 dB(A), among both adults aged ≥25 years (relative risk: 1.03, 95%CI: 1.00 to 1.06) and elderly aged ≥75 years (relative risk: 1.04, 95%CI: 1.01 to 1.07).

Previous studies were limited by sample size when investigating interaction effects, and the results were rather mixed. In the subgroup analyses of this large study, I found that the effects of daytime road traffic noise on incident IHD were higher among females, those aged <60 years, those with a BMI greater than 30 kg/$m^2$ and current-smokers. Although the interaction terms for these analyses were all statistically significant, the 95% confidence interval in each strata was overlapping. Moreover, only the association among current-smokers reached significance (HR: 1.018, 95%CI: 1.003 to 1.033).

My findings of stronger associations in females and those aged less than 60 years are somewhat contradictory to previous findings, but are plausible as women and those participants aged <60 years may be more sensitive to noise exposure. An analysis from the Whitehall II study among 3,630 civil

servants in the UK showed that women and those aged 50-55 years had higher noise sensitivity(Stansfeld & Shipley, 2015). However, in this PhD study, noise sensitivity data were not available to investigate this directly.

As with my air pollution analyses, I further investigated the role of daytime road traffic noise on acute coronary heart events (ICD10: I20.0, I21, I23, I24). A similar finding to the main analysis was observed. For each 10 dB(A) increase of daytime road traffic noise, incident acute coronary heart events increased by 4% (HR: 1.04, 95%CI: 0.91 to 1.91).

One finding from this PhD work is that I found an increased (although non-significant) risk for incident IHD or total CVD events among those exposed to the range between 50-55 dB(A) of daytime road traffic noise. In Europe, a large number of general populations may be exposed to a noise level within the range of 50-55dB(A) according to a report from European Environment Agency (http://www.eea.europa.eu/publications/noise-in-europe-2014, accessed April 2016). Although only small excess risks were found, findings suggest that exposure to road traffic noise between 50-55 dB(A) may potentially have a substantial impact on the burden of cardiovascular disease.

Besides the effect of day-time road traffic noise, I also found that higher night-time noise exposure levels were non-significantly associated with higher risks for incident IHD. However, since the correlation between daytime and night-time noise was close to unity, it is difficult to disentangle the independent effects of each. Nevertheless, night-time noise is of particular potential importance as it may disrupt sleep, and poor sleep quality has been associated with incident myocardial infarction and other cardiovascular diseases(Jackson, Redline & Emmons, 2015b).

In summary, this PhD work provides further evidence of the role of road traffic noise on incident ischaemic heart disease. Findings were based on three large European cohorts and are in line with previous studies mostly undertaken in Europe, and although almost all were not statistically significant they report similar effect sizes.

### 6.2.3. Possible biological mechanisms

Possible biological mechanisms linking air pollution and CVD outcomes have been well documented. Three major biological mechanisms hypothesised for air pollution effects include increased systemic inflammation, activation of the autonomic nervous system (ANS), and a direct transfer of fine particles from lung to blood circulation(Brook et al., 2010). In recent years, increasing studies also showed that exposure to air pollution can lead to epigenetic modifications of DNA, which results in either up-regulation or down-regulation of certain gene expressions in relation to cardiovascular diseases(Chin, 2015). Currently, the best-supported hypothesised mechanism is the inflammatory response in relation to sub-acute and chronic air pollution exposure(Chin, 2015). Many animal and human studies have suggested that inhaled air pollutants can provoke a local inflammatory response in lung, which promotes oxidative stress and the release of pro-inflammatory mediators such as cytokines from lung-based cells. These pro-inflammatory mediators then spill over from lung into the blood circulation, leading to systemic oxidative stress and inflammation(Chin, 2015; Brook et al., 2010). As a result, a range of physiological changes in metabolism and blood circulation becomes biologically plausible. This includes, but is not limited to, insulin-resistance, dyslipidaemia, impaired HDL function, increased coagulation, thrombosis, and decreased fibrinolysis. In Chapter 5, I have demonstrated that long-term exposure to air pollution is significantly associated with elevated levels of hsCRP, HDL cholesterol, triglycerides and blood glucose, which lends further support to these proposed biological mechanisms and their implications for manifest CVD outcomes.

As briefly mentioned in the first Chapter, it was proposed that environmental noise exert its adverse cardiovascular effects via both direct and indirect pathways as a general stressor(Babisch, 2014; Munzel et al., 2014). The direct pathway mainly involves noise-induced sleep disturbance, whilst the indirect pathways may include cognitive and emotional responses such as annoyance. To cope with this 'stressful' situation, two classic reactions will be activated, as described by Henry and Stephens in the 1970s in thier general psychophysiological stress reaction model(Henry, 1992). One is the 'fight-flight' reaction, in which adrenalin will be released to actively 'fight' against the stressor in order to

remove it or help the organism to 'flee' from this stressor. The other is the 'defeat' reaction, which means the 'fight' was not successful or the organism did not manage to 'flee', when cortisol is released to help organism mitigate the damages that the stressor causes. Recently, an extended version of this stress model was proposed by Recio et al(Recio et al., 2016). In this updated stress model, Recio et al describe comprehensively the biological mechanisms of road traffic noise effects on cardiovascular outcomes. In their proposal, if the 'defeat reaction' persists, in other words, if the stress is sustained, then it will lead to a state that Recio et al termed as 'emotional flight'. In general, emotional flight involves a change of stress processing from the psychological level to the physiological level; the latter is processed by the central autonomic network, for example, by hypothalamus. In response to this sustained stress (e.g. chronic exposure to road traffic noise), the physiological processing usually leads to an excess allostatic load, sometimes referred to as imbalanced homoeostasis. In return, this imbalanced homoeostasis will lead to a range of physiological changes including to blood pressure, blood glucose, blood lipids and thrombosis. In addition, according to Recio's review, noise as a stressor can also cause other defensive responses, such as oxidative stress, immune system activation and systemic inflammation. Taken together, these physiological changes will ultimately lead to manifest CVD outcomes such as hypertension, ischaemic heart disease or stroke in the long term.

In the biochemistry analyses described in Chapter 5, I observed some significant positive associations between daytime road traffic noise and hsCRP, triglycerides in the pooled analyses of the HUNT3 and LifeLines cohorts and fasting glucose in LifeLines analysis only. In particular for blood glucose, the association was robust to further adjustment for air pollution.

Another important risk factor for CVD is hypertension. A quantitative review of 24 studies up to year 2012 was published by van Kempen and Babisch, which reported a statistically significant dose-response relationship between road traffic noise and hypertension(van Kempen & Babisch, 2012). These findings, from this PhD work and others, suggested positive associations between road traffic

noise and intermediate risk factors of CVD, which provide novel insights into the links between road traffic noise and CVD outcomes.

Indeed, some mechanisms and their underlying physiological changes in relation to CVD outcomes are shared by both air pollution and road traffic noise independently, as I have demonstrated in Chapter 5. More recently, in the German Heinz Nixdorf Recall study of 4238 participants, both $PM_{2.5}$ and night-time traffic noise were both independently associated with thoracic aortic calcification, a measure of atherosclerosis(Kalsch et al., 2014). Although the statistical evidence was not particularly robust (P~0.05) for traffic noise, this important study complements previous studies using a proxy of noise exposure (e.g. traffic intensity exposure) and provides new evidence that long-term exposure to traffic-related fine particles and noise may act independently in the development of CVD outcomes.

### 6.2.4. Limitations

Some limitations of this PhD work are acknowledged.

First, modelled air pollution and noise estimates at home address will inevitably have some misclassification (e.g. due to time spent away from home, travel outside the house during the day, exposure modification via housing characteristics and window-opening etc.). Performance for the pan-European LUR air pollution model was moderate, compared to the ESCAPE-LUR model and the ESCAPE model itself was based on a limited measuring campaign. Effect estimates may have been underestimated when using the pan-European air pollution metrics because of the greater potential exposure misclassification. In addition, for the noise model, some simplified input variables were used to enable a harmonised approach across the cohorts investigated, and this may have resulted in non-differential misclassification of noise exposure that would be expected to bias results toward null. Further, traffic flow data were not available as inputs into the model for some secondary roads, which may result in an underestimation of noise exposure in these areas and particular misclassifications at lower noise levels. Nevertheless, given the broad geographic regions that this PhD work covers, common LUR air pollution and noise models were developed for Europe to minimise differences

between cohorts that would otherwise be introduced by having different exposure assessment methods. In the air pollution analyses, I did not run co-pollutant models because the correlation between $NO_2$ and $PM_{10}$ in this PhD study was high (r>0.7).

Second, I used air pollution and noise estimates from a single year to represent long-term annual average estimates in the incident CVD analyses. Residential air pollution and noise estimates were modelled for the years 2007 and 2009 respectively, which may differ by up to 16 years from when recruitment took place across the three cohorts (1993-2010). This PhD therefore has to rely on the assumptions that air pollution spatial contrast has not changed substantially in the last 10-14 years in the respective countries and that the estimates from the LUR model for year 2007 were representative of the baseline spatial contrast in each cohort. The same limitation was also noted in the ESCAPE studies where air pollution LUR model was built for 2008-2011 and baseline recruitment in most European cohorts took place in the mid-1990s. In the ESCAPE studies, investigators either used back-extrapolated air pollution estimates based on the current LUR modelled estimates or conducted sensitivity analyses among the most recently established cohorts only, but no differences were shown for the main results. Also, some Scandinavian studies with a complete residential history of modelled air pollution estimates demonstrated that there were no clear differences in the effects on incident CVD, using either estimates at baseline recruitment period or in most recent years(Korek et al., 2015; Nafstad et al., 2004). In fact, some recent LUR studies in Europe showed that LUR models based on current air pollution data can predict spatial contrast in the preceding 10 years(Gulliver et al., 2013; Eeftens et al., 2011; Madsen et al., 2011). Similarly, for noise exposures, the assumption had to be made that spatial contrasts in road traffic noise levels will have been relatively stable over the last decade. Another potential source of exposure misclassification may be due to participants changing their residential location over the follow-up periods. Restricting analyses to those who were residing in the same address for more than 10 years showed slightly stronger associations in some analyses, but overall the results were not materially different.

Third, although many efforts were made to harmonise key covariates across cohorts, estimates presented in this PhD work may include residual confounding, as information about variables such as environmental tobacco smoke, diet, physical activity and neighbourhood socioeconomic status was not able to be harmonised, and therefore were not included in the models. Nevertheless, the adjusted variables included in this PhD analysis have been largely comparable to those included in many previous studies.

Fourth, effect estimates may also have been affected by 'healthy survivor' bias, particularly in the incident CVD analyses. Compared to those who lost to follow-up or died prematurely, those who remained in the study were more likely healthy individuals.

Fifth, as I stated in section 2.1 of Chapter 2, by not taking into account of the possible clustering effects in LifeLines, effect estimates yielded from the analyses of CVD biochemistry markers (Chapter 5) may have been biased for this particular cohort. The consequence of ignoring clustering effects in the analyses of clustered data depends on within-cluster correlation(Galbraith, Daniel & Vissel, 2010). Within-cluster correlation, as the case for LifeLines cohort, may possibly be positive, and may cause standard errors to be underestimated when the exposure of interest is fixed for the cluster (i.e. members from the same family/household have same levels of air pollution exposure) and overestimated when the exposure varies within cluster (i.e. members from the same family/household have different levels of air pollution exposure)(Desai & Begg, 2008). Either case is possible for the LifeLines cohort. Whilst this issue was not specifically investigated in this PhD study, there are some statistical approaches based on regression techniques that can address this problem. For example, random effects modelling approach is one of the most common methods that is used to incorporate a random intercept term which accounts for correlation within cluster(Desai & Begg, 2008).

Sixth, I did not consider other factors that may have specifically affected noise estimates in my analyses. Only residential noise data were available in each cohort, and therefore noise exposure in

other places such as at work was not available. However, this is likely to be a non-differential misclassification in both cases and non-cases of incident CVD and therefore bias the estimates towards null. Also, data about behaviours that individuals might adopt to mitigate noise exposures, as well as other personal characteristics, were not available. For example, window opening habits, hearing impairment, noise sensitivity or bedroom location. Estimates may have been biased without considering these potential effect modifiers. In the present PhD analysis, I have considered noise effects only from road traffic, but not from aircraft or railway traffic. By not considering the total traffic noise exposures, estimates may have been underestimated.

Lastly, many analyses have been conducted in this PhD work, it is therefore possible that some observations are simply due to chance.

### 6.2.5. Chapter summary

In summary, in line with previous European studies, my findings suggest a possible effect of both long-term exposure to air pollution and road traffic noise on incident ischaemic heart disease and cerebrovascular disease. These associations were observed in a large study sample from three European cohorts, and were independent of adjustments for covariates. It is further suggested that air pollution and road traffic noise effects on CVD outcomes are likely independent from each other.

# Chapter 7 Conclusion

## 7.1. Overall summary

Findings from this large project, involving four established cohorts in the UK, the Netherlands and Norway, contribute to the current literature and address some important research gaps regarding the contribution of air pollution and noise research to CVD epidemiology.

Pooling data from HUNT3 and LifeLines cohorts, I found that both air pollution and road traffic noise was significantly associated with a marker of systemic inflammation (hsCRP). This is, to my knowledge, the first study to report a positive significant association between road traffic noise and systemic inflammation in a population-based study. This finding suggests that the link between noise and CVD may operate via systemic inflammation, which is also supported by a recent review on the relevant biological mechanisms of road traffic noise effects on CVD(Recio et al., 2016). However, statistical significance for this road traffic noise effect was lost after air pollution was adjusted for, suggesting somewhat confounding effects by air pollution for this specific association.

Another finding from these two cohorts is that both air pollution and noise were significantly and independently associated with HDL cholesterol, but not associated with total cholesterol. Although I have mentioned the possible mechanisms explaining these findings in Chapter 5, more studies on the biological mechanisms are needed, particularly for increased noise and elevated HDL cholesterol levels. I also found that air pollution was significantly associated with triglyceride levels, and effect was independent from road traffic noise. In contrast, the positive effect of road traffic noise on triglyceride levels was confounded by air pollution.

In the LifeLines cohort only, I found that both air pollution and noise were significantly and independently associated with increased fasting blood glucose. These important findings provide some novel evidence not only on the associations between air pollution/noise and CVD outcomes, but also on other chronic health outcomes such as diabetes. However, I did not observe corresponding associations with HbA1c.

Based on these cross-sectional associations between air pollution, noise and CVD biochemistry markers, I further investigated the direct links between air pollution, noise and incident CVD outcomes. Specifically, I investigated incident cerebrovascular disease, ischaemic heart disease and total CVD in relation to air pollution and road traffic noise, by pooling data EPIC-Oxford, HUNT2, and UK Biobank, which are linked to medical records to permit ascertainment of incident cases (fatal and non-fatal) since recruitment.

In general, I found suggestive evidence of an association between long-term air pollution exposure and incident cerebrovascular disease. Only a few European studies have been conducted previously, and the effect estimates from this PhD fell within the range of most of these previous studies. I conducted a meta-analysis of all the available estimates for $NO_2$ and incident cerebrovascular disease from European studies, including this current PhD, and found a borderline significant positive association between $NO_2$ and incident cerebrovascular disease. A meta-analysis of PM effects was recently published and reported a significant positive association(Scheers et al., 2015).

However, I found only very weak evidence of air pollution effects on incident ischaemic heart disease. This association may possibly be driven by the associations seen in HUNT2, as I found unexpected significantly negative associations in this cohort. Whilst for both EPIC-Oxford and UK Biobank, small but non-significant increased risks were observed, in keeping with most previous European studies.

For the $PM_{10}$ effects on total incident CVD, significant and stronger positive associations were found among men, those aged 60 years and over, those with a BMI greater than 30 $kg/m^2$. The large sample size of this PhD study allowed investigations of these interaction effects. These findings in susceptible groups were consistent with some previous studies.

Findings for road traffic noise effects on incident cerebrovascular disease were somewhat contradicted when continuous versus categorical noise estimates were used in the analyses. Null associations were reported with continuous noise estimates, however, in categorical analyses, there

was a tendency for higher daytime, but not night-time, noise exposure to be associated with higher risks for incident cerebrovascular diseases. Despite these inconclusive findings, since only one previous population-based cohort study has been conducted to date, results from this PhD represent important contributions to the current knowledge.

I found a small but non-significantly increased risk for incident ischaemic heart disease in relation to either daytime or night-time noise exposures, which was in line with what previously reported findings.

This PhD study has several strengths. First, this is the largest study on air pollution and noise exposure and cardiovascular risk factors and outcomes to date. Tremendous efforts were made to harmonise data across some of the largest cohorts in Europe, allowing an integrated large sample size and sufficient statistical power to investigate interaction effects.

Second, analyses were conducted using DataSHIELD, a novel 'compute to the data' approach, which could help maximise the scientific potential of established cohorts by pooling personal data robustly yet ethically for research. Furthermore, results from DataSHIELD have been validated here by both cohort-specific meta-analysis and physical pooling of cohort data, which showed a consistency of the reported estimates.

Third, mutual adjustments were made for modelled residential air pollution and noise data in all the analyses, which further clarify the possible independent and joint roles of each on the development of CVD outcomes.

## 7.2. Future research and policy implications

More research is needed on both air pollution and road traffic noise effects on CVD outcomes.

My analyses on CVD biochemistry markers in Chapter 5 were cross-sectional in design, and longitudinal studies investigating changes of these biochemistry markers in relation to both air pollution and noise exposures are required to confirm these findings.

In terms of air pollution-incident CVD analyses, more studies in other populations (e.g. those in developing countries with higher levels of exposures) as well as studies on the possible independent effects of both PM and $NO_2$ are needed to better clarify the current knowledge. Additionally, studies into the sources and components of air pollution mixtures and their associated CVD effects will be extremely informative for policy making.

More attention should be focused on studying traffic noise effects on less studied CVD outcomes (e.g. stroke). A number of noise studies are in progress within Europe at the time of writing this thesis. The NORAH (Noise-Related Annoyance, Cognition, and Health) study in Germany has several manuscripts in preparation, detailing different sources (rail, road, aviation) of noise effects on a range of health outcomes including CVD, blood pressure, quality of life, sleep and children's intellectual development and reading abilities (http://www.laermstudie.de/en/norah-study/overview/, accessed June 2016). Wilma Zijlema, one of the BioSHaRE colleagues, has submited a manuscript on road noise effects on blood pressure and heart rate in a study of three European cohorts. In London, a recently concluded study was conducted in Whitehall 2 and SABRE (the Southhall and Brent Revisited Study) cohorts about road noise and air pollution effects on both blood pressure and carotid intima-media thickness(CIMT) (personal communications from Anna Hansell). Meta-analysis of noise effects on different CVD outcomes will be helpful in future to update those by Vienneau(Vienneau et al., 2015) and Babisch(Babisch, 2014).

Besides CVD outcomes, emerging studies in the last 2 to 3 years showed that other health outcomes such as cognitive performance, depression, obesity and diabetes are also related to both air pollution and noise exposures, which warrant further investigation.

Ambient air pollution not only has a broad range of acute and chronic health impacts, but also has an impact on economic costs associated with diseases and deaths. At the time of writing this thesis, a new WHO study showed that in 2010 approximately 600,000 premature deaths were due to air pollution in the WHO Europe region, which cost the economy 1.6 trillion US dollars a year(WHO, 2015).

More specifically, for the United Kingdom and the Netherlands, this cost accounted for 3.7% and 3.3% of the country's Gross Domestic Product (GDP) in 2010. This PhD study, in line with other European studies on air pollution, again highlights the importance of collective efforts to control the persistent air pollution problem. There is indeed a need to call for health-effective and cost-efficient air pollution control measures for the years and decades to come. For example, future environmental policies will be better informed if the chemical and physical properties of the air pollution mixture are able to be characterised to identify the responsible components that related to certain health effects. Equally important is the need to further clarify different sources of air pollution and their associated health effects. To refine air pollution control policies in the context of a climate change era, it will increasingly require more and more collaborative efforts from various scientific disciplines including epidemiology, toxicology and atmospheric science.

After air pollution, traffic-related noise is the second largest environmental risk factor for health in Europe. As estimated in a WHO 2011 report, in the high-income Western European countries, at least 1 million healthy life years were lost annually due to exposure to traffic noise(European Office WHO, 2011). Among those healthy life years lost, an estimated 61,000 years were accounted for by ischaemic heart disease. Findings from this PhD, in line with just a handful of other studies, provide further scientific evidence to refine risk estimates for policy formulation to mitigate the harmful effects of transport noise on certain CVD outcomes, particularly hypertension, ischaemic heart disease and possibly stroke. Effects of noise on annoyance and on sleep disturbance are well recognised, but for policy formulation special attention is also needed for other potential clinical outcomes including respiratory health, obesity and diabetes. Specific health effects of different sources of noise exposure (aviation, rail, and road) should also be better studied to inform the relevant policy making.

Technology, such as DataSHIELD used in this PhD project, has particular usefulness to combine information from very large population cohorts to permit investigation of small increases in risk

related to environmental exposures and to better identify vulnerable groups in subgroup analyses. This will help set guideline exposure limits that protect the most sensitive groups in the population.

With the rapid developments in science and technology, in together with our willingness to solve these problems, we really can have a tomorrow to enjoy a healthy living in a healthy environment. We can achieve this if all of us-governments, industries, scientists and the general public-work together to seek sustainable options for our environment and health, for now and for the future.

# References


Adar, S. D., Sheppard, L., Vedal, S., Polak, J. F., Sampson, P. D., Diez Roux, A. V., Budoff, M., Jacobs, D. R.,Jr, Barr, R. G., Watson, K. & Kaufman, J. D. (2013) Fine particulate air pollution and the progression of carotid intima-medial thickness: a prospective cohort study from the multi-ethnic study of atherosclerosis and air pollution. *PLoS Medicine.* 10 (4), e1001430.

Allen, N. E., Sudlow, C., Peakman, T., Collins, R. & UK Biobank. (2014) UK biobank data: come and get it. *Science Translational Medicine.* 6 (224), 224ed4.

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

Andersen, Z. J., Raaschou-Nielsen, O., Ketzel, M., Jensen, S. S., Hvidberg, M., Loft, S., Tjonneland, A., Overvad, K. & Sorensen, M. (2012b) Diabetes incidence and long-term exposure to air pollution: a cohort study. *Diabetes Care.* 35 (1), 92-98.

Araujo, J. A., Barajas, B., Kleinman, M., Wang, X., Bennett, B. J., Gong, K. W., Navab, M., Harkema, J., Sioutas, C., Lusis, A. J. & Nel, A. E. (2008) Ambient particulate pollutants in the ultrafine range promote early atherosclerosis and systemic oxidative stress. *Circulation Research.* 102 (5), 589-596.

Argalasova-Sobotova, L., Lekaviciute, J., Jeram, S., Sevcikova, L. & Jurkovicova, J. (2013) Environmental noise and cardiovascular disease in adults: research in Central, Eastern and South-Eastern Europe and Newly Independent States. *Noise & Health.* 15 (62), 22-31.

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

Babisch, W. (2014) Updated exposure-response relationship between road traffic noise and coronary heart diseases: a meta-analysis. *Noise & Health.* 16 (68), 1-9.

Babisch, W. (2008) Road traffic noise and cardiovascular risk. *Noise & Health.* 10 (38), 27-33.

Babisch, W. (2006) Transportation noise and cardiovascular risk: updated review and synthesis of epidemiological studies indicate that the evidence has increased. *Noise & Health.* 8 (30), 1-29.

Babisch, W. (2000) Traffic Noise and Cardiovascular Disease: Epidemiological Review and Synthesis. *Noise & Health.* 2 (8), 9-32.

Babisch, W., Beule, B., Schust, M., Kersten, N. & Ising, H. (2005) Traffic noise and risk of myocardial infarction. *Epidemiology (Cambridge, Mass.).* 16 (1), 33-40.

Babisch, W., Elwood, P. C., Ising, H. & Kruppa, B. (1993a) Traffic noise as a risk factor for myocardial infarction. *Schriftenreihe Des Vereins Fur Wasser-, Boden- Und Lufthygiene.* 88135-166.

Babisch, W., Ising, H., Gallacher, J. E., Sharp, D. S. & Baker, I. A. (1993b) Traffic noise and cardiovascular risk: the Speedwell study, first phase. Outdoor noise levels and risk factors. *Archives of Environmental Health.* 48 (6), 401-405.

Babisch, W., Ising, H., Gallacher, J. E., Sweetnam, P. M. & Elwood, P. C. (1999) Traffic noise and cardiovascular risk: the Caerphilly and Speedwell studies, third phase--10-year follow up. *Archives of Environmental Health.* 54 (3), 210-216.

Babisch, W., Ising, H., Gallacher, J. E. J., Elwood, P. C., Sweetnam, P. M., Yarnell, J. W. G., Bainton, D. & Baker, I. A. (1990) Public Health Implications of Environmental Noise Traffic noise, work noise and cardiovascular risk factors: The Caerphilly and Speedwell collaborative heart disease studies. *Environment International.* 16 (4), 425-435. Available from: http://www.sciencedirect.com/science/article/pii/016041209090011T. Available from: doi: http://dx.doi.org/10.1016/0160-4120(90)90011-T.

Babisch, W., Pershagen, G., Selander, J., Houthuijs, D., Breugelmans, O., Cadum, E., Vigna-Taglianti, F., Katsouyanni, K., Haralabidis, A. S., Dimakopoulou, K., Sourtzi, P., Floud, S. & Hansell, A. L. (2013) Noise annoyance--a modifier of the association between noise level and cardiovascular health? *The Science of the Total Environment.* 452-45350-57.

Babisch, W., Ising, H., Kruppa, B. & Wiens, D. (1994) The incidence of myocardial infarction and its relation to road traffic noise— the Berlin case-control studies. *Environment International.* 20 (4), 469-474. Available from: http://www.sciencedirect.com/science/article/pii/0160412094901953. Available from: doi: http://dx.doi.org/10.1016/0160-4120(94)90195-3.

Basagaña, X., Rivera, M., Aguilera, I., Agis, D., Bouso, L., Elosua, R., Foraster, M., de Nazelle, A., Nieuwenhuijsen, M., Vila, J. & Künzli, N. (2012) Effect of the number of measurement sites on land use regression models in estimating local air pollution. *Atmospheric Environment.* 54634-642.

Basner, M., Babisch, W., Davis, A., Brink, M., Clark, C., Janssen, S. & Stansfeld, S. (2014) Auditory and non-auditory effects of noise on health. *Lancet (London, England).* 383 (9925), 1325-1332.

Baxter, L. K., Dionisio, K. L., Burke, J., Ebelt Sarnat, S., Sarnat, J. A., Hodas, N., Rich, D. Q., Turpin, B. J., Jones, R. R., Mannshardt, E., Kumar, N., Beevers, S. D. & Ozkaynak, H. (2013) Exposure prediction approaches used in air pollution epidemiology studies: key findings and future recommendations. *Journal of Exposure Science & Environmental Epidemiology.* 23 (6), 654-659.

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

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

Beelen, R., Raaschou-Nielsen, O., Stafoggia, M., Andersen, Z. J., Weinmayr, G., Hoffmann, B., Wolf, K., Samoli, E., Fischer, P., Nieuwenhuijsen, M., Vineis, P., Xun, W. W., Katsouyanni, K., Dimakopoulou, K., Oudin, A., Forsberg, B., Modig, L., Havulinna, A. S., Lanki, T., Turunen, A., Oftedal, B., Nystad, W., Nafstad, P., De Faire, U., Pedersen, N. L., Ostenson, C. G., Fratiglioni, L., Penell, J., Korek, M., Pershagen, G., Eriksen, K. T., Overvad, K., Ellermann, T., Eeftens, M., Peeters, P. H., Meliefste, K.,

Wang, M., Bueno-de-Mesquita, B., Sugiri, D., Kramer, U., Heinrich, J., de Hoogh, K., Key, T., Peters, A., Hampel, R., Concin, H., Nagel, G., Ineichen, A., Schaffner, E., Probst-Hensch, N., Kunzli, N., Schindler, C., Schikowski, T., Adam, M., Phuleria, H., Vilier, A., Clavel-Chapelon, F., Declercq, C., Grioni, S., Krogh, V., Tsai, M. Y., 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. (2014a) Effects of long-term exposure to air pollution on natural-cause mortality: an analysis of 22 European cohorts within the multicentre ESCAPE project. *Lancet (London, England).* 383 (9919), 785-795.

Beelen, R., Stafoggia, M., Raaschou-Nielsen, O., Andersen, Z. J., Xun, W. W., 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, N. L., Ostenson, C. G., Fratiglioni, L., Penell, J., Korek, M., Pyko, A., Eriksen, K. T., Tjonneland, A., Becker, T., Eeftens, M., Bots, M., Meliefste, K., Wang, M., Bueno-de-Mesquita, B., Sugiri, D., Kramer, 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, M. Y., 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. (2014b) Long-term exposure to air pollution and cardiovascular mortality: an analysis of 22 European cohorts. *Epidemiology (Cambridge, Mass.).* 25 (3), 368-378.

Beelen, R., Hoek, G., Vienneau, D., Eeftens, M., Dimakopoulou, K., Pedeli, X., Tsai, M., Künzli, N., Schikowski, T., Marcon, A., Eriksen, K. T., Raaschou-Nielsen, O., Stephanou, E., Patelarou, E., Lanki, T., Yli-Tuomi, T., Declercq, C., Falq, G., Stempfelet, M., Birk, M., Cyrys, J., von Klot, S., Nádor, G., Varró, M. J., Dėdelė, A., Gražulevičienė, R., Mölter, A., Lindley, S., Madsen, C., Cesaroni, G., Ranzi, A., Badaloni, C., Hoffmann, B., Nonnemacher, M., Krämer, U., Kuhlbusch, T., Cirach, M., de Nazelle, A., Nieuwenhuijsen, M., Bellander, T., Korek, M., Olsson, D., Strömgren, M., Dons, E., Jerrett, M., Fischer, P., Wang, M., Brunekreef, B. & de Hoogh, K. (2013) Development of NO2 and NOx land use regression models for estimating air pollution exposure in 36 study areas in Europe – The ESCAPE project. *Atmospheric Environment.* 72 (0), 10-23.

Bell, M. L., Davis, D. L. & Fletcher, T. (2004) A retrospective assessment of mortality from the London smog episode of 1952: the role of influenza and pollution. *Environmental Health Perspectives.* 112 (1), 6-8.

Belojevic, G., Paunovic, K., Jakovljevic, B., Stojanov, V., Ilic, J., Slepcevic, V. & Saric-Tanaskovic, M. (2011) Cardiovascular effects of environmental noise: research in Serbia. *Noise & Health.* 13 (52), 217-220.

Bluhm, G. & Eriksson, C. (2011) Cardiovascular effects of environmental noise: research in Sweden. *Noise & Health.* 13 (52), 212-216.

Bodin, T., Bjork, J., Mattisson, K., Bottai, M., Rittner, R., Gustavsson, P., Jakobsson, K., Ostergren, P. O. & Albin, M. (2016) Road traffic noise, air pollution and myocardial infarction: a prospective cohort study. *International Archives of Occupational and Environmental Health.*

Boldo, E., Linares, C., Lumbreras, J., Borge, R., Narros, A., Garcia-Perez, J., Fernandez-Navarro, P., Perez-Gomez, B., Aragones, N., Ramis, R., Pollan, M., Moreno, T., Karanasiou, A. & Lopez-Abente, G. (2011) Health impact assessment of a reduction in ambient PM(2.5) levels in Spain. *Environment International.* 37 (2), 342-348.

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

Briggs DJ, Collins S, Elliott P, Fischer P, Kingham S, Lebert E, et al. (1997) Mapping urban air pollution using GIS: a regression-based approach. *International Journal of Geographical Information Science.* 11 (7), 699-718.

Briggs, D. J., de Hoogh, C., Gulliver, J., Wills, J., Elliott, P., Kingham, S. & Smallbone, K. (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.

Brindley, D. N., McCann, B. S., Niaura, R., Stoney, C. M. & Suarez, E. C. (1993) Stress and lipoprotein metabolism: modulators and mechanisms. *Metabolism: Clinical and Experimental.* 42 (9 Suppl 1), 3-15.

Brook, R. D., Jerrett, M., Brook, J. R., Bard, R. L. & Finkelstein, M. M. (2008) The relationship between diabetes mellitus and traffic-related air pollution. *Journal of Occupational and Environmental Medicine / American College of Occupational and Environmental Medicine.* 50 (1), 32-38.

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

Brook, R. D., Xu, X., Bard, R. L., Dvonch, J. T., Morishita, M., Kaciroti, N., Sun, Q., Harkema, J. & Rajagopalan, S. (2013) Reduced metabolic insulin sensitivity following sub-acute exposures to low levels of ambient fine particulate matter air pollution. *The Science of the Total Environment.* 44866-71.

Brown, D. M., Wilson, M. R., MacNee, W., Stone, V. & Donaldson, K. (2001) Size-dependent proinflammatory effects of ultrafine polystyrene particles: a role for surface area and oxidative stress in the enhanced activity of ultrafines. *Toxicology and Applied Pharmacology.* 175 (3), 191-199.

Carey, I. M., Atkinson, R. W., Kent, A. J., van Staa, T., Cook, D. G. & Anderson, H. R. (2013) Mortality associations with long-term exposure to outdoor air pollution in a national English cohort. *American Journal of Respiratory and Critical Care Medicine.* 187 (11), 1226-1233.

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

Cesaroni, G., Forastiere, F., Stafoggia, M., Andersen, Z. J., Badaloni, C., Beelen, R., Caracciolo, B., de Faire, U., Erbel, R., Eriksen, K. T., Fratiglioni, L., Galassi, C., Hampel, R., Heier, M., Hennig, F., Hilding, A., Hoffmann, B., Houthuijs, D., Jockel, K. H., Korek, M., Lanki, T., Leander, K., Magnusson, P. K., Migliore, E., Ostenson, C. G., Overvad, K., Pedersen, N. L., J, J. P., Penell, J., Pershagen, G., Pyko, A., Raaschou-Nielsen, O., Ranzi, A., Ricceri, F., Sacerdote, C., Salomaa, V., Swart, W., Turunen, A. W., Vineis, P., Weinmayr, G., Wolf, K., de Hoogh, K., Hoek, G., Brunekreef, B. & Peters, A. (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 (Clinical Research Ed.).* 348f7412.

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

Chen, L., Zhou, Y., Li, S., Williams, G., Kan, H., Marks, G. B., Morawska, L., Abramson, M. J., Chen, S., Yao, T., Qin, T., Wu, S. & Guo, Y. (2016) Air pollution and fasting blood glucose: A longitudinal study in China. *The Science of the Total Environment.* 541750-755.

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

Chuang, K. J., Chan, C. C., Su, T. C., Lee, C. T. & Tang, C. S. (2007) The effect of urban air pollution on inflammation, oxidative stress, coagulation, and autonomic dysfunction in young adults. *American Journal of Respiratory and Critical Care Medicine.* 176 (4), 370-376.

Chuang, K. J., Yan, Y. H. & Cheng, T. J. (2010) Effect of air pollution on blood pressure, blood lipids, and blood sugar: a population-based approach. *Journal of Occupational and Environmental Medicine / American College of Occupational and Environmental Medicine.* 52 (3), 258-262.

Chuang, K. J., Yan, Y. H., Chiu, S. Y. & Cheng, T. J. (2011) Long-term air pollution exposure and risk factors for cardiovascular diseases among the elderly in Taiwan. *Occupational and Environmental Medicine.* 68 (1), 64-68.

Clark, C., Martin, R., van Kempen, E., Alfred, T., Head, J., Davies, H. W., Haines, M. M., Lopez Barrio, I., Matheson, M. & Stansfeld, S. A. (2006) Exposure-effect relations between aircraft and road traffic noise exposure at school and reading comprehension: the RANCH project. *American Journal of Epidemiology.* 163 (1), 27-37.

Clougherty, J. E., Wright, R. J., Baxter, L. K. & Levy, J. I. (2008) Land use regression modeling of intra-urban residential variability in multiple traffic-related air pollutants. *Environmental Health : A Global Access Science Source.* 717-069X-7-17.

Collins, F. S. & Manolio, T. A. (2007) Merging and emerging cohorts: necessary but not sufficient. *Nature.* 445 (7125), 259.

Coogan, P. F., White, L. F., Jerrett, M., Brook, R. D., Su, J. G., Seto, E., Burnett, R., Palmer, J. R. & Rosenberg, L. (2012) Air pollution and incidence of hypertension and diabetes mellitus in black women living in Los Angeles. *Circulation.* 125 (6), 767-772.

Correia, A. W., Peters, J. L., Levy, J. I., Melly, S. & Dominici, F. (2013) Residential exposure to aircraft noise and hospital admissions for cardiovascular diseases: multi-airport retrospective study. *BMJ (Clinical Research Ed.).* 347f5561.

Crouse, D. L., Peters, P. A., Villeneuve, P. J., Proux, M. O., Shin, H. H., Goldberg, M. S., Johnson, M., Wheeler, A. J., Allen, R. W., Atari, D. O., Jerrett, M., Brauer, M., Brook, J. R., Cakmak, S. & Burnett, R. T. (2015) Within- and between-city contrasts in nitrogen dioxide and mortality in 10 Canadian cities;

a subset of the Canadian Census Health and Environment Cohort (CanCHEC). *Journal of Exposure Science & Environmental Epidemiology.* 25 (5), 482-489.

Crowe, F. L., Appleby, P. N., Travis, R. C. & Key, T. J. (2013) Risk of hospitalization or death from ischemic heart disease among British vegetarians and nonvegetarians: results from the EPIC-Oxford cohort study. *The American Journal of Clinical Nutrition.* 97 (3), 597-603.

Cyrys, J., Eeftens, M., Heinrich, J., Ampe, C., Armengaud, A., Beelen, R., Bellander, T., Beregszaszi, T., Birk, M., Cesaroni, G., Cirach, M., de Hoogh, K., De Nazelle, A., de Vocht, F., Declercq, C., Dėdelė, A., Dimakopoulou, K., Eriksen, K., Galassi, C., Grąulevičienė, R., Grivas, G., Gruzieva, O., Gustafsson, A. H., Hoffmann, B., Iakovides, M., Ineichen, A., Krämer, U., Lanki, T., Lozano, P., Madsen, C., Meliefste, K., Modig, L., Mölter, A., Mosler, G., Nieuwenhuijsen, M., Nonnemacher, M., Oldenwening, M., Peters, A., Pontet, S., Probst-Hensch, N., Quass, U., Raaschou-Nielsen, O., Ranzi, A., Sugiri, D., Stephanou, E. G., Taimisto, P., Tsai, M., Vaskövi, É, Villani, S., Wang, M., Brunekreef, B. & Hoek, G. (2012) Variation of NO2 and NOx concentrations between and within 36 European study areas: Results from the ESCAPE study. *Atmospheric Environment.* 62374-390.

Davey, G. K., Spencer, E. A., Appleby, P. N., Allen, N. E., Knox, K. H. & Key, T. J. (2003) EPIC-Oxford: lifestyle characteristics and nutrient intakes in a cohort of 33 883 meat-eaters and 31 546 non meat-eaters in the UK. *Public Health Nutrition.* 6 (3), 259-269.

Dawes, P., Cruickshanks, K. J., Moore, D. R., Edmondson-Jones, M., McCormack, A., Fortnum, H. & Munro, K. J. (2014) Cigarette smoking, passive smoking, alcohol consumption, and hearing loss. *Journal of the Association for Research in Otolaryngology : JARO.* 15 (4), 663-674.

de Hoogh, K., Korek, M., Vienneau, D., Keuken, M., Kukkonen, J., Nieuwenhuijsen, M. J., Badaloni, C., Beelen, R., Bolignano, A., Cesaroni, G., Pradas, M. C., Cyrys, J., Douros, J., Eeftens, M., Forastiere, F., Forsberg, B., Fuks, K., Gehring, U., Gryparis, A., Gulliver, J., Hansell, A. L., Hoffmann, B., Johansson, C., Jonkers, S., Kangas, L., Katsouyanni, K., Kunzli, N., Lanki, T., Memmesheimer, M., Moussiopoulos, N., Modig, L., Pershagen, G., Probst-Hensch, N., Schindler, C., Schikowski, T., Sugiri, D., Teixido, O., Tsai, M. Y., Yli-Tuomi, T., Brunekreef, B., Hoek, G. & Bellander, T. (2014) Comparing land use regression and dispersion modelling to assess residential exposure to ambient air pollution for epidemiological studies. *Environment International.* 73382-392.

de Kluizenaar, Y., van Lenthe, F. J., Visschedijk, A. J., Zandveld, P. Y., Miedema, H. M. & Mackenbach, J. P. (2013) Road traffic noise, air pollution components and cardiovascular events. *Noise & Health.* 15 (67), 388-397.

Desai, M. & Begg, M. D. (2008) A comparison of regression approaches for analyzing clustered data. *American Journal of Public Health.* 98 (8), 1425-1429.

Diez Roux, A. V., Auchincloss, A. H., Astor, B., Barr, R. G., Cushman, M., Dvonch, T., Jacobs, D. R.,Jr, Kaufman, J., Lin, X. & Samson, P. (2006) Recent exposure to particulate matter and C-reactive protein concentration in the multi-ethnic study of atherosclerosis. *American Journal of Epidemiology.* 164 (5), 437-448.

Dijkema, M. B., Mallant, S. F., Gehring, U., van den Hurk, K., Alssema, M., van Strien, R. T., Fischer, P. H., Nijpels, G., Stehouwer, C. D., Hoek, G., Dekker, J. M. & Brunekreef, B. (2011) Long-term exposure to traffic-related air pollution and type 2 diabetes prevalence in a cross-sectional screening-study in the Netherlands. *Environmental Health : A Global Access Science Source.* 1076-069X-10-76.

Dimakopoulou, K., Samoli, E., Beelen, R., Stafoggia, M., Andersen, Z. J., Hoffmann, B., Fischer, P., Nieuwenhuijsen, M., Vineis, P., Xun, W., Hoek, G., Raaschou-Nielsen, O., Oudin, A., Forsberg, B., Modig, L., Jousilahti, P., Lanki, T., Turunen, A., Oftedal, B., Nafstad, P., Schwarze, P. E., Penell, J., Fratiglioni, L., Andersson, N., Pedersen, N., Korek, M., De Faire, U., Eriksen, K. T., Tjonneland, A., Becker, T., Wang, M., Bueno-de-Mesquita, B., Tsai, M. Y., Eeftens, M., Peeters, P. H., Meliefste, K., Marcon, A., Kramer, U., Kuhlbusch, T. A., Vossoughi, M., Key, T., de Hoogh, K., Hampel, R., Peters, A., Heinrich, J., Weinmayr, G., Concin, H., Nagel, G., Ineichen, A., Jacquemin, B., Stempfelet, M., Vilier, A., Ricceri, F., Sacerdote, C., Pedeli, X., Katsoulis, M., Trichopoulou, A., Brunekreef, B. & Katsouyanni, K. (2014) Air pollution and nonmalignant respiratory mortality in 16 cohorts within the ESCAPE project. *American Journal of Respiratory and Critical Care Medicine.* 189 (6), 684-696.

Directive 2002/49/EC. (2002) Directive of the European Parliament and of the Council of 25 June 2002 relating to the assessment and management of environmental noise. *Official J Eur Communities.* L18912-25.

Dockery, D. W., Pope, C. A.,3rd, Xu, X., Spengler, J. D., Ware, J. H., Fay, M. E., Ferris, B. G.,Jr & Speizer, F. E. (1993) An association between air pollution and mortality in six U.S. cities. *The New England Journal of Medicine.* 329 (24), 1753-1759.

Eeftens, M., Beelen, R., de Hoogh, K., Bellander, T., Cesaroni, G., Cirach, M., Declercq, C., Dedele, A., Dons, E., de Nazelle, A., Dimakopoulou, K., Eriksen, K., Falq, G., Fischer, P., Galassi, C., Grazuleviciene, R., Heinrich, J., Hoffmann, B., Jerrett, M., Keidel, D., Korek, M., Lanki, T., Lindley, S., Madsen, C., Molter, A., Nador, 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, M. Y., Yli-Tuomi, T., Varro, M. J., Vienneau, D., Klot, S., Wolf, K., Brunekreef, B. & Hoek, G. (2012a) 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. *Environmental Science & Technology.* 46 (20), 11195-11205.

Eeftens, M., Beelen, R., Fischer, P., Brunekreef, B., Meliefste, K. & Hoek, G. (2011) Stability of measured and modelled spatial contrasts in NO(2) over time. *Occupational and Environmental Medicine.* 68 (10), 765-770.

Eeftens, M., Beekhuizen, J., Beelen, R., Wang, M., Vermeulen, R., Brunekreef, B., Huss, A. & Hoek, G. (2013) Quantifying urban street configuration for improvements in air pollution models. *Atmospheric Environment.* 721-9.

Eeftens, M., Tsai, M., Ampe, C., Anwander, B., Beelen, R., Bellander, T., Cesaroni, G., Cirach, M., Cyrys, J., de Hoogh, K., De Nazelle, A., de Vocht, F., Declercq, C., Dėdelė, A., Eriksen, K., Galassi, C., Gražulevičienė, R., Grivas, G., Heinrich, J., Hoffmann, B., Iakovides, M., Ineichen, A., Katsouyanni, K., Korek, M., Krämer, U., Kuhlbusch, T., Lanki, T., Madsen, C., Meliefste, K., Mölter, A., Mosler, G., Nieuwenhuijsen, M., Oldenwening, M., Pennanen, A., Probst-Hensch, N., Quass, U., Raaschou-Nielsen, O., Ranzi, A., Stephanou, E., Sugiri, D., Udvardy, O., Vaskövi, É, Weinmayr, G., Brunekreef, B. & Hoek, G. (2012b) Spatial variation of PM2.5, PM10, PM2.5 absorbance and PMcoarse concentrations between and within 20 European study areas and the relationship with NO2 – Results of the ESCAPE project. *Atmospheric Environment.* 62303-317.

Elliott, P., Shaddick, G., Wakefield, J. C., de Hoogh, C. & Briggs, D. J. (2007) Long-term associations of outdoor air pollution with mortality in Great Britain. *Thorax.* 62 (12), 1088-1094.

Eriksson, C., Hilding, A., Pyko, A., Bluhm, G., Pershagen, G. & Ostenson, C. G. (2014) Long-term aircraft noise exposure and body mass index, waist circumference, and type 2 diabetes: a prospective study. *Environmental Health Perspectives.* 122 (7), 687-694.

Eriksson, C., Rosenlund, M., Pershagen, G., Hilding, A., Ostenson, C. G. & Bluhm, G. (2007) Aircraft noise and incidence of hypertension. *Epidemiology (Cambridge, Mass.).* 18 (6), 716-721.

European Commission. (2013) *Final Report ENNAH-European Network on Noise and Health.*

European Commission. (2010) *Special Eurobarometer 347 on Electromagnetic Fields* .

European Commission. (2009) *Perception survey on quality of life in European cities: Analytical report.*

European Environment Agency. (2013) *TERM 2013: transport indicators tracking progress towards environmental targets in Europe.*

European Environment Agency. (2012) *The contribution of transport to air quality - TERM 2012: Transport indicators tracking progress towards environmental targets in Europe.*

European Office WHO. (2011) *Burden of disease from environmental noise: quantification of healthy life years lost in Europe.*

European Union. Available from: http://ec.europa.eu/environment/air/quality/standards.htm .

Evrard, A. S., Lefevre, M., Champelovier, P., Lambert, J. & Laumon, B. (2016) Does aircraft noise exposure increase the risk of hypertension in the population living near airports in France? *Occupational and Environmental Medicine.*

Eze, I. C., Hemkens, L. G., Bucher, H. C., Hoffmann, B., Schindler, C., Kunzli, N., Schikowski, T. & Probst-Hensch, N. M. (2015) Association between ambient air pollution and diabetes mellitus in Europe and North America: systematic review and meta-analysis. *Environmental Health Perspectives.* 123 (5), 381-389.

Fagernaes, C. F., Heuch, I., Zwart, J. A., Winsvold, B. S., Linde, M. & Hagen, K. (2015) Blood pressure as a risk factor for headache and migraine: a prospective population-based study. *European Journal of Neurology.* 22 (1), 156-62, e10-1.

Faustini, A., Rapp, R. & Forastiere, F. (2014) Nitrogen dioxide and mortality: review and meta-analysis of long-term studies. *The European Respiratory Journal.*

Filleul, L., Rondeau, V., Vandentorren, S., Le Moual, N., Cantagrel, A., Annesi-Maesano, I., Charpin, D., Declercq, C., Neukirch, F., Paris, C., Vervloet, D., Brochard, P., Tessier, J. F., Kauffmann, F. & Baldi, I. (2005) Twenty five year mortality and air pollution: results from the French PAARC survey. *Occupational and Environmental Medicine.* 62 (7), 453-460.

Floud, S., Blangiardo, M., Clark, C., de Hoogh, K., Babisch, W., Houthuijs, D., Swart, W., Pershagen, G., Katsouyanni, K., Velonakis, M., Vigna-Taglianti, F., Cadum, E. & Hansell, A. L. (2013) Exposure to aircraft and road traffic noise and associations with heart disease and stroke in six European countries: a cross-sectional study. *Environmental Health : A Global Access Science Source.* 1289-069X-12-89.

Foraster, M. (2013) Is it traffic-related air pollution or road traffic noise, or both? Key questions not yet settled! *International Journal of Public Health.* 58 (5), 647-648.

Foraster, M., Basagana, X., Aguilera, I., Rivera, M., Agis, D., Bouso, L., Deltell, A., Marrugat, J., Ramos, R., Sunyer, J., Vila, J., Elosua, R. & Kunzli, N. (2014) Association of long-term exposure to traffic-related air pollution with blood pressure and hypertension in an adult population-based cohort in Spain (the REGICOR study). *Environmental Health Perspectives.* 122 (4), 404-411.

Forbes, L. J., Patel, M. D., Rudnicka, A. R., Cook, D. G., Bush, T., Stedman, J. R., Whincup, P. H., Strachan, D. P. & Anderson, H. R. (2009a) Chronic exposure to outdoor air pollution and diagnosed cardiovascular disease: meta-analysis of three large cross-sectional surveys. *Environmental Health : A Global Access Science Source.* 830-069X-8-30.

Forbes, L. J., Patel, M. D., Rudnicka, A. R., Cook, D. G., Bush, T., Stedman, J. R., Whincup, P. H., Strachan, D. P. & Anderson, R. H. (2009b) Chronic exposure to outdoor air pollution and markers of systemic inflammation. *Epidemiology (Cambridge, Mass.).* 20 (2), 245-253.

Fortier, I., Burton, P. R., Robson, P. J., Ferretti, V., Little, J., L'Heureux, F., Deschenes, M., Knoppers, B. M., Doiron, D., Keers, J. C., Linksted, P., Harris, J. R., Lachance, G., Boileau, C., Pedersen, N. L., Hamilton, C. M., Hveem, K., Borugian, M. J., Gallagher, R. P., McLaughlin, J., Parker, L., Potter, J. D., Gallacher, J., Kaaks, R., Liu, B., Sprosen, T., Vilain, A., Atkinson, S. A., Rengifo, A., Morton, R., Metspalu, A., Wichmann, H. E., Tremblay, M., Chisholm, R. L., Garcia-Montero, A., Hillege, H., Litton, J. E., Palmer, L. J., Perola, M., Wolffenbuttel, B. H., Peltonen, L. & Hudson, T. J. (2010) Quality, quantity and harmony: the DataSHaPER approach to integrating data across bioclinical studies. *International Journal of Epidemiology.* 39 (5), 1383-1393.

Fortier, I., Doiron, D., Little, J., Ferretti, V., L'Heureux, F., Stolk, R. P., Knoppers, B. M., Hudson, T. J., Burton, P. R. & International Harmonization Initiative. (2011) Is rigorous retrospective harmonization possible? Application of the DataSHaPER approach across 53 large studies. *International Journal of Epidemiology.* 40 (5), 1314-1328.

Fuks, K., Moebus, S., Hertel, S., Viehmann, A., Nonnemacher, M., Dragano, N., Mohlenkamp, S., Jakobs, H., Kessler, C., Erbel, R., Hoffmann, B. & Heinz Nixdorf Recall Study Investigative Group. (2011) Long-term urban particulate air pollution, traffic noise, and arterial blood pressure. *Environmental Health Perspectives.* 119 (12), 1706-1711.

Galbraith, S., Daniel, J. A. & Vissel, B. (2010) A study of clustered data and approaches to its analysis. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience.* 30 (32), 10601-10608.

Gan, W. Q., Davies, H. W., Koehoorn, M. & Brauer, M. (2012) Association of long-term exposure to community noise and traffic-related air pollution with coronary heart disease mortality. *American Journal of Epidemiology.* 175 (9), 898-906.

Gan, W. Q., Koehoorn, M., Davies, H. W., Demers, P. A., Tamburic, L. & Brauer, M. (2011) Long-term exposure to traffic-related air pollution and the risk of coronary heart disease hospitalization and mortality. *Environmental Health Perspectives.* 119 (4), 501-507.

Garg, N. & Maji, S. (2014) A critical review of principal traffic noise models: Strategies and implications. *Environmental Impact Assessment Review.* 4668-81.

Gaye, A., Marcon, Y., Isaeva, J., LaFlamme, P., Turner, A., Jones, E. M., Minion, J., Boyd, A. W., Newby, C. J., Nuotio, M. L., Wilson, R., Butters, O., Murtagh, B., Demir, I., Doiron, D., Giepmans, L., Wallace, S. E., Budin-Ljosne, I., Oliver Schmidt, C., Boffetta, P., Boniol, M., Bota, M., Carter, K. W., deKlerk, N., Dibben, C., Francis, R. W., Hiekkalinna, T., Hveem, K., Kvaloy, K., Millar, S., Perry, I. J., Peters, A., Phillips, C. M., Popham, F., Raab, G., Reischl, E., Sheehan, N., Waldenberger, M., Perola, M., van den Heuvel, E., Macleod, J., Knoppers, B. M., Stolk, R. P., Fortier, I., Harris, J. R., Woffenbuttel, B. H., Murtagh, M. J., Ferretti, V. & Burton, P. R. (2014) DataSHIELD: taking the analysis to the data, not the data to the analysis. *International Journal of Epidemiology.* 43 (6), 1929-1944.

Genc, S., Zadeoglulari, Z., Fuss, S. H. & Genc, K. (2012) The adverse effects of air pollution on the nervous system. *Journal of Toxicology.* 2012782462.

Gilbert, N. L., Goldberg, M. S., Beckerman, B., Brook, J. R. & Jerrett, M. (2005) Assessing spatial variability of ambient nitrogen dioxide in Montreal, Canada, with a land-use regression model. *Journal of the Air & Waste Management Association (1995).* 55 (8), 1059-1063.

Gulliver, J., de Hoogh, K., Hansell, A. & Vienneau, D. (2013) Development and back-extrapolation of NO2 land use regression models for historic exposure assessment in Great Britain. *Environmental Science & Technology.* 47 (14), 7804-7811.

Gulliver, J., de Hoogh, K., Fecht, D., Vienneau, D. & Briggs, D. (2011) Comparative assessment of GIS-based methods and metrics for estimating long-term exposures to air pollution. *Atmospheric Environment.* 45 (39), 7072-7080.

Hajat, A., Allison, M., Diez-Roux, A. V., Jenny, N. S., Jorgensen, N. W., Szpiro, A. A., Vedal, S. & Kaufman, J. D. (2015) Long-term exposure to air pollution and markers of inflammation, coagulation, and endothelial activation: a repeat-measures analysis in the Multi-Ethnic Study of Atherosclerosis (MESA). *Epidemiology (Cambridge, Mass.).* 26 (3), 310-320.

Hales, S., Blakely, T. & Woodward, A. (2012) Air pollution and mortality in New Zealand: cohort study. *Journal of Epidemiology and Community Health.* 66 (5), 468-473.

Halonen, J. I., Hansell, A. L., Gulliver, J., Morley, D., Blangiardo, M., Fecht, D., Toledano, M. B., Beevers, S. D., Anderson, H. R., Kelly, F. J. & Tonne, C. (2015) Road traffic noise is associated with increased cardiovascular morbidity and mortality and all-cause mortality in London. *European Heart Journal.* 36 (39), 2653-2661.

Hanninen, O., Knol, A. B., Jantunen, M., Lim, T. A., Conrad, A., Rappolder, M., Carrer, P., Fanetti, A. C., Kim, R., Buekers, J., Torfs, R., Iavarone, I., Classen, T., Hornberg, C., Mekel, O. C. & EBoDE Working Group. (2014) Environmental burden of disease in Europe: assessing nine risk factors in six countries. *Environmental Health Perspectives.* 122 (5), 439-446.

Hansell, A. L., Blangiardo, M., Fortunato, L., Floud, S., de Hoogh, K., Fecht, D., Ghosh, R. E., Laszlo, H. E., Pearson, C., Beale, L., Beevers, S., Gulliver, J., Best, N., Richardson, S. & Elliott, P. (2013) Aircraft noise and cardiovascular disease near Heathrow airport in London: small area study. *BMJ (Clinical Research Ed.).* 347f5432.

Hart, J. E., Puett, R. C., Rexrode, K. M., Albert, C. M. & Laden, F. (2015) Effect Modification of Long-Term Air Pollution Exposures and the Risk of Incident Cardiovascular Disease in US Women. *Journal of the American Heart Association.* 4 (12), 10.1161/JAHA.115.002301.

Hart, J. E., Rimm, E. B., Rexrode, K. M. & Laden, F. (2013) Changes in traffic exposure and the risk of incident myocardial infarction and all-cause mortality. *Epidemiology (Cambridge, Mass.).* 24 (5), 734-742.

Heinecke, J. W. (2012) The not-so-simple HDL story: A new era for quantifying HDL and cardiovascular risk? *Nature Medicine.* 18 (9), 1346-1347.

Heinrich, J., Thiering, E., Rzehak, P., Kramer, U., Hochadel, M., Rauchfuss, K. M., Gehring, U. & Wichmann, H. E. (2013) Long-term exposure to NO2 and PM10 and all-cause and cause-specific mortality in a prospective cohort of women. *Occupational and Environmental Medicine.* 70 (3), 179-186.

Henderson, S. B., Beckerman, B., Jerrett, M. & Brauer, M. (2007) Application of land use regression to estimate long-term concentrations of traffic-related nitrogen oxides and fine particulate matter. *Environmental Science & Technology.* 41 (7), 2422-2428.

Hennig, F., Fuks, K., Moebus, S., Weinmayr, G., Memmesheimer, M., Jakobs, H., Brocker-Preuss, M., Fuhrer-Sakel, D., Mohlenkamp, S., Erbel, R., Jockel, K. H., Hoffmann, B. & Heinz Nixdorf Recall Study Investigative Group. (2014) Association between source-specific particulate matter air pollution and hs-CRP: local traffic and industrial emissions. *Environmental Health Perspectives.* 122 (7), 703-710.

Henry, J. P. (1992) Biological basis of the stress response. *Integrative Physiological and Behavioral Science : The Official Journal of the Pavlovian Society.* 27 (1), 66-83.

Hoek, G., Brunekreef, B., Goldbohm, S., Fischer, P. & van den Brandt, P. A. (2002) Association between mortality and indicators of traffic-related air pollution in the Netherlands: a cohort study. *Lancet (London, England).* 360 (9341), 1203-1209.

Hoek, G., Krishnan, R. M., Beelen, R., Peters, A., Ostro, B., Brunekreef, B. & Kaufman, J. D. (2013) Long-term air pollution exposure and cardio- respiratory mortality: a review. *Environmental Health : A Global Access Science Source.* 12 (1), 43-069X-12-43.

Hoek, G., Beelen, R., de Hoogh, K., Vienneau, D., Gulliver, J., Fischer, P. & Briggs, D. (2008) A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmospheric Environment.* 42 (33), 7561-7578.

Hoffmann, B., Moebus, S., Dragano, N., Stang, A., Mohlenkamp, S., Schmermund, A., Memmesheimer, M., Brocker-Preuss, M., Mann, K., Erbel, R. & Jockel, K. H. (2009) Chronic residential exposure to particulate matter air pollution and systemic inflammatory markers. *Environmental Health Perspectives.* 117 (8), 1302-1308.

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

Huang, D., Song, X., Cui, Q., Tian, J., Wang, Q. & Yang, K. (2015) Is there an association between aircraft noise exposure and the incidence of hypertension? A meta-analysis of 16784 participants. *Noise & Health.* 17 (75), 93-97.

Huerta, J. M., Tormo, M. J., Egea-Caparros, J. M., Ortola-Devesa, J. B. & Navarro, C. (2009) Accuracy of self-reported diabetes, hypertension and hyperlipidemia in the adult Spanish population. DINO study findings. *Revista Espanola De Cardiologia.* 62 (2), 143-152.

Huss, A., Spoerri, A., Egger, M., Roosli, M. & Swiss National Cohort Study Group. (2010) Aircraft noise, air pollution, and mortality from myocardial infarction. *Epidemiology (Cambridge, Mass.).* 21 (6), 829-836.

Huttunen, K., Siponen, T., Salonen, I., Yli-Tuomi, T., Aurela, M., Dufva, H., Hillamo, R., Linkola, E., Pekkanen, J., Pennanen, A., Peters, A., Salonen, R. O., Schneider, A., Tiittanen, P., Hirvonen, M. R. & Lanki, T. (2012) Low-level exposure to ambient particulate matter is associated with systemic inflammation in ischemic heart disease patients. *Environmental Research.* 11644-51.

Hveem, K. (2011) Creation of a new prospective research biobank: the example of HUNT3. *Methods in Molecular Biology (Clifton, N.J.).* 675231-239.

Jackson, C. L., Redline, S. & Emmons, K. M. (2015a) Sleep as a potential fundamental contributor to disparities in cardiovascular health. *Annual Review of Public Health.* 36417-440.

Jackson, C. L., Redline, S. & Emmons, K. M. (2015b) Sleep as a potential fundamental contributor to disparities in cardiovascular health. *Annual Review of Public Health.* 36417-440.

Janke, K., Propper, C. & Henderson, J. (2009) Do current levels of air pollution kill? The impact of air pollution on population mortality in England. *Health Economics.* 18 (9), 1031-1055.

Jarup, L., Babisch, W., Houthuijs, D., Pershagen, G., Katsouyanni, K., Cadum, E., Dudley, M. L., Savigny, P., Seiffert, I., Swart, W., Breugelmans, O., Bluhm, G., Selander, J., Haralabidis, A., Dimakopoulou, K., Sourtzi, P., Velonakis, M., Vigna-Taglianti, F. & HYENA study team. (2008) Hypertension and exposure to noise near airports: the HYENA study. *Environmental Health Perspectives.* 116 (3), 329-333.

Jerrett, M., Arain, M. A., Kanaroglou, P., Beckerman, B., Crouse, D., Gilbert, N. L., Brook, J. R., Finkelstein, N. & Finkelstein, M. M. (2007) Modeling the intraurban variability of ambient traffic pollution in Toronto, Canada. *Journal of Toxicology and Environmental Health.Part A.* 70 (3-4), 200-212.

Jerrett, M., Burnett, R. T., Beckerman, B. S., Turner, M. C., Krewski, D., Thurston, G., Martin, R. V., van Donkelaar, A., Hughes, E., Shi, Y., Gapstur, S. M., Thun, M. J. & Pope, C. A.,3rd. (2013) Spatial analysis of air pollution and mortality in California. *American Journal of Respiratory and Critical Care Medicine.* 188 (5), 593-599.

Jorgensen, L., Jenssen, T., Joakimsen, O., Heuch, I., Ingebretsen, O. C. & Jacobsen, B. K. (2004) Glycated hemoglobin level is strongly related to the prevalence of carotid artery plaques with high echogenicity in nondiabetic individuals: the Tromso study. *Circulation.* 110 (4), 466-470.

Kalsch, H., Hennig, F., Moebus, S., Mohlenkamp, S., Dragano, N., Jakobs, H., Memmesheimer, M., Erbel, R., Jockel, K. H., Hoffmann, B. & Heinz Nixdorf Recall Study Investigative Group. (2014) Are air pollution and traffic noise independently associated with atherosclerosis: the Heinz Nixdorf Recall Study. *European Heart Journal.* 35 (13), 853-860.

Kan, H., Heiss, G., Rose, K. M., Whitsel, E. A., Lurmann, F. & London, S. J. (2008) Prospective analysis of traffic exposure as a risk factor for incident coronary heart disease: the Atherosclerosis Risk in Communities (ARIC) study. *Environmental Health Perspectives.* 116 (11), 1463-1468.

Kannel, W. B. (1990) Bishop lecture. Contribution of the Framingham Study to preventive cardiology. *Journal of the American College of Cardiology.* 15 (1), 206-211.

Kaptoge, S., Seshasai, S. R., Gao, P., Freitag, D. F., Butterworth, A. S., Borglykke, A., Di Angelantonio, E., Gudnason, V., Rumley, A., Lowe, G. D., Jorgensen, T. & Danesh, J. (2014) Inflammatory cytokines and risk of coronary heart disease: new prospective study and updated meta-analysis. *European Heart Journal.* 35 (9), 578-589.

Katsoulis, M., Dimakopoulou, K., Pedeli, X., Trichopoulos, D., Gryparis, A., Trichopoulou, A. & Katsouyanni, K. (2014) Long-term exposure to traffic-related air pollution and cardiovascular health in a Greek cohort study. *The Science of the Total Environment.* 490934-940.

Katsouyanni, K. (2003) Ambient air pollution and health. *British Medical Bulletin.* 68143-156.

Kelly, F. J. & Fussell, J. C. (2012a) Size, source and chemical composition as determinants of toxicity attributable to ambient particulate matter. *Atmospheric Environment.* 60504-526.

Kelly, F. J. & Fussell, J. C. (2012b) Size, source and chemical composition as determinants of toxicity attributable to ambient particulate matter. *Atmospheric Environment.* 60504-526.

Kempen, E. (2011) Cardiovascular effects of environmental noise: research in The Netherlands. *Noise & Health.* 13 (52), 221-228.

Kephalopoulos, S., Paviotti, M., Anfosso-Ledee, F., Van Maercke, D., Shilton, S. & Jones, N. (2014) Advances in the development of common noise assessment methods in Europe: The CNOSSOS-EU framework for strategic environmental noise mapping. *The Science of the Total Environment.* 482-483400-410.

Khafaie, M. A., Salvi, S. S., Ojha, A., Khafaie, B., Gore, S. S. & Yajnik, C. S. (2013) Systemic inflammation (C-reactive protein) in type 2 diabetic patients is associated with ambient air pollution in Pune City, India. *Diabetes Care.* 36 (3), 625-630.

Korek, M. J., Bellander, T. D., Lind, T., Bottai, M., Eneroth, K. M., Caracciolo, B., de Faire, U. H., Fratiglioni, L., Hilding, A., Leander, K., Magnusson, P. K., Pedersen, N. L., Ostenson, C. G., Pershagen, G. & Penell, J. C. (2015) Traffic-related air pollution exposure and incidence of stroke in four cohorts from Stockholm. *Journal of Exposure Science & Environmental Epidemiology.* 25 (5), 517-523.

Kramer, U., Herder, C., Sugiri, D., Strassburger, K., Schikowski, T., Ranft, U. & Rathmann, W. (2010) Traffic-related air pollution and incident type 2 diabetes: results from the SALIA cohort study. *Environmental Health Perspectives.* 118 (9), 1273-1279.

Krokstad, S., Langhammer, A., Hveem, K., Holmen, T. L., Midthjell, K., Stene, T. R., Bratberg, G., Heggland, J. & Holmen, J. (2013) Cohort Profile: the HUNT Study, Norway. *International Journal of Epidemiology.* 42 (4), 968-977.

Kuehn, B. M. (2014) WHO: More than 7 million air pollution deaths each year. *JAMA : The Journal of the American Medical Association.* 311 (15), 1486.

Kunzli, N. (2014) Effects of near-road and regional air pollution: the challenge of separation. *Thorax.* 69 (6), 503-504.

Kunzli, N., Jerrett, M., Garcia-Esteban, R., Basagana, X., Beckermann, B., Gilliland, F., Medina, M., Peters, J., Hodis, H. N. & Mack, W. J. (2010) Ambient air pollution and the progression of atherosclerosis in adults. *PloS One.* 5 (2), e9096.

Laden, F., Schwartz, J., Speizer, F. E. & Dockery, D. W. (2006) Reduction in fine particulate air pollution and mortality: Extended follow-up of the Harvard Six Cities study. *American Journal of Respiratory and Critical Care Medicine.* 173 (6), 667-672.

Lanki, T., Hampel, R., Tiittanen, P., Andrich, S., Beelen, R., Brunekreef, B., Dratva, J., De Faire, U., Fuks, K. B., Hoffmann, B., Imboden, M., Jousilahti, P., Koenig, W., Mahabadi, A. A., Kunzli, N., Pedersen, N. L., Penell, J., Pershagen, G., Probst-Hensch, N. M., Schaffner, E., Schindler, C., Sugiri, D., Swart, W. J., Tsai, M. Y., Turunen, A. W., Weinmayr, G., Wolf, K., Yli-Tuomi, T. & Peters, A. (2015) Air Pollution from Road Traffic and Systemic Inflammation in Adults: A Cross-Sectional Analysis in the European ESCAPE Project. *Environmental Health Perspectives.* 123 (8), 785-791.

Lepeule, J., Laden, F., Dockery, D. & Schwartz, J. (2012) Chronic exposure to fine particles and mortality: an extended follow-up of the Harvard Six Cities study from 1974 to 2009. *Environmental Health Perspectives.* 120 (7), 965-970.

Lercher, P., Botteldooren, D., Widmann, U., Uhrner, U. & Kammeringer, E. (2011) Cardiovascular effects of environmental noise: research in Austria. *Noise & Health.* 13 (52), 234-250.

Lercher, P. (1996) Environmental noise and health: An integrated research perspective. *Environment International.* 22 (1), 117-129.

Li, Y., Rittenhouse-Olson, K., Scheider, W. L. & Mu, L. (2012) Effect of particulate matter air pollution on C-reactive protein: a review of epidemiologic studies. *Reviews on Environmental Health.* 27 (2-3), 133-149.

Libby, P. (2002) Inflammation in atherosclerosis. *Nature.* 420 (6917), 868-874.

Lipsett, M. J., Ostro, B. D., Reynolds, P., Goldberg, D., Hertz, A., Jerrett, M., Smith, D. F., Garcia, C., Chang, E. T. & Bernstein, L. (2011) Long-term exposure to air pollution and cardiorespiratory disease in the California teachers study cohort. *American Journal of Respiratory and Critical Care Medicine.* 184 (7), 828-835.

Liu, C., Yang, C., Zhao, Y., Ma, Z., Bi, J., Liu, Y., Meng, X., Wang, Y., Cai, J., Kan, H. & Chen, R. (2016) Associations between long-term exposure to ambient particulate air pollution and type 2 diabetes prevalence, blood glucose and glycosylated hemoglobin levels in China. *Environment International.* 92-93416-421.

Madsen, C., Gehring, U., Håberg, S. E., Nafstad, P., Meliefste, K., Nystad, W., Lødrup Carlsen, K. C. & Brunekreef, B. (2011) Comparison of land-use regression models for predicting spatial NOx contrasts over a three year period in Oslo, Norway. *Atmospheric Environment.* 45 (21), 3576-3583.

Maheswaran, R. & Elliott, P. (2003) Stroke mortality associated with living near main roads in England and wales: a geographical study. *Stroke; a Journal of Cerebral Circulation.* 34 (12), 2776-2780.

Maheswaran, R., Haining, R. P., Brindley, P., Law, J., Pearson, T., Fryers, P. R., Wise, S. & Campbell, M. J. (2005) Outdoor air pollution and stroke in Sheffield, United Kingdom: a small-area level geographical study. *Stroke; a Journal of Cerebral Circulation.* 36 (2), 239-243.

Maheswaran, R., Pearson, T., Smeeton, N. C., Beevers, S. D., Campbell, M. J. & Wolfe, C. D. (2012) Outdoor air pollution and incidence of ischemic and hemorrhagic stroke: a small-area level ecological study. *Stroke; a Journal of Cerebral Circulation.* 43 (1), 22-27.

Maheswaran, R., Pearson, T., Smeeton, N. C., Beevers, S. D., Campbell, M. J. & Wolfe, C. D. (2010) Impact of outdoor air pollution on survival after stroke: population-based cohort study. *Stroke; a Journal of Cerebral Circulation.* 41 (5), 869-877.

Mark Jenner. (1995) The politics of London air John Evelyn's Fumifugium and the Restoration. *The Historical Journal.* 38535-551.

Maschke, C. (2011) Cardiovascular effects of environmental noise: research in Germany. *Noise & Health.* 13 (52), 205-211.

Maschke, C., Rupp, T. & Hecht, K. (2000) The influence of stressors on biochemical reactions--a review of present scientific findings with noise. *International Journal of Hygiene and Environmental Health.* 203 (1), 45-53.

Melamed, S., Froom, P., Kristal-Boneh, E., Gofer, D. & Ribak, J. (1997) Industrial noise exposure, noise annoyance, and serum lipid levels in blue-collar workers--the CORDIS Study. *Archives of Environmental Health.* 52 (4), 292-298.

Miller, K. A., Siscovick, D. S., Sheppard, L., Shepherd, K., Sullivan, J. H., Anderson, G. L. & Kaufman, J. D. (2007) Long-term exposure to air pollution and incidence of cardiovascular events in women. *The New England Journal of Medicine.* 356 (5), 447-458.

Mölter, A., Lindley, S., de Vocht, F., Simpson, A. & Agius, R. (2010) Modelling air pollution for epidemiologic research – Part II: Predicting temporal variation through land use regression. *Science of the Total Environment.* 409 (1), 211-217.

Morgenstern, V., Zutavern, A., Cyrys, J., Brockow, I., Gehring, U., Koletzko, S., Bauer, C. P., Reinhardt, D., Wichmann, H. E. & Heinrich, J. (2007) Respiratory health and individual estimated exposure to traffic-related air pollutants in a cohort of young children. *Occupational and Environmental Medicine.* 64 (1), 8-16.

Morley, D. W., de Hoogh, K., Fecht, D., Fabbri, F., Bell, M., Goodman, P. S., Elliott, P., Hodgson, S., Hansell, A. & Gulliver, J. (2015) International scale implementation of the CNOSSOS-EU road traffic noise prediction model for epidemiological studies. *Environmental Pollution (Barking, Essex : 1987).* 206332-341.

Munzel, T., Gori, T., Babisch, W. & Basner, M. (2014) Cardiovascular effects of environmental noise exposure. *European Heart Journal.* 35 (13), 829-836.

Murphy M, K. E. (2014) *Environmental Noise Pollution: noise mapping, public health and policy. . ,* Elsevier.

Murray, C. J., Vos, T., Lozano, R., Naghavi, M., Flaxman, A. D., Michaud, C., Ezzati, M., Shibuya, K., Salomon, J. A., Abdalla, S., Aboyans, V., Abraham, J., Ackerman, I., Aggarwal, R., Ahn, S. Y., Ali, M. K., Alvarado, M., Anderson, H. R., Anderson, L. M., Andrews, K. G., Atkinson, C., Baddour, L. M., Bahalim, A. N., Barker-Collo, S., Barrero, L. H., Bartels, D. H., Basanez, M. G., Baxter, A., Bell, M. L., Benjamin, E. J., Bennett, D., Bernabe, E., Bhalla, K., Bhandari, B., Bikbov, B., Bin Abdulhak, A., Birbeck, G., Black, J. A., Blencowe, H., Blore, J. D., Blyth, F., Bolliger, I., Bonaventure, A., Boufous, S., Bourne, R., Boussinesq, M., Braithwaite, T., Brayne, C., Bridgett, L., Brooker, S., Brooks, P., Brugha, T. S., Bryan-Hancock, C., Bucello, C., Buchbinder, R., Buckle, G., Budke, C. M., Burch, M., Burney, P., Burstein, R., Calabria, B., Campbell, B., Canter, C. E., Carabin, H., Carapetis, J., Carmona, L., Cella, C., Charlson, F., Chen, H., Cheng, A. T., Chou, D., Chugh, S. S., Coffeng, L. E., Colan, S. D., Colquhoun, S., Colson, K. E., Condon, J., Connor, M. D., Cooper, L. T., Corriere, M., Cortinovis, M., de Vaccaro, K. C., Couser, W., Cowie, B. C., Criqui, M. H., Cross, M., Dabhadkar, K. C., Dahiya, M., Dahodwala, N., Damsere-Derry, J., Danaei, G., Davis, A., De Leo, D., Degenhardt, L., Dellavalle, R., Delossantos, A., Denenberg, J., Derrett, S., Des Jarlais, D. C., Dharmaratne, S. D., Dherani, M., Diaz-Torne, C., Dolk, H., Dorsey, E. R., Driscoll, T., Duber, H., Ebel, B., Edmond, K., Elbaz, A., Ali, S. E., Erskine, H., Erwin, P. J., Espindola, P., Ewoigbokhan, S. E., Farzadfar, F., Feigin, V., Felson, D. T., Ferrari, A., Ferri, C. P., Fevre, E. M., Finucane, M. M., Flaxman, S., Flood, L., Foreman, K., Forouzanfar, M. H., Fowkes, F. G., Fransen, M., Freeman, M. K., Gabbe, B. J., Gabriel, S. E., Gakidou, E., Ganatra, H. A., Garcia, B., Gaspari, F., Gillum, R. F., Gmel, G., Gonzalez-Medina, D., Gosselin, R., Grainger, R., Grant, B., Groeger, J., Guillemin, F., Gunnell, D., Gupta, R., Haagsma, J., Hagan, H., Halasa, Y. A., Hall, W., Haring, D., Haro, J. M., Harrison, J. E., Havmoeller, R., Hay, R. J., Higashi, H., Hill, C., Hoen, B., Hoffman, H., Hotez, P. J., Hoy, D., Huang, J. J., Ibeanusi, S. E., Jacobsen, K. H., James, S. L., Jarvis, D., Jasrasaria, R., Jayaraman, S., Johns, N., Jonas, J. B., Karthikeyan, G., Kassebaum, N., Kawakami, N., Keren, A., Khoo, J. P., King, C. H., Knowlton, L. M., Kobusingye, O., Koranteng, A., Krishnamurthi, R., Laden, F., Lalloo, R., Laslett, L. L., Lathlean, T., Leasher, J. L., Lee, Y. Y., Leigh, J., Levinson, D., Lim, S. S., Limb, E., Lin, J. K., Lipnick, M., Lipshultz, S. E., Liu, W., Loane, M., Ohno, S. L., Lyons, R., Mabweijano, J., MacIntyre, M. F., Malekzadeh, R., Mallinger, L., Manivannan, S., Marcenes, W., March, L., Margolis, D. J., Marks, G. B., Marks, R., Matsumori, A., Matzopoulos, R., Mayosi, B. M., McAnulty, J. H., McDermott, M. M., McGill, N., McGrath, J., Medina-Mora, M. E., Meltzer, M., Mensah, G. A., Merriman, T. R., Meyer, A. C., Miglioli, V., Miller, M., Miller, T. R., Mitchell, P. B., Mock, C., Mocumbi, A. O., Moffitt, T. E., Mokdad, A. A., Monasta, L., Montico, M., Moradi-Lakeh, M., Moran, A., Morawska, L., Mori, R., Murdoch, M. E., Mwaniki, M. K., Naidoo, K., Nair, M. N., Naldi, L., Narayan, K. M., Nelson, P. K., Nelson, R. G., Nevitt, M. C., Newton, C. R., Nolte, S., Norman, P., Norman, R., O'Donnell, M., O'Hanlon, S., Olives, C., Omer, S. B., Ortblad, K., Osborne, R., Ozgediz, D., Page, A., Pahari, B., Pandian, J. D., Rivero, A. P., Patten, S. B., Pearce, N., Padilla, R. P., Perez-Ruiz, F., Perico, N., Pesudovs, K., Phillips, D., Phillips, M. R., Pierce, K., Pion, S., Polanczyk, G. V., Polinder, S., Pope, C. A.,3rd, Popova, S., Porrini, E., Pourmalek, F., Prince, M., Pullan, R. L., Ramaiah, K. D., Ranganathan, D., Razavi, H., Regan, M., Rehm, J. T., Rein, D. B., Remuzzi, G., Richardson, K., Rivara, F. P., Roberts, T., Robinson, C., De Leon, F. R., Ronfani, L., Room, R., Rosenfeld, L. C., Rushton, L., Sacco, R. L., Saha, S., Sampson, U., Sanchez-Riera, L., Sanman, E., Schwebel, D. C., Scott, J. G., Segui-Gomez, M., Shahraz, S., Shepard, D. S., Shin, H., Shivakoti, R., Singh, D., Singh, G. M., Singh, J. A., Singleton, J., Sleet, D. A., Sliwa, K., Smith, E., Smith, J. L., Stapelberg, N. J., Steer, A., Steiner, T., Stolk, W. A., Stovner, L. J., Sudfeld, C., Syed, S., Tamburlini, G., Tavakkoli, M., Taylor, H. R., Taylor, J. A., Taylor, W. J., Thomas, B., Thomson, W. M., Thurston, G. D., Tleyjeh, I. M., Tonelli, M., Towbin, J. A., Truelsen, T., Tsilimbaris, M. K., Ubeda, C., Undurraga, E. A., van der Werf, M. J., van Os, J., Vavilala, M. S., Venketasubramanian, N., Wang, M., Wang, W., Watt, K., Weatherall, D. J., Weinstock, M. A., Weintraub, R., Weisskopf, M. G., Weissman, M. M., White, R. A., Whiteford, H., Wiebe, N., Wiersma, S. T., Wilkinson, J. D., Williams, H. C., Williams, S. R., Witt, E., Wolfe, F., Woolf, A. D., Wulf, S., Yeh, P. H., Zaidi, A. K., Zheng, Z. J., Zonies, D., Lopez, A. D., AlMazroa, M. A. & Memish, Z. A. (2012) Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a

systematic analysis for the Global Burden of Disease Study 2010. *Lancet (London, England).* 380 (9859), 2197-2223.

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

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

Niaura, R., Stoney, C. M. & Herbert, P. N. (1992) Lipids in psychological research: the last decade. *Biological Psychology.* 34 (1), 1-43.

Nicholas Hewitt, C. (1991) Spatial variations in nitrogen dioxide concentrations in an urban area. *Atmospheric Environment.Part B.Urban Atmosphere.* 25 (3), 429-434. Available from: http://www.sciencedirect.com/science/article/pii/0957127291900146. Available from: doi: http://dx.doi.org/10.1016/0957-1272(91)90014-6.

Niwa, Y., Hiura, Y., Sawamura, H. & Iwai, N. (2008) Inhalation exposure to carbon black induces inflammatory response in rats. *Circulation Journal : Official Journal of the Japanese Circulation Society.* 72 (1), 144-149.

Nonogaki, K., Fuller, G. M., Fuentes, N. L., Moser, A. H., Staprans, I., Grunfeld, C. & Feingold, K. R. (1995) Interleukin-6 stimulates hepatic triglyceride secretion in rats. *Endocrinology.* 136 (5), 2143-2149.

Ostro, B., Malig, B., Broadwin, R., Basu, R., Gold, E. B., Bromberger, J. T., Derby, C., Feinstein, S., Greendale, G. A., Jackson, E. A., Kravitz, H. M., Matthews, K. A., Sternfeld, B., Tomey, K., Green, R. R. & Green, R. (2014) Chronic PM2.5 exposure and inflammation: determining sensitive subgroups in mid-life women. *Environmental Research.* 132168-175.

Panasevich, S., Leander, K., Rosenlund, M., Ljungman, P., Bellander, T., de Faire, U., Pershagen, G. & Nyberg, F. (2009) Associations of long- and short-term air pollution exposure with markers of inflammation and coagulation in a population sample. *Occupational and Environmental Medicine.* 66 (11), 747-753.

Passchier-Vermeer, W. & Passchier, W. F. (2000) Noise exposure and public health. *Environmental Health Perspectives.* 108 Suppl 1123-131.

Pearson, J. F., Bachireddy, C., Shyamprasad, S., Goldfine, A. B. & Brownstein, J. S. (2010) Association between fine particulate matter and diabetes prevalence in the U.S. *Diabetes Care.* 33 (10), 2196-2201.

Pope, C. A.,3rd. (2001) Particulate air pollution, C-reactive protein, and cardiac risk. *European Heart Journal.* 22 (14), 1149-1150.

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

Pope, C. A.,3rd, Hansen, M. L., Long, R. W., Nielsen, K. R., Eatough, N. L., Wilson, W. E. & Eatough, D. J. (2004b) Ambient particulate air pollution, heart rate variability, and blood markers of inflammation in a panel of elderly subjects. *Environmental Health Perspectives.* 112 (3), 339-345.

Pope, C. A.,3rd, Thun, M. J., Namboodiri, M. M., Dockery, D. W., Evans, J. S., Speizer, F. E. & Heath, C. W.,Jr. (1995) Particulate air pollution as a predictor of mortality in a prospective study of U.S. adults. *American Journal of Respiratory and Critical Care Medicine.* 151 (3 Pt 1), 669-674.

Poursafa, P., Mansourian, M., Motlagh, M. E., Ardalan, G. & Kelishadi, R. (2014a) Is air quality index associated with cardiometabolic risk factors in adolescents? The CASPIAN-III Study. *Environmental Research.* 134105-109.

Poursafa, P., Mansourian, M., Motlagh, M. E., Ardalan, G. & Kelishadi, R. (2014b) Is air quality index associated with cardiometabolic risk factors in adolescents? The CASPIAN-III Study. *Environmental Research.* 134105-109.

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

Raaschou-Nielsen, O., Andersen, Z. J., Jensen, S. S., Ketzel, M., Sorensen, M., Hansen, J., Loft, S., Tjonneland, A. & Overvad, K. (2012) Traffic air pollution and mortality from cardiovascular disease and all causes: a Danish cohort study. *Environmental Health : A Global Access Science Source.* 1160-069X-11-60.

Rader, D. J. & Tall, A. R. (2012) The not-so-simple HDL story: Is it time to revise the HDL cholesterol hypothesis? *Nature Medicine.* 18 (9), 1344-1346.

Rasouli, B., Ahlbom, A., Andersson, T., Grill, V., Midthjell, K., Olsson, L. & Carlsson, S. (2013) Alcohol consumption is associated with reduced risk of Type 2 diabetes and autoimmune diabetes in adults: results from the Nord-Trondelag health study. *Diabetic Medicine : A Journal of the British Diabetic Association.* 30 (1), 56-64.

Ray, K. K., Seshasai, S. R., Wijesuriya, S., Sivakumaran, R., Nethercott, S., Preiss, D., Erqou, S. & Sattar, N. (2009) Effect of intensive control of glucose on cardiovascular outcomes and death in patients with diabetes mellitus: a meta-analysis of randomised controlled trials. *Lancet (London, England).* 373 (9677), 1765-1772.

Recio, A., Linares, C., Banegas, J. R. & Diaz, J. (2016) Road traffic noise effects on cardiovascular, respiratory, and metabolic health: An integrative model of biological mechanisms. *Environmental Research.* 146359-370.

Riboli, E. & Kaaks, R. (1997) The EPIC Project: rationale and study design. European Prospective Investigation into Cancer and Nutrition. *International Journal of Epidemiology.* 26 Suppl 1S6-14.

Ridker, P. M. (2003) Clinical application of C-reactive protein for cardiovascular disease detection and prevention. *Circulation.* 107 (3), 363-369.

Riedl, M. & Diaz-Sanchez, D. (2005) Biology of diesel exhaust effects on respiratory function. *The Journal of Allergy and Clinical Immunology.* 115 (2), 221-8; quiz 229.

Roetzheim, R. G., Freund, K. M., Corle, D. K., Murray, D. M., Snyder, F. R., Kronman, A. C., Jean-Pierre, P., Raich, P. C., Holden, A. E., Darnell, J. S., Warren-Mears, V., Patierno, S. & PNRP Design and Analysis Committee for the Patient Navigation Research Program Investigators. (2012) Analysis of combined data from heterogeneous study designs: an applied example from the patient navigation research program. *Clinical Trials (London, England).* 9 (2), 176-187.

Rohleder, N. (2014) Stimulation of systemic low-grade inflammation by psychosocial stress. *Psychosomatic Medicine.* 76 (3), 181-189.

Rohr, A. C., Wagner, J. G., Morishita, M., Kamal, A., Keeler, G. J. & Harkema, J. R. (2010) Cardiopulmonary responses in spontaneously hypertensive and Wistar-Kyoto rats exposed to concentrated ambient particles from Detroit, Michigan. *Inhalation Toxicology.* 22 (6), 522-533.

Rosenlund, M., Berglind, N., Pershagen, G., Hallqvist, J., Jonson, T. & Bellander, T. (2006) Long-term exposure to urban air pollution and myocardial infarction. *Epidemiology (Cambridge, Mass.).* 17 (4), 383-390.

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

Ruckerl, R., Greven, S., Ljungman, P., Aalto, P., Antoniades, C., Bellander, T., Berglind, N., Chrysohoou, C., Forastiere, F., Jacquemin, B., von Klot, S., Koenig, W., Kuchenhoff, H., Lanki, T., Pekkanen, J., Perucci, C. A., Schneider, A., Sunyer, J., Peters, A. & AIRGENE Study Group. (2007) Air pollution and inflammation (interleukin-6, C-reactive protein, fibrinogen) in myocardial infarction survivors. *Environmental Health Perspectives.* 115 (7), 1072-1080.

Ruckerl, R., Hampel, R., Breitner, S., Cyrys, J., Kraus, U., Carter, J., Dailey, L., Devlin, R. B., Diaz-Sanchez, D., Koenig, W., Phipps, R., Silbajoris, R., Soentgen, J., Soukup, J., Peters, A. & Schneider, A. (2014) Associations between ambient air pollution and blood markers of inflammation and coagulation/fibrinolysis in susceptible populations. *Environment International.* 7032-49.

Ruckerl, R., Ibald-Mulli, A., Koenig, W., Schneider, A., Woelke, G., Cyrys, J., Heinrich, J., Marder, V., Frampton, M., Wichmann, H. E. & Peters, A. (2006) Air pollution and markers of inflammation and coagulation in patients with coronary heart disease. *American Journal of Respiratory and Critical Care Medicine.* 173 (4), 432-441.

Rudez, G., Janssen, N. A., Kilinc, E., Leebeek, F. W., Gerlofs-Nijland, M. E., Spronk, H. M., ten Cate, H., Cassee, F. R. & de Maat, M. P. (2009) Effects of ambient air pollution on hemostasis and inflammation. *Environmental Health Perspectives.* 117 (6), 995-1001.

Ryan, P. H. & LeMasters, G. K. (2007) A review of land-use regression models for characterizing intraurban air pollution exposure. *Inhalation Toxicology.* 19 Suppl 1127-133.

Sade, M. Y., Kloog, I., Liberty, I. F., Katra, I., Novack, L. & Novack, V. (2015) Air Pollution and Serum Glucose Levels: A Population-Based Study. *Medicine.* 94 (27), e1093.

Sahsuvaroglu, T., Arain, A., Kanaroglou, P., Finkelstein, N., Newbold, B., Jerrett, M., Beckerman, B., Brook, J., Finkelstein, M. & Gilbert, N. L. (2006) A land use regression model for predicting ambient concentrations of nitrogen dioxide in Hamilton, Ontario, Canada. *Journal of the Air & Waste Management Association (1995).* 56 (8), 1059-1069.

Sateia, M. J., Doghramji, K., Hauri, P. J. & Morin, C. M. (2000) Evaluation of chronic insomnia. An American Academy of Sleep Medicine review. *Sleep.* 23 (2), 243-308.

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

Scholtens, S., Smidt, N., Swertz, M. A., Bakker, S. J., Dotinga, A., Vonk, J. M., van Dijk, F., van Zon, S. K., Wijmenga, C., Wolffenbuttel, B. H. & Stolk, R. P. (2015) Cohort Profile: LifeLines, a three-generation cohort study and biobank. *International Journal of Epidemiology.* 44 (4), 1172-1180.

Selander, J., Nilsson, M. E., Bluhm, G., Rosenlund, M., Lindqvist, M., Nise, G. & Pershagen, G. (2009) Long-term exposure to road traffic noise and myocardial infarction. *Epidemiology (Cambridge, Mass.).* 20 (2), 272-279.

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

Shanley, R. P., Hayes, R. B., Cromar, K. R., Ito, K., Gordon, T. & Ahn, J. (2016) Particulate Air Pollution and Clinical Cardiovascular Disease Risk Factors. *Epidemiology (Cambridge, Mass.).* 27 (2), 291-298.

Slagter, S. N., van Vliet-Ostaptchouk, J. V., Vonk, J. M., Boezen, H. M., Dullaart, R. P., Kobold, A. C., Feskens, E. J., van Beek, A. P., van der Klauw, M. M. & Wolffenbuttel, B. H. (2014) Combined effects of smoking and alcohol on metabolic syndrome: the LifeLines cohort study. *PloS One.* 9 (4), e96406.

Smith-Warner, S. A., Spiegelman, D., Ritz, J., Albanes, D., Beeson, W. L., Bernstein, L., Berrino, F., van den Brandt, P. A., Buring, J. E., Cho, E., Colditz, G. A., Folsom, A. R., Freudenheim, J. L., Giovannucci, E., Goldbohm, R. A., Graham, S., Harnack, L., Horn-Ross, P. L., Krogh, V., Leitzmann, M. F., McCullough, M. L., Miller, A. B., Rodriguez, C., Rohan, T. E., Schatzkin, A., Shore, R., Virtanen, M., Willett, W. C., Wolk, A., Zeleniuch-Jacquotte, A., Zhang, S. M. & Hunter, D. J. (2006) Methods for pooling results of epidemiologic studies: the Pooling Project of Prospective Studies of Diet and Cancer. *American Journal of Epidemiology.* 163 (11), 1053-1064.

Soares, S. R., Carvalho-Oliveira, R., Ramos-Sanchez, E., Catanozi, S., da Silva, L. F., Mauad, T., Gidlund, M., Goto, H. & Garcia, M. L. (2009) Air pollution and antibodies against modified lipoproteins are associated with atherosclerosis and vascular remodeling in hyperlipemic mice. *Atherosclerosis.* 207 (2), 368-373.

Sorensen, M., Andersen, Z. J., Nordsborg, R. B., Becker, T., Tjonneland, A., Overvad, K. & Raaschou-Nielsen, O. (2013) Long-term exposure to road traffic noise and incident diabetes: a cohort study. *Environmental Health Perspectives.* 121 (2), 217-222.

Sorensen, M., Andersen, Z. J., Nordsborg, R. B., Jensen, S. S., Lillelund, K. G., Beelen, R., Schmidt, E. B., Tjonneland, A., Overvad, K. & Raaschou-Nielsen, O. (2012a) Road traffic noise and incident myocardial infarction: a prospective cohort study. *PloS One.* 7 (6), e39283.

Sorensen, M., Hjortebjerg, D., Eriksen, K. T., Ketzel, M., Tjonneland, A., Overvad, K. & Raaschou-Nielsen, O. (2015) Exposure to long-term air pollution and road traffic noise in relation to cholesterol: A cross-sectional study. *Environment International.* 85238-243.

Sorensen, M., Hoffmann, B., Hvidberg, M., Ketzel, M., Jensen, S. S., Andersen, Z. J., Tjonneland, A., Overvad, K. & Raaschou-Nielsen, O. (2012b) Long-term exposure to traffic-related air pollution associated with blood pressure and self-reported hypertension in a Danish cohort. *Environmental Health Perspectives.* 120 (3), 418-424.

Sorensen, M., Hvidberg, M., Andersen, Z. J., Nordsborg, R. B., Lillelund, K. G., Jakobsen, J., Tjonneland, A., Overvad, K. & Raaschou-Nielsen, O. (2011) Road traffic noise and stroke: a prospective cohort study. *European Heart Journal.* 32 (6), 737-744.

Sorensen, M., Luhdorf, P., Ketzel, M., Andersen, Z. J., Tjonneland, A., Overvad, K. & Raaschou-Nielsen, O. (2014) Combined effects of road traffic noise and ambient air pollution in relation to risk for stroke? *Environmental Research.* 13349-55.

Spencer, E. A., Appleby, P. N., Davey, G. K. & Key, T. J. (2002) Validity of self-reported height and weight in 4808 EPIC-Oxford participants. *Public Health Nutrition.* 5 (4), 561-565.

Spencer, E. A., Roddam, A. W. & Key, T. J. (2004) Accuracy of self-reported waist and hip measurements in 4492 EPIC-Oxford participants. *Public Health Nutrition.* 7 (6), 723-727.

Stafoggia, M., Cesaroni, G., Peters, A., Andersen, Z. J., Badaloni, C., Beelen, R., Caracciolo, B., Cyrys, J., de Faire, U., de Hoogh, K., Eriksen, K. T., Fratiglioni, L., Galassi, C., Gigante, B., Havulinna, A. S., Hennig, F., Hilding, A., Hoek, G., Hoffmann, B., Houthuijs, D., Korek, M., Lanki, T., Leander, K., Magnusson, P. K., Meisinger, C., Migliore, E., Overvad, K., Ostenson, C. G., Pedersen, N. L., Pekkanen, J., Penell, J., Pershagen, G., Pundt, N., Pyko, A., Raaschou-Nielsen, O., Ranzi, A., Ricceri, F., Sacerdote, C., Swart, W. J., Turunen, A. W., Vineis, P., Weimar, C., Weinmayr, G., Wolf, K., Brunekreef, B. & Forastiere, F. (2014) Long-term exposure to ambient air pollution and incidence of cerebrovascular events: results from 11 European cohorts within the ESCAPE project. *Environmental Health Perspectives.* 122 (9), 919-925.

Stansfeld, S. & Crombie, R. (2011) Cardiovascular effects of environmental noise: research in the United Kingdom. *Noise & Health.* 13 (52), 229-233.

Stansfeld, S. A. & Shipley, M. (2015) Noise sensitivity and future risk of illness and mortality. *The Science of the Total Environment.* 520114-119.

Steinvil, A., Kordova-Biezuner, L., Shapira, I., Berliner, S. & Rogowski, O. (2008) Short-term exposure to air pollution and inflammation-sensitive biomarkers. *Environmental Research.* 106 (1), 51-61.

Steptoe, A. & Brydon, L. (2005) Associations between acute lipid stress responses and fasting lipid levels 3 years later. *Health Psychology : Official Journal of the Division of Health Psychology, American Psychological Association.* 24 (6), 601-607.

Stockfelt, L., Andersson, E. M., Molnar, P., Rosengren, A., Wilhelmsen, L., Sallsten, G. & Barregard, L. (2015) Long term effects of residential NO(x) exposure on total and cause-specific mortality and incidence of myocardial infarction in a Swedish cohort. *Environmental Research.* 142197-206.

Sullivan, J. H., Hubbard, R., Liu, S. L., Shepherd, K., Trenga, C. A., Koenig, J. Q., Chandler, W. L. & Kaufman, J. D. (2007) A community study of the effect of particulate matter on blood measures of inflammation and thrombosis in an elderly population. *Environmental Health : A Global Access Science Source.* 63.

Sun, Q., Wang, A., Jin, X., Natanzon, A., Duquaine, D., Brook, R. D., Aguinaldo, J. G., Fayad, Z. A., Fuster, V., Lippmann, M., Chen, L. C. & Rajagopalan, S. (2005) Long-term air pollution exposure and acceleration of atherosclerosis and vascular inflammation in an animal model. *Jama.* 294 (23), 3003-3010.

Sun, Q., Yue, P., Deiuliis, J. A., Lumeng, C. N., Kampfrath, T., Mikolaj, M. B., Cai, Y., Ostrowski, M. C., Lu, B., Parthasarathy, S., Brook, R. D., Moffatt-Bruce, S. D., Chen, L. C. & Rajagopalan, S. (2009) Ambient air pollution exaggerates adipose inflammation and insulin resistance in a mouse model of diet-induced obesity. *Circulation.* 119 (4), 538-546.

Suwa, T., Hogg, J. C., Quinlan, K. B., Ohgami, A., Vincent, R. & van Eeden, S. F. (2002) Particulate air pollution induces progression of atherosclerosis. *Journal of the American College of Cardiology.* 39 (6), 935-942.

Tamayo, T., Rathmann, W., Kramer, U., Sugiri, D., Grabert, M. & Holl, R. W. (2014) Is particle pollution in outdoor air associated with metabolic control in type 2 diabetes? *PloS One.* 9 (3), e91639.

Tasali, E., Leproult, R. & Spiegel, K. (2009) Reduced sleep duration or quality: relationships with insulin resistance and type 2 diabetes. *Progress in Cardiovascular Diseases.* 51 (5), 381-391.

Teichert, T., Vossoughi, M., Vierkotter, A., Sugiri, D., Schikowski, T., Schulte, T., Roden, M., Luckhaus, C., Herder, C. & Kramer, U. (2013) Association between traffic-related air pollution, subclinical inflammation and impaired glucose metabolism: results from the SALIA study. *PloS One.* 8 (12), e83042.

Tetreault, L. F., Perron, S. & Smargiassi, A. (2013) Cardiovascular health, traffic-related air pollution and noise: are associations mutually confounded? A systematic review. *International Journal of Public Health.* 58 (5), 649-666.

Thiering, E. & Heinrich, J. (2015) Epidemiology of air pollution and diabetes. *Trends in Endocrinology and Metabolism: TEM.* 26 (7), 384-394.

Tobias, A., Diaz, J., Recio, A. & Linares, C. (2015) Traffic noise and risk of mortality from diabetes. *Acta Diabetologica.* 52 (1), 187-188.

Tomao, E., Tiziana, P. B., Rosati, V., Marcellini, L. & Tomei, F. (2002) The effects of air pollution on the lipid balance of traffic police personnel. *Annals of Saudi Medicine.* 22 (5-6), 287-290.

Tsai, D. H., Amyai, N., Marques-Vidal, P., Wang, J. L., Riediker, M., Mooser, V., Paccaud, F., Waeber, G., Vollenweider, P. & Bochud, M. (2012) Effects of particulate matter on inflammatory markers in the general adult population. *Particle and Fibre Toxicology.* 924-8977-9-24.

Ueda, K., Nagasawa, S. Y., Nitta, H., Miura, K., Ueshima, H. & NIPPON DATA80 Research Group. (2012) Exposure to particulate matter and long-term risk of cardiovascular mortality in Japan: NIPPON DATA80. *Journal of Atherosclerosis and Thrombosis.* 19 (3), 246-254.

UK Committee on the Medical Effects of Air Pollutants. (December 2015) INTERIM STATEMENT ON QUANTIFYING THE ASSOCIATION OF LONG-TERM AVERAGE CONCENTRATIONS OF NITROGEN DIOXIDE AND MORTALITY.

Upadhyay, S., Ganguly, K., Stoeger, T., Semmler-Bhenke, M., Takenaka, S., Kreyling, W. G., Pitz, M., Reitmeir, P., Peters, A., Eickelberg, O., Wichmann, H. E. & Schulz, H. (2010) Cardiovascular and inflammatory effects of intratracheally instilled ambient dust from Augsburg, Germany, in spontaneously hypertensive rats (SHRs). *Particle and Fibre Toxicology.* 727-8977-7-27.

Van de Ker Ckhove. (2016) What factors are associated with noise sensitivity in the UK population? *16-19 November,* .

van Kempen, E. & Babisch, W. (2012) The quantitative relationship between road traffic noise and hypertension: a meta-analysis. *Journal of Hypertension.* 30 (6), 1075-1086.

van Kempen, E. E., Kruize, H., Boshuizen, H. C., Ameling, C. B., Staatsen, B. A. & de Hollander, A. E. (2002) The association between noise exposure and blood pressure and ischemic heart disease: a meta-analysis. *Environmental Health Perspectives.* 110 (3), 307-317.

van Vliet-Ostaptchouk, J. V., Nuotio, M. L., Slagter, S. N., Doiron, D., Fischer, K., Foco, L., Gaye, A., Gogele, M., Heier, M., Hiekkalinna, T., Joensuu, A., Newby, C., Pang, C., Partinen, E., Reischl, E., Schwienbacher, C., Tammesoo, M. L., Swertz, M. A., Burton, P., Ferretti, V., Fortier, I., Giepmans, L., Harris, J. R., Hillege, H. L., Holmen, J., Jula, A., Kootstra-Ros, J. E., Kvaloy, K., Holmen, T. L., Mannisto, S., Metspalu, A., Midthjell, K., Murtagh, M. J., Peters, A., Pramstaller, P. P., Saaristo, T., Salomaa, V., Stolk, R. P., Uusitupa, M., van der Harst, P., van der Klauw, M. M., Waldenberger, M., Perola, M. & Wolffenbuttel, B. H. (2014) The prevalence of metabolic syndrome and metabolically healthy obesity in Europe: a collaborative analysis of ten large cohort studies. *BMC Endocrine Disorders.* 149-6823-14-9.

Vardoulakis, S., Fisher, B. E. A., Pericleous, K. & Gonzalez-Flesca, N. (2003) Modelling air quality in street canyons: a review. *Atmospheric Environment.* 37 (2), 155-182.

Viehmann, A., Hertel, S., Fuks, K., Eisele, L., Moebus, S., Mohlenkamp, S., Nonnemacher, M., Jakobs, H., Erbel, R., Jockel, K. H., Hoffmann, B. & Heinz Nixdorf Recall Investigator Group. (2015) Long-term residential exposure to urban air pollution, and repeated measures of systemic blood markers of inflammation and coagulation. *Occupational and Environmental Medicine.* 72 (9), 656-663.

Vienneau, D., de Hoogh, K., Bechle, M. J., Beelen, R., van Donkelaar, A., Martin, R. V., Millet, D. B., Hoek, G. & Marshall, J. D. (2013) Western European land use regression incorporating satellite- and ground-based measurements of NO2 and PM10. *Environmental Science & Technology.* 47 (23), 13555-13564.

Vienneau, D., Schindler, C., Perez, L., Probst-Hensch, N. & Roosli, M. (2015) The relationship between transportation noise exposure and ischemic heart disease: a meta-analysis. *Environmental Research.* 138372-380.

Visser, M., Bouter, L. M., McQuillan, G. M., Wener, M. H. & Harris, T. B. (1999) Elevated C-reactive protein levels in overweight and obese adults. *Jama.* 282 (22), 2131-2135.

Wallace, S. E., Gaye, A., Shoush, O. & Burton, P. R. (2014) Protecting personal data in epidemiological research: DataSHIELD and UK law. *Public Health Genomics.* 17 (3), 149-157.

Wang, M., Beelen, R., Eeftens, M., Meliefste, K., Hoek, G. & Brunekreef, B. (2012) Systematic evaluation of land use regression models for NO(2). *Environmental Science & Technology.* 46 (8), 4481-4489.

Wang, X. Y., Hu, W. & Tong, S. (2009) Long-term exposure to gaseous air pollutants and cardio-respiratory mortality in Brisbane, Australia. *Geospatial Health.* 3 (2), 257-263.

WHO. (2015) *Economic cost of the health impact of air pollution in Europe: Clean air, health and wealth.*

WHO. (2005) *Principles of characterizing and applying human exposure models.*

Willett, W. C., Blot, W. J., Colditz, G. A., Folsom, A. R., Henderson, B. E. & Stampfer, M. J. (2007) Merging and emerging cohorts: not worth the wait. *Nature.* 445 (7125), 257-258.

Wilson, P. W., D'Agostino, R. B., Levy, D., Belanger, A. M., Silbershatz, H. & Kannel, W. B. (1998) Prediction of coronary heart disease using risk factor categories. *Circulation.* 97 (18), 1837-1847.

Wolfson, M., Wallace, S. E., Masca, N., Rowe, G., Sheehan, N. A., Ferretti, V., LaFlamme, P., Tobin, M. D., Macleod, J., Little, J., Fortier, I., Knoppers, B. M. & Burton, P. R. (2010) DataSHIELD: resolving a conflict in contemporary bioscience--performing a pooled analysis of individual-level data without sharing the data. *International Journal of Epidemiology.* 39 (5), 1372-1382.

Woodfield, R., UK Biobank Stroke Outcomes Group, UK Biobank Follow-up and Outcomes Working Group & Sudlow, C. L. (2015) Accuracy of Patient Self-Report of Stroke: A Systematic Review from the UK Biobank Stroke Outcomes Group. *PloS One.* 10 (9), e0137538.

World Health Organisation Europe. (2013) *Review of evidence on health aspects of air pollution-REVIHAAP project: final technical report.*

World Health Organisation Europe. (2009) *Night-time noise guidelines for Europe.*

Xu, X., Liu, C., Xu, Z., Tzan, K., Zhong, M., Wang, A., Lippmann, M., Chen, L. C., Rajagopalan, S. & Sun, Q. (2011) Long-term exposure to ambient fine particulate pollution induces insulin resistance and mitochondrial alteration in adipose tissue. *Toxicological Sciences : An Official Journal of the Society of Toxicology.* 124 (1), 88-98.

Yap, C., Beverland, I. J., Heal, M. R., Cohen, G. R., Robertson, C., Henderson, D. E., Ferguson, N. S., Hart, C. L., Morris, G. & Agius, R. M. (2012) Association between long-term exposure to air pollution and specific causes of mortality in Scotland. *Occupational and Environmental Medicine.* 69 (12), 916-924.

Yeatts, K., Svendsen, E., Creason, J., Alexis, N., Herbst, M., Scott, J., Kupper, L., Williams, R., Neas, L., Cascio, W., Devlin, R. B. & Peden, D. B. (2007) Coarse particulate matter (PM2.5-10) affects heart

rate variability, blood lipids, and circulating eosinophils in adults with asthma. *Environmental Health Perspectives.* 115 (5), 709-714.

Yin, F., Lawal, A., Ricks, J., Fox, J. R., Larson, T., Navab, M., Fogelman, A. M., Rosenfeld, M. E. & Araujo, J. A. (2013) Diesel exhaust induces systemic lipid peroxidation and development of dysfunctional pro-oxidant and pro-inflammatory high-density lipoprotein. *Arteriosclerosis, Thrombosis, and Vascular Biology.* 33 (6), 1153-1161.

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

Zeka, A., Sullivan, J. R., Vokonas, P. S., Sparrow, D. & Schwartz, J. (2006) Inflammatory markers and particulate air pollution: characterizing the pathway to disease. *International Journal of Epidemiology.* 35 (5), 1347-1354.

Zhang, L. W., Chen, X., Xue, X. D., Sun, M., Han, B., Li, C. P., Ma, J., Yu, H., Sun, Z. R., Zhao, L. J., Zhao, B. X., Liu, Y. M., Chen, J., Wang, P. P., Bai, Z. P. & Tang, N. J. (2014) Long-term exposure to high particulate matter pollution and cardiovascular mortality: a 12-year cohort study in four cities in northern China. *Environment International.* 6241-47.

Zhou, M., Liu, Y., Wang, L., Kuang, X., Xu, X. & Kan, H. (2014) Particulate air pollution and mortality in a cohort of Chinese men. *Environmental Pollution (Barking, Essex : 1987).* 1861-6.

Zhu, Y., Hinds, W. C., Kim, S. & Sioutas, C. (2002) Concentration and size distribution of ultrafine particles near a major highway. *Journal of the Air & Waste Management Association (1995).* 52 (9), 1032-1042.

# Appendices

## 1. Chapter 5

**Appendix-5.1 Main analyses by exposure and by outcome: Model1 & Model2**

| | | **$PM_{10}$ , µg/$m^3$, per IQR** | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | N | Model1 | | | N | Model2 | | |
| **Pooled analysis** | | ES | 95%CI | | | ES | 95%CI | |
| **hsCRP** | 87622 | 1.001 | 0.991 | 1.012 | 87622 | 1.013 | 1.002 | 1.024 |
| **Total cholesterol** | 110836 | -0.060 | -0.068 | -0.052 | 110836 | -0.021 | -0.029 | -0.013 |
| **HDL cholesterol** | 110834 | 0.003 | 0.000 | 0.006 | 110834 | 0.004 | 0.001 | 0.006 |
| **Triglycerides** | 111547 | -0.001 | -0.008 | 0.006 | 111547 | 0.015 | 0.009 | 0.022 |
| | | | | | | | | |
| **LifeLines only** | | | | | | | | |
| **Fasting blood glucose** | 59898 | -0.015 | -0.024 | -0.006 | 59898 | 0.029 | 0.020 | 0.037 |
| **HbA1c** | 47628 | -0.323 | -0.380 | -0.266 | 47628 | 0.018 | -0.037 | 0.072 |
| | | **$NO_2$, µg/$m^3$, per IQR** | | | | | | |
| | | Model1 | | | | Model2 | | |
| **Pooled analysis** | | ES | 95%CI | | | ES | 95%CI | |
| **hsCRP** | 87957 | 0.999 | 0.987 | 1.010 | 87957 | 1.010 | 0.999 | 1.022 |
| **Total cholesterol** | 111180 | -0.062 | -0.071 | -0.053 | 111180 | -0.017 | -0.026 | -0.009 |
| **HDL cholesterol** | 111178 | 0.008 | 0.004 | 0.011 | 111178 | 0.007 | 0.004 | 0.011 |
| **Triglycerides** | 111893 | -0.004 | -0.012 | 0.003 | 111893 | 0.016 | 0.009 | 0.024 |
| | | | | | | | | |
| **LifeLines only** | | | | | | | | |
| **Fasting blood glucose** | 60182 | -0.021 | -0.030 | -0.011 | 60182 | 0.030 | 0.020 | 0.039 |
| **HbA1c** | 47686 | -0.398 | -0.462 | -0.334 | 47686 | -0.014 | -0.075 | 0.047 |

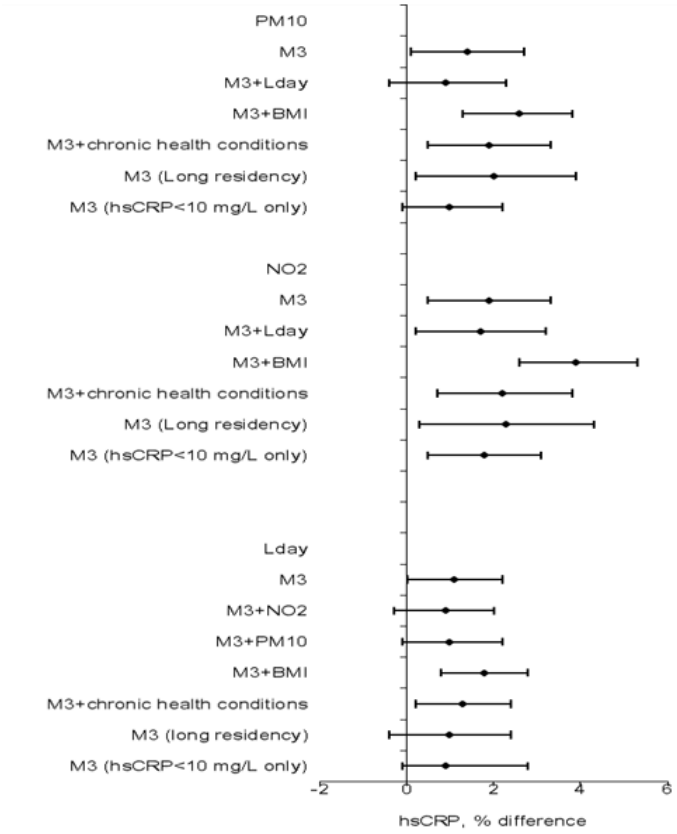
| | | LDAY, dB(A), per IQR | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | Model1 | | | | Model2 | | |
| **Pooled analysis** | | ES | 95%CI | | | ES | 95%CI | |
| **hsCRP** | 90689 | 1.007 | 0.999 | 1.015 | 90689 | 1.004 | 1.001 | 1.008 |
| **Total cholesterol** | 118868 | -0.022 | -0.029 | -0.015 | 118868 | -0.004 | -0.010 | 0.003 |
| **HDL cholesterol** | 118866 | 0.004 | 0.002 | 0.007 | 118866 | 0.005 | 0.003 | 0.008 |
| **Triglycerides** | 119464 | 0.004 | -0.002 | 0.010 | 119464 | 0.027 | 0.015 | 0.039 |
| | | | | | | | | |
| **LifeLines only** | | | | | | | | |
| **Fasting blood glucose** | 72401 | -0.009 | -0.015 | -0.002 | 72401 | 0.011 | 0.005 | 0.017 |
| **HbA1c** | 56026 | -0.163 | -0.204 | -0.122 | 56026 | -0.011 | -0.050 | 0.028 |

Model1: adjusted for study (in pooled analyses on hsCRP and blood lipids) or unadjusted (in LifeLines analyses on blood glucose and HbA1c); Model2: adjusted for (study), age and sex. IQR for $PM_{10}$ and $NO_2$ was 2 and 7.4 $\mu g/m^3$ for analyses of hsCRP and blood lipids; IQR for $PM_{10}$ and $NO_2$ was 2.4 and 8.8 $\mu g/m^3$ for analyses of fasting glucose and HbA1c. Unit HsCRP (mg/L), total cholesterol (mmol/l), HDL cholesterol (mmol/l), triglycerides (mmol/l), fasting glucose (mmol/l), HbA1c (mmol/mol)

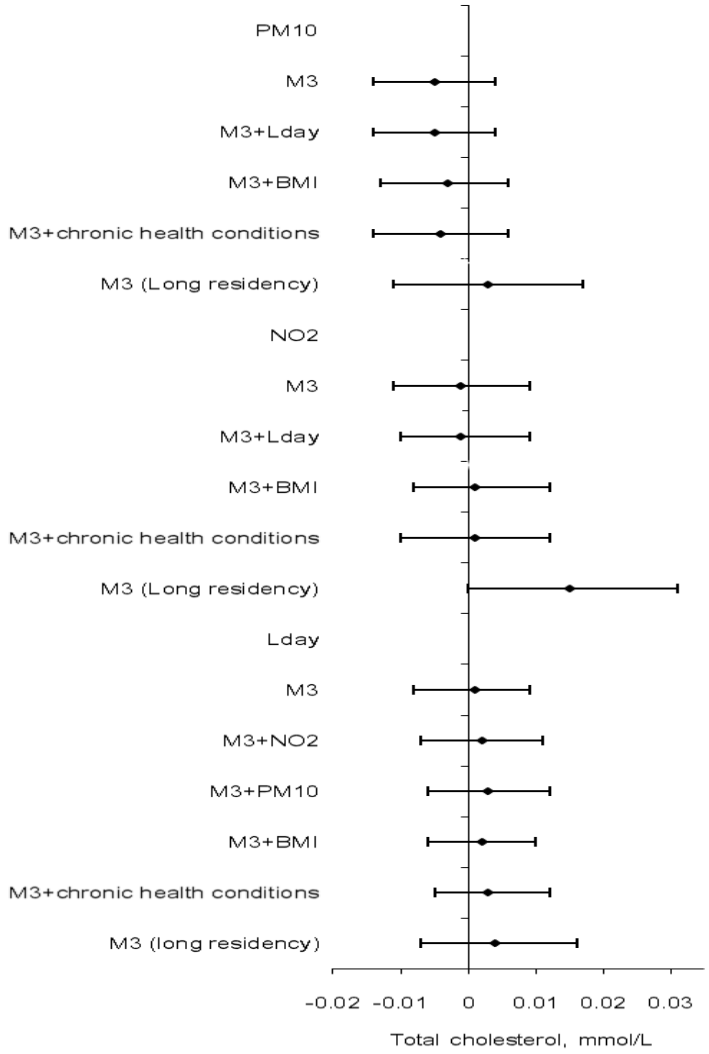
**Appendix-5.2 Sensitivity analyses based on Model 3 (M3) by outcome and by exposure**

For hsCRP and blood lipids (pooled analyses), effect estimate expressed by per each 7.4 µg/$m^3$ increase in $NO_2$, per each 2.0 µg/$m^3$ increase in $PM_{10}$, per each 5.1 dB(A) increase in Lday, Sensitivity analyses based on Model 3 (M3), adjusted for age, sex, education, employment, smoking status, smoking pack-years, alcohol consumption, and study. BMI: body mass index; Chronic health conditions: self-reported ever-had diabetes and hypertension; Long residency: living at the same address for more than 10 years at recruitment.

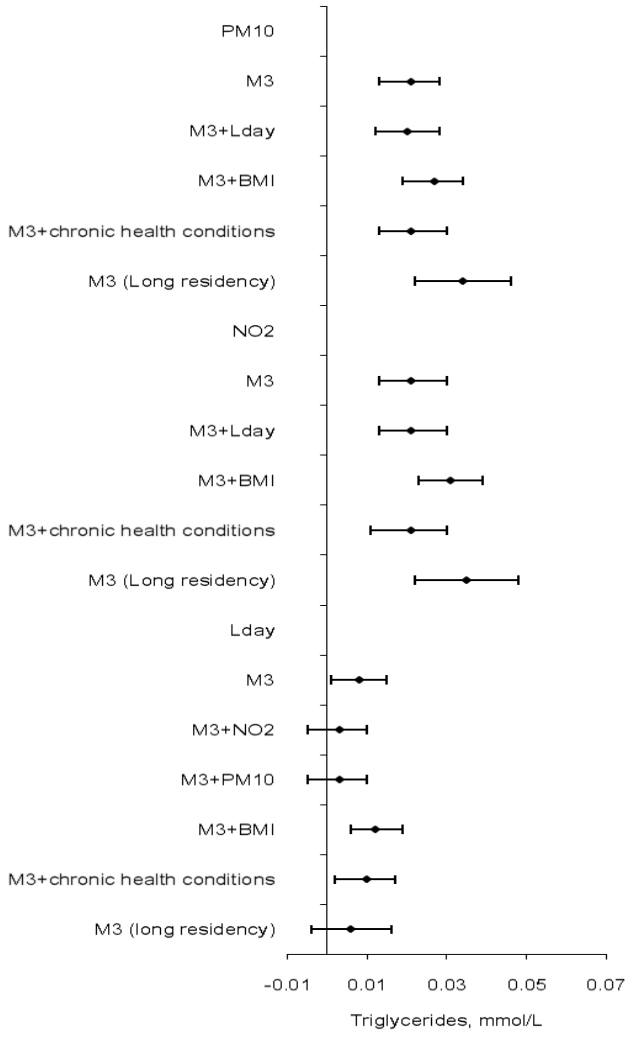
- HsCRP



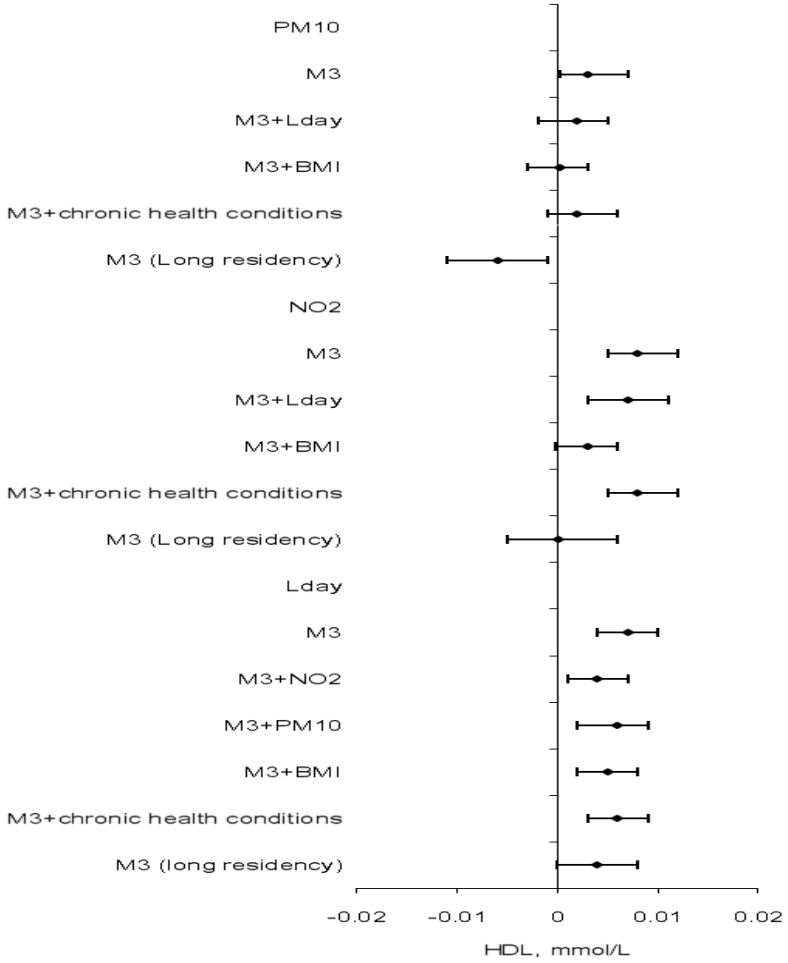
- Total cholesterol




Triglycerides

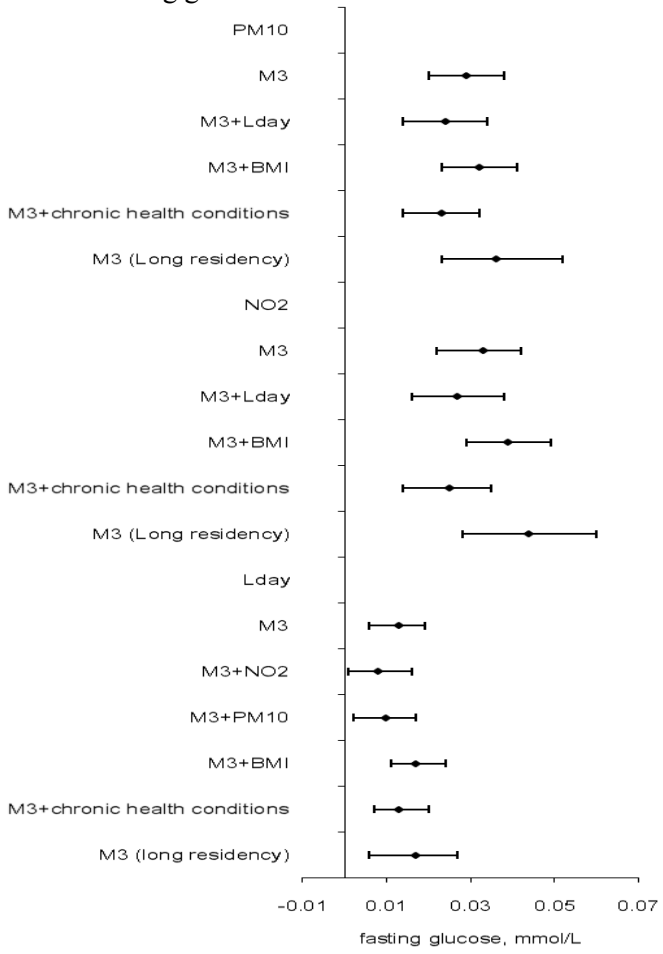


- HDL cholesterol

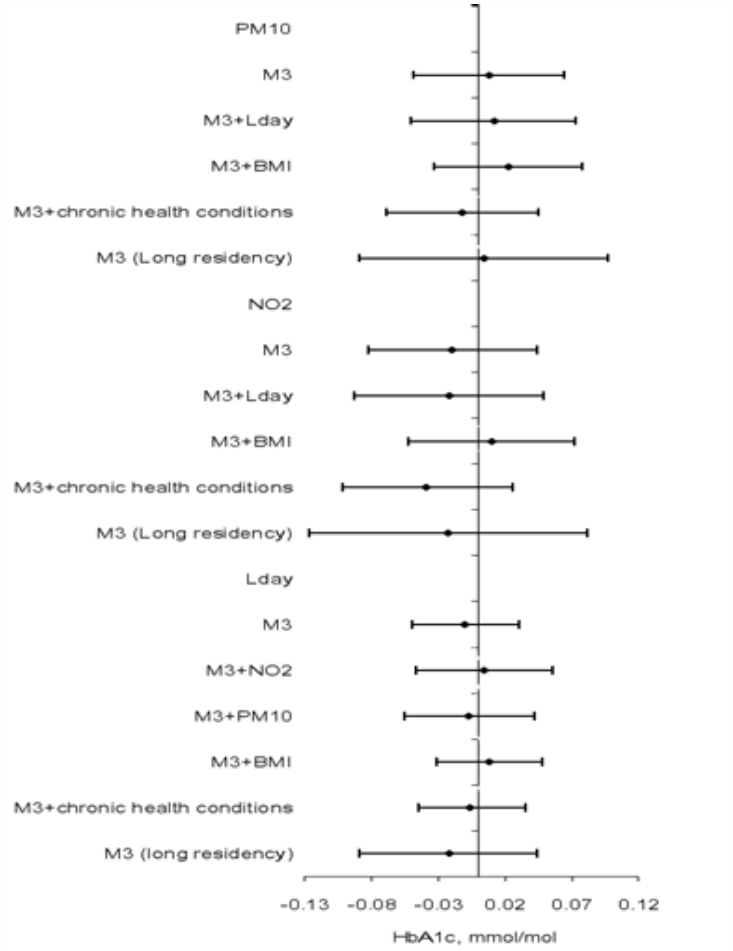


**For analyses of fasting glucose and HbA1c, effect estimate expressed by per each 8.8 µg/m$^3$ increase in $NO_2$, per each 2.4 µg/m$^3$ increase in $PM_{10}$, per each 4.2 dB(A) increase in Lday.**

- Fasting glucose




HbA1c



**Appendix-5.3 Subgroup analyses based on Model 3 (M3) by outcome and by exposure**

**For hsCRP and blood lipids (pooled analyses), effect estimate expressed by per each 7.4 μg/m$^3$ increase in $NO_2$, per each 2.0 μg/m$^3$ increase in $PM_{10}$, per each 5.1 dB(A) increase in Lday.**

| **hsCRP** | **$NO_2$** | | | | | **$PM_{10}$** | | | | | **Lday** | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | N | ES | lo95 | up95 | P* | N | ES | lo95 | up95 | P* | N | ES | lo95 | up95 | P* |
| Men | 22221 | 1.05% | -0.93% | 3.15% | 0.00 | 22116 | 1.03% | -0.84% | 2.94% | 0.00 | 24053 | 0.87% | -0.67% | 2.47% | 0.00 |
| Women | 29238 | 2.82% | 0.92% | 4.77% | | 29122 | 1.93% | 0.19% | 3.70% | | 31877 | 1.27% | -0.20% | 2.73% | |
| <60y | 39623 | 1.78% | 0.26% | 3.34% | 0.00 | 39453 | 0.90% | -0.52% | 2.56% | 0.00 | 44184 | 1.14% | -0.13% | 2.40% | 0.00 |
| >=60y | 11836 | 0.59% | -2.59% | 3.92% | | 11785 | 1.73% | -1.42% | 4.90% | | 11746 | 0.54% | -1.55% | 2.73% | |
| BMI<25 | 20229 | 3.99% | 1.91% | 6.12% | 0.48 | 20143 | 2.88% | 0.97% | 4.78% | 0.06 | 21906 | 2.53% | 0.80% | 4.25% | 0.95 |
| BMI: 25-30 | 21778 | 3.15% | 1.12% | 5.28% | | 21685 | 1.67% | -0.26% | 3.64% | | 23571 | 1.20% | -0.34% | 2.80% | |
| BMI>=30 | 9360 | 3.86% | 0.59% | 7.20% | | 9318 | 2.75% | -0.26% | 5.84% | | 10362 | 0.67% | -1.48% | 2.93% | |
| Diabetes | 1679 | 6.50% | -2.73% | 16.56% | 0.01 | 1673 | 8.88% | 0.04% | 18.44% | 0.00 | 1757 | 3.33% | -2.70% | 9.64% | 0.02 |
| No diabetes | 49726 | 1.78% | 0.40% | 3.21% | | 49511 | 1.22% | -0.06% | 2.56% | | 54105 | 1.00% | -0.13% | 2.07% | |
| Hypertension | 11621 | 6.69% | 3.47% | 9.93% | 0.53 | 11574 | 5.90% | 2.94% | 8.88% | 0.09 | 12577 | 0.60% | -1.62% | 2.93% | 0.03 |
| no Hypertension | 35081 | 1.12% | -0.66% | 2.89% | | 34935 | 0.77% | -0.84% | 2.37% | | 37683 | 1.60% | 0.27% | 2.93% | |

P*: p-value for interaction

| **Total cholesterol** | **$NO_2$** | | | | | **$PM_{10}$** | | | | | **Lday** | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | N | ES | lo95 | up95 | P* | N | ES | lo95 | up95 | P* | N | ES | lo95 | up95 | P* |
| Men | 31278 | -0.029 | -0.045 | -0.013 | 0.00 | 31170 | -0.026 | -0.040 | -0.011 | 0.00 | 34961 | -0.009 | -0.022 | 0.003 | 0.00 |
| Women | 41505 | 0.020 | 0.008 | 0.033 | | 41381 | 0.012 | 0.001 | 0.024 | | 46629 | 0.011 | 0.001 | 0.022 | |
| <60y | 60225 | 0.017 | 0.006 | 0.027 | 0.00 | 60046 | 0.012 | 0.003 | 0.021 | 0.00 | 68967 | 0.008 | -0.0002 | 0.017 | 0.00 |
| >=60y | 12558 | 0.013 | -0.018 | 0.045 | | 12505 | -0.003 | -0.034 | 0.027 | | 12623 | 0.008 | -0.014 | 0.029 | |
| BMI<25 | 31197 | 0.019 | 0.006 | 0.032 | 0.00 | 31106 | 0.019 | 0.007 | 0.031 | 0.00 | 34636 | 0.014 | 0.003 | 0.026 | 0.00 |
| BMI: 25-30 | 29321 | -0.007 | -0.024 | 0.010 | | 29225 | -0.017 | -0.032 | -0.001 | | 32865 | 0.002 | -0.011 | 0.016 | |
| BMI>=30 | 12170 | 0.019 | -0.009 | 0.048 | | 12125 | 0.002 | -0.024 | 0.028 | | 13996 | -0.005 | -0.026 | 0.015 | |
| Diabetes | 1994 | 0.036 | -0.040 | 0.111 | 0.00 | 1988 | -0.004 | -0.072 | 0.065 | 0.00 | 2151 | -0.010 | -0.064 | 0.043 | 0.19 |
| No diabetes | 70718 | 0.001 | -0.009 | 0.011 | | 70492 | -0.003 | -0.012 | 0.006 | | 79348 | 0.003 | -0.005 | 0.011 | |
| Hypertension | 15545 | 0.008 | -0.017 | 0.033 | 0.00 | 15495 | -0.009 | -0.031 | 0.014 | 0.00 | 17343 | 0.004 | -0.015 | 0.023 | 0.00 |
| no Hypertension | 48990 | 0.002 | -0.010 | 0.014 | | 48838 | 0.001 | -0.010 | 0.012 | | 54398 | 0.005 | -0.004 | 0.015 | |

P*: p-value for interaction

| **Triglycerides** | **$NO_2$** | | | | | **$PM_{10}$** | | | | | **Lday** | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | N | ES | lo95 | up95 | P* | N | ES | lo95 | up95 | P* | N | ES | lo95 | up95 | P* |
| Men | 31373 | 0.019 | 0.003 | 0.035 | 0.28 | 31265 | 0.022 | 0.008 | 0.037 | 0.14 | 35045 | 0.0002 | -0.013 | 0.013 | 0.63 |
| Women | 41653 | 0.024 | 0.016 | 0.033 | | 41529 | 0.021 | 0.014 | 0.029 | | 46754 | 0.015 | 0.008 | 0.022 | |
| <60y | 60389 | 0.025 | 0.016 | 0.034 | 0.00 | 60210 | 0.023 | 0.015 | 0.032 | 0.00 | 69109 | 0.011 | 0.0035 | 0.018 | 0.00 |
| >=60y | 12637 | 0.031 | 0.007 | 0.055 | | 12584 | 0.038 | 0.015 | 0.061 | | 12690 | 0.005 | -0.011 | 0.021 | |
| BMI<25 | 31279 | 0.025 | 0.017 | 0.033 | 0.00 | 31188 | 0.020 | 0.013 | 0.028 | 0.00 | 34704 | 0.011 | 0.004 | 0.018 | 0.00 |
| BMI: 25-30 | 29432 | 0.031 | 0.016 | 0.046 | | 29336 | 0.024 | 0.011 | 0.038 | | 32962 | 0.011 | -0.0001 | 0.023 | |
| BMI>=30 | 12217 | 0.060 | 0.031 | 0.089 | | 12172 | 0.062 | 0.035 | 0.088 | | 14036 | 0.010 | -0.010 | 0.031 | |
| Diabetes | 2006 | 0.042 | -0.043 | 0.128 | 0.60 | 2000 | 0.070 | -0.008 | 0.148 | 0.65 | 2163 | -0.007 | -0.064 | 0.050 | 0.27 |
| No diabetes | 70949 | 0.020 | 0.012 | 0.029 | | 70723 | 0.019 | 0.011 | 0.026 | | 79545 | 0.008 | 0.001 | 0.015 | |
| Hypertension | 15597 | 0.037 | 0.015 | 0.059 | 0.80 | 15547 | 0.041 | 0.021 | 0.062 | 0.89 | 17385 | 0.016 | -0.001 | 0.033 | 0.22 |
| no Hypertension | 49181 | 0.018 | 0.008 | 0.027 | | 49029 | 0.018 | 0.009 | 0.026 | | 54565 | 0.009 | 0.001 | 0.016 | |

P*: p-value for interaction

| **HDL** | **$NO_2$** | | | | | **$PM_{10}$** | | | | | **Lday** | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | N | ES | lo95 | up95 | P* | N | ES | lo95 | up95 | P* | N | ES | lo95 | up95 | P* |
| Men | 31278 | -0.0002 | -0.005 | 0.005 | 0.00 | 31170 | -0.003 | -0.007 | 0.002 | 0.00 | 34961 | 0.004 | 0.001 | 0.008 | 0.00 |
| Women | 41505 | 0.013 | 0.008 | 0.018 | | 41381 | 0.006 | 0.002 | 0.011 | | 46629 | 0.008 | 0.004 | 0.012 | |
| <60y | 60225 | 0.011 | 0.007 | 0.015 | 0.32 | 60046 | 0.005 | 0.002 | 0.009 | 0.09 | 68967 | 0.009 | 0.0059 | 0.012 | 0.18 |
| >=60y | 12558 | 0.005 | -0.006 | 0.015 | | 12505 | 0.001 | -0.009 | 0.011 | | 12623 | 0.003 | -0.004 | 0.010 | |
| BMI<25 | 31197 | 0.006 | 0.001 | 0.011 | 0.03 | 31106 | 0.003 | -0.001 | 0.008 | 0.00 | 34636 | 0.008 | 0.004 | 0.013 | 0.03 |
| BMI: 25-30 | 29321 | 0.002 | -0.003 | 0.007 | | 29225 | -0.0002 | -0.005 | 0.005 | | 32865 | 0.002 | -0.0018 | 0.007 | |
| BMI>=30 | 12170 | 0.003 | -0.005 | 0.011 | | 12125 | -0.005 | -0.012 | 0.003 | | 13996 | 0.007 | 0.001 | 0.012 | |
| Diabetes | 1994 | 0.023 | -0.002 | 0.047 | 0.35 | 1988 | 0.013 | -0.009 | 0.035 | 0.03 | 2151 | 0.002 | -0.015 | 0.020 | 0.04 |
| No diabetes | 70718 | 0.009 | 0.005 | 0.012 | | 70492 | 0.004 | 0.000 | 0.007 | | 79348 | 0.008 | 0.005 | 0.011 | |
| Hypertension | 15545 | 0.006 | -0.003 | 0.014 | 0.32 | 15495 | -0.005 | -0.012 | 0.003 | 0.12 | 17343 | 0.005 | -0.002 | 0.011 | 0.42 |
| no Hypertension | 48890 | 0.009 | 0.005 | 0.014 | | 48838 | 0.004 | 0.0002 | 0.008 | | 54398 | 0.007 | 0.003 | 0.010 | |

P*: p-value for interaction

**For analyses of fasting glucose and HbA1c, effect estimate expressed by per each 8.8 μg/m$^3$ increase in $NO_2$, per each 2.4 μg/m$^3$ increase in $PM_{10}$, per each 4.2 dB(A) increase in Lday.**

| **Fasting Glucose** | **$NO_2$** | | | | | **$PM_{10}$** | | | | | **Lday** | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | N | ES | lo95 | up95 | P* | N | ES | lo95 | up95 | P* | N | ES | lo95 | up95 | P* |
| Men | 22009 | 0.020 | 0.002 | 0.038 | 0.02 | 21908 | 0.020 | 0.004 | 0.036 | 0.04 | 26349 | 0.009 | -0.002 | 0.020 | 0.32 |
| Women | 30444 | 0.042 | 0.030 | 0.054 | | 30326 | 0.037 | 0.026 | 0.048 | | 36416 | 0.015 | 0.008 | 0.023 | |
| <60y | 47744 | 0.034 | 0.024 | 0.044 | 0.03 | 47577 | 0.028 | 0.019 | 0.037 | 0.00 | 57144 | 0.013 | 0.007 | 0.020 | 0.00 |
| >=60y | 4709 | 0.027 | -0.019 | 0.073 | | 4657 | 0.050 | 0.005 | 0.094 | | 5621 | 0.009 | -0.020 | 0.037 | |
| BMI<25 | 25019 | 0.026 | 0.015 | 0.036 | 0.08 | 24935 | 0.021 | 0.011 | 0.031 | 0.09 | 28920 | 0.012 | 0.005 | 0.019 | 0.02 |
| BMI: 25-30 | 20056 | 0.043 | 0.027 | 0.059 | | 19962 | 0.037 | 0.022 | 0.052 | | 24302 | 0.014 | 0.004 | 0.024 | |
| BMI>=30 | 7364 | 0.039 | -0.003 | 0.082 | | 7323 | 0.026 | -0.013 | 0.066 | | 9528 | 0.019 | -0.006 | 0.044 | |
| Diabetes | 954 | 0.073 | -0.199 | 0.345 | 0.05 | 949 | 0.086 | -0.160 | 0.332 | 0.02 | 1208 | 0.098 | -0.067 | 0.262 | 0.00 |
| No diabetes | 51435 | 0.026 | 0.018 | 0.033 | | 51221 | 0.024 | 0.017 | 0.031 | | 61473 | 0.010 | 0.005 | 0.015 | |
| Hypertension | 10523 | 0.047 | 0.015 | 0.080 | 0.26 | 10474 | 0.033 | 0.003 | 0.063 | 0.84 | 12765 | 0.021 | 0.001 | 0.041 | 0.48 |
| no Hypertension | 33894 | 0.026 | 0.015 | 0.038 | | 33754 | 0.026 | 0.015 | 0.036 | | 40400 | 0.012 | 0.005 | 0.020 | |

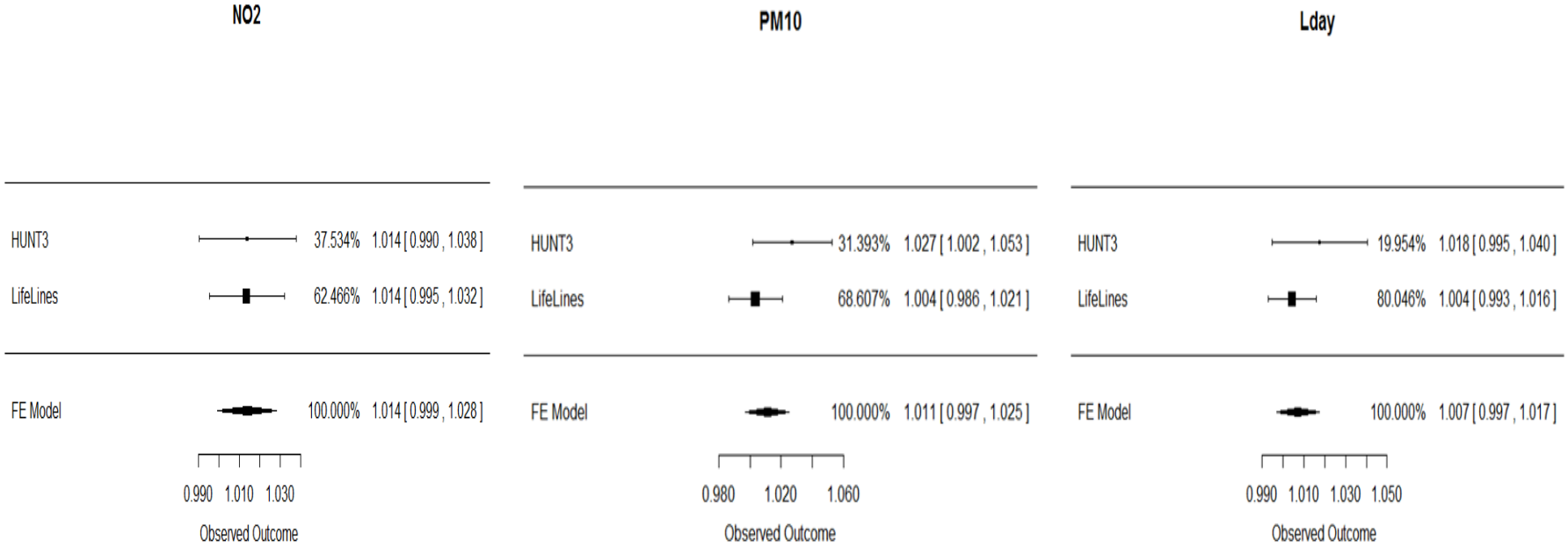
P*: p-value for interaction

| **HbA1c** | **$NO_2$** | | | | | **$PM_{10}$** | | | | | **Lday** | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | N | ES | lo95 | up95 | P* | N | ES | lo95 | up95 | P* | N | ES | lo95 | up95 | P* |
| Men | 18138 | -0.072 | -0.180 | 0.035 | 0.26 | 18115 | 0.020 | -0.075 | 0.116 | 0.63 | 21204 | -0.022 | -0.089 | 0.045 | 0.91 |
| Women | 25399 | 0.029 | -0.048 | 0.106 | | 25366 | 0.013 | -0.056 | 0.082 | | 29710 | 0.005 | -0.044 | 0.054 | |
| <60y | 39901 | -0.039 | -0.104 | 0.026 | 0.59 | 39857 | -0.0003 | -0.058 | 0.058 | 0.75 | 46684 | -0.006 | -0.047 | 0.035 | 0.85 |
| >=60y | 3636 | 0.107 | -0.161 | 0.376 | | 3624 | 0.017 | -0.233 | 0.267 | | 4230 | -0.106 | -0.273 | 0.062 | |
| BMI<25 | 20767 | -0.026 | -0.099 | 0.046 | 0.03 | 20751 | -0.018 | -0.083 | 0.047 | 0.06 | 23621 | 0.012 | -0.035 | 0.058 | 0.24 |
| BMI: 25-30 | 16561 | -0.029 | -0.133 | 0.075 | | 16536 | 0.017 | -0.076 | 0.110 | | 19583 | -0.077 | -0.143 | -0.011 | |
| BMI>=30 | 6196 | -0.012 | -0.268 | 0.244 | | 6181 | -0.029 | -0.259 | 0.201 | | 7697 | 0.058 | -0.090 | 0.206 | |
| Diabetes | 831 | 0.601 | -0.725 | 1.927 | 0.00 | 831 | 0.363 | -0.790 | 1.517 | 0.01 | 1003 | 0.467 | -0.363 | 1.297 | 0.00 |
| No diabetes | 42667 | -0.060 | -0.113 | -0.007 | | 42611 | -0.016 | -0.063 | 0.032 | | 49862 | -0.022 | -0.055 | 0.011 | |
| Hypertension | 8698 | 0.136 | -0.059 | 0.332 | 0.07 | 8685 | 0.092 | -0.082 | 0.267 | 0.27 | 10290 | 0.004 | -0.117 | 0.125 | 0.94 |
| no Hypertension | 28163 | -0.045 | -0.117 | 0.027 | | 28130 | -0.017 | -0.081 | 0.048 | | 32826 | -0.003 | -0.049 | 0.042 | |

P*: p-value for interaction

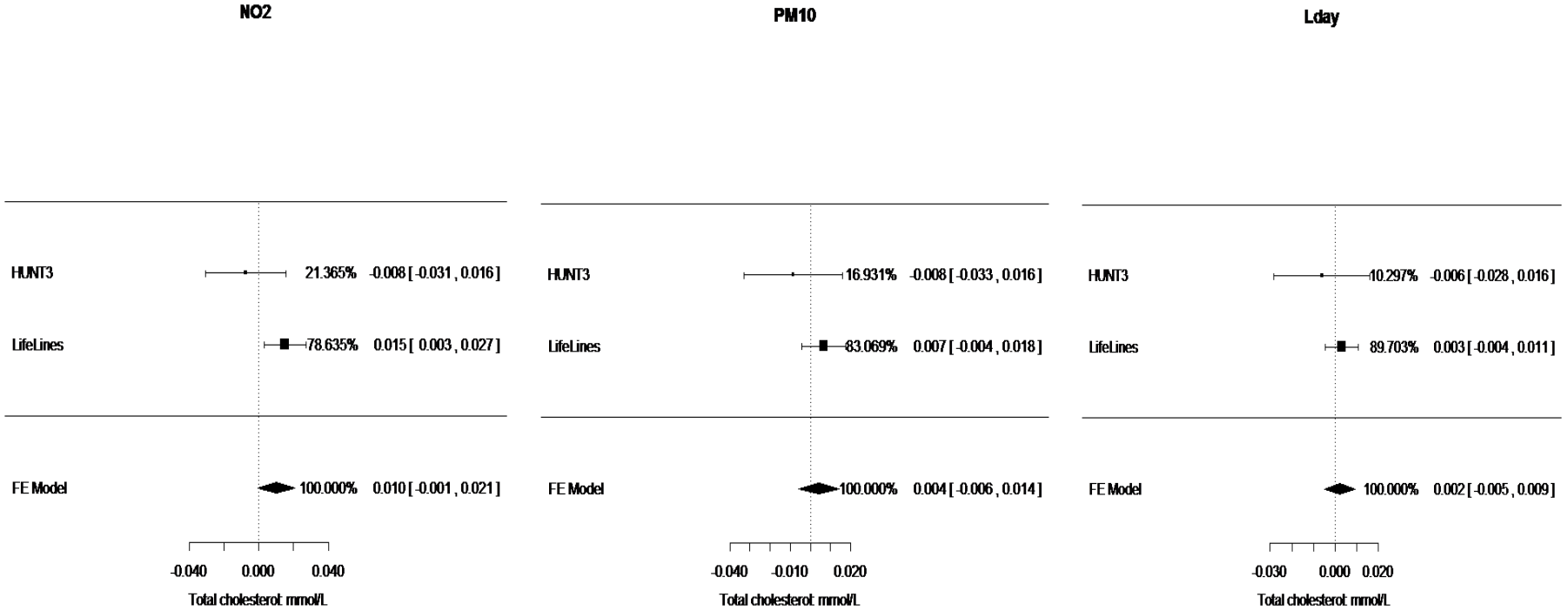
**Appendix-5.4 cohort-specific meta-analysis by outcome and by exposure**

- **hsCRP**



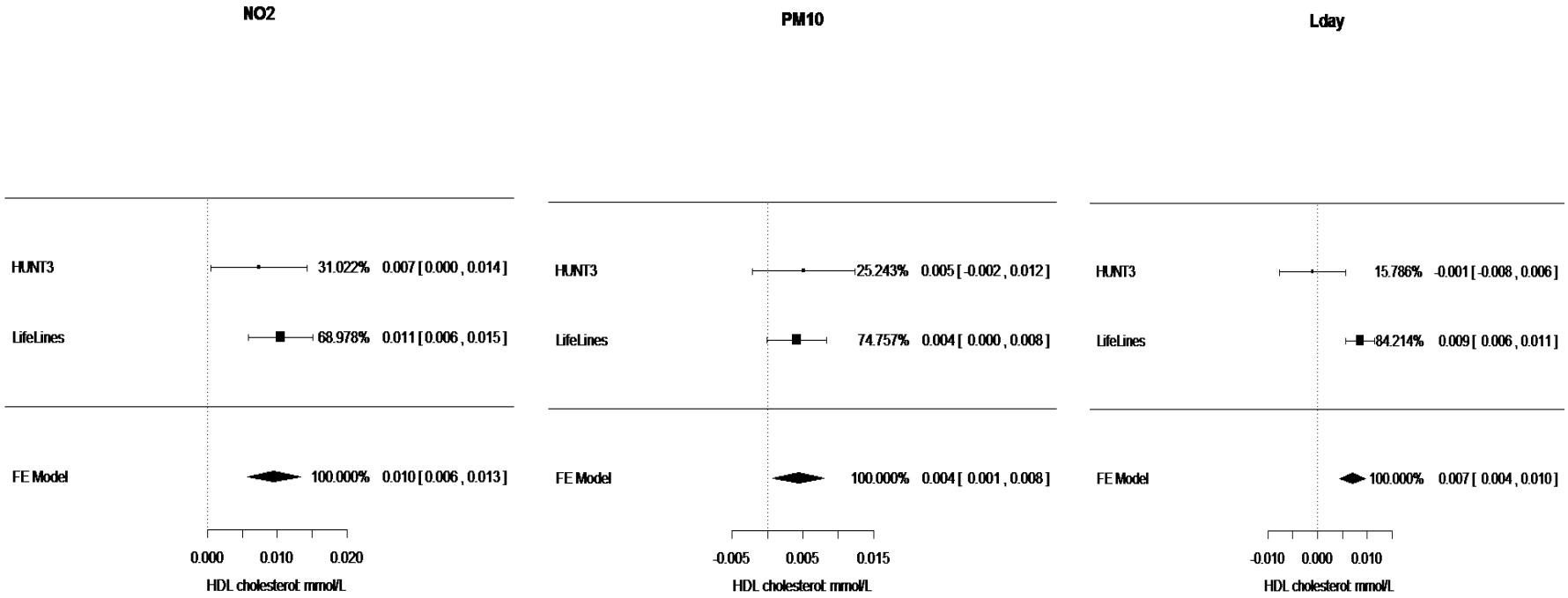

Effect estimates expressed as per IQR (Table 5.3) increase for each exposure in each cohort

- **Total cholesterol**



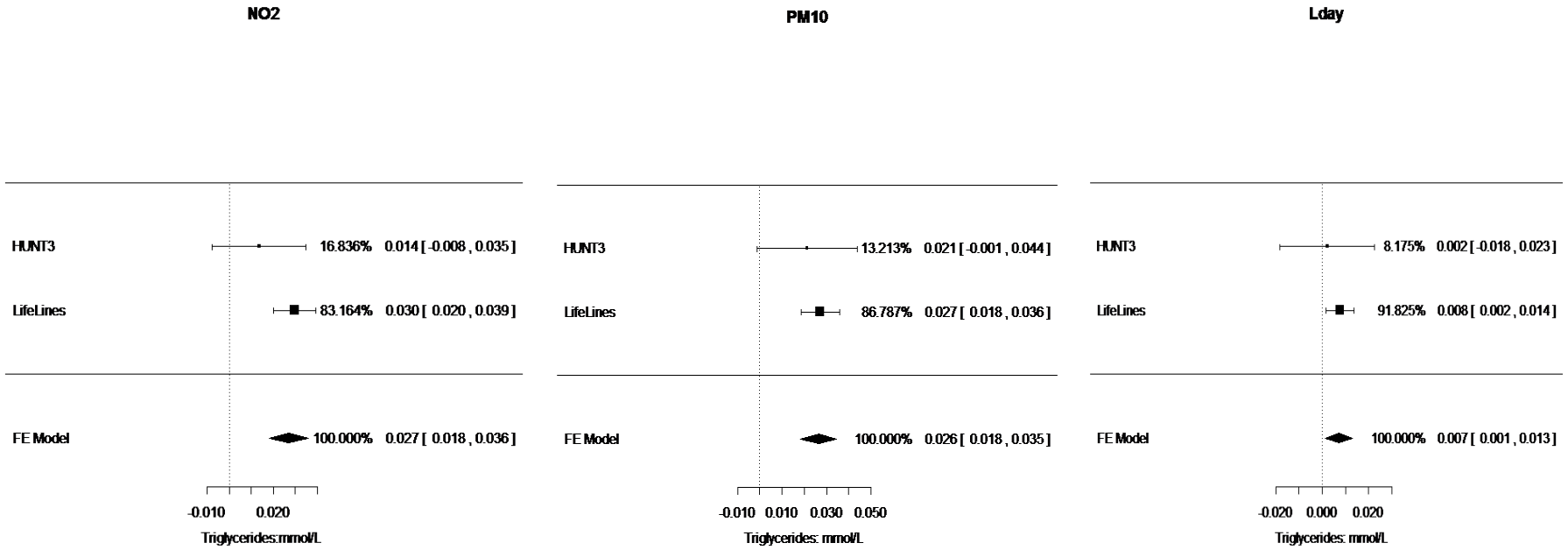

Effect estimates expressed as per IQR (Table 5.3) increase for each exposure in each cohort

- **HDL cholesterol**




Effect estimates expressed as per IQR (Table 5.3) increase for each exposure in each cohort

- **Triglycerides**




Effect estimates expressed as per IQR (Table 5.3) increase for each exposure in each cohort

**Appendix-5.5 ESCAPE LUR modelled air pollutants: analyses in LifeLines only**

| | M3 | M3+LDAY | M3+BMI | M3+Long-term health conditions | M3 (Long residency) |
|---|---|---|---|---|---|
| **hsCRP (N=32266)** | | | | | |
| **$NO_2$** | 1.0% (-1.4% to 3.4%) | 0.5% (-2.6% to 3.6%) | **3.8% (1.5% to 6.1%)** | 1.3% (-1.4% to 4.0%) | 3.8% (0.01% to 7.7%) |
| **$PM_{10}$** | -1.1% (-2.8% to 0.6%) | -2.7% (-4.8% to -0.5%) | 1.4% (-0.2% to 3.1%) | -0.9% (-2.8% to 1.0%) | -0.9% (-3.6% to 1.9%) |
| **$PM_{2.5}$** | 0.03% (-0.8% to 0.8%) | -0.2% (-1.1% to 0.7%) | 0.4% (-0.4% to 1.2%) | 0.1% (-0.8% to 1.0%) | -0.8% (-2.1% to 0.4%) |
| **PMcoarse** | -1.7% (-3.3% to 0%) | -3.1% (-5.0% to -1.1%) | 1.4% (-0.2% to 3.0%) | -1.5% (-3.3% to 0.4%) | 0.2% (-2.5% to 3.0%) |
| **$PM_{2.5absorbance}$** | -0.6% (-2.6% to 1.4%) | -2.2% (-4.8% to 0.4%) | **2.3% (0.4% to 4.2%)** | -0.5% (-2.7% to 1.7%) | -0.8% (-3.9% to 2.4%) |
| **total cholesterol (N=53807)** | | | | | |
| **$NO_2$** | 0.012 (-0.004 to 0.028) | 0.010 (-0.010 to 0.030) | **0.017 (0.001 to 0.033)** | **0.020 (0.003 to 0.037)** | **0.027 (0.0003 to 0.053)** |
| **$PM_{10}$** | 0.011 (-0.001 to 0.022) | 0.010 (-0.004 to 0.025) | **0.016 (0.003 to 0.026)** | **0.016 (0.004 to 0.029)** | 0.021 (-0.0001 to 0.041) |
| **$PM_{2.5}$** | **0.006 (0.001 to 0.011)** | 0.006 (-0.0002 to 0.012) | **0.007 (0.001 to 0.012)** | **0.006 (0.001 to 0.012)** | **0.010 (0.001 to 0.019)** |
| **PMcoarse** | 0.005 (-0.006 to 0.017) | 0.002 (-0.012 to 0.016) | 0.010 (-0.001 to 0.021) | **0.012 (0.0001 to 0.025)** | 0.010 (-0.010 to 0.031) |
| **$PM_{2.5absorbance}$** | **0.017 (0.004 to 0.030)** | **0.021 (0.003 to 0.039)** | **0.021 (0.008 to 0.034)** | **0.022 (0.007 to 0.036)** | **0.030 (0.007 to 0.053)** |
| **HDL cholesterol (N=53807)** | | | | | |
| **$NO_2$** | **0.019 (0.013 to 0.025)** | **0.015 (0.007 to 0.023)** | **0.010 (0.004 to 0.016)** | **0.019 (0.012 to 0.025)** | -0.001 (-0.011 to 0.009) |
| **$PM_{10}$** | **0.019 (0.015 to 0.024)** | **0.021 (0.015 to 0.026)** | **0.012 (0.008 to 0.016)** | **0.019 (0.014 to 0.024)** | **0.012 (0.004 to 0.020)** |
| **$PM_{2.5}$** | **0.005 (0.003 to 0.007)** | **0.003 (0.0005 to 0.005)** | **0.004 (0.002 to 0.006)** | **0.004 (0.002 to 0.006)** | **0.003 (0.0002 to 0.007)** |

| | | | | | |
|---|---|---|---|---|---|
| **PMcoarse** | **0.021 (0.016 to 0.025)** | **0.021 (0.016 to 0.026)** | **0.011 (0.007 to 0.015)** | **0.021 (0.016 to 0.025)** | **0.011 (0.003 to 0.018)** |
| **$PM_{2.5absorbance}$** | **0.021 (0.016 to 0.026)** | **0.023 (0.016 to 0.030)** | **0.013 (0.008 to 0.018)** | **0.020 (0.015 to 0.026)** | **0.013 (0.004 to 0.022)** |
| **Triglycerides (N=53807)** | | | | | |
| **$NO_2$** | **0.035 (0.022 to 0.047)** | **0.055 (0.039 to 0.070)** | **0.048 (0.036 to 0.061)** | **0.036 (0.023 to 0.050)** | **0.061 (0.041 to 0.082)** |
| **$PM_{10}$** | 0.007 (-0.002 to 0.016) | **0.013 (0.001 to 0.024)** | **0.018 (0.010 to 0.027)** | **0.012 (0.003 to 0.022)** | 0.014 (-0.001 to 0.031) |
| **$PM_{2.5}$** | 0.001 (-0.003 to 0.006) | -0.001 (-0.006 to 0.003) | 0.003 (-0.001 to 0.007) | 0.004 (-0.001 to 0.008) | 0.00004 (-0.007 to 0.007) |
| **PMcoarse** | 0.008 (-0.001 to 0.017) | **0.019 (0.008 to 0.029)** | **0.023 (0.014 to 0.031)** | **0.012 (0.002 to 0.022)** | **0.021 (0.006 to 0.037)** |
| **$PM_{2.5absorbance}$** | **0.011 (0.001 to 0.022)** | **0.020 (0.006 to 0.033)** | **0.024 (0.013 to 0.034)** | **0.017 (0.005 to 0.028)** | **0.018 (0.0004 to 0.036)** |
| **fasting blood glucose (N=52453)** | | | | | |
| **$NO_2$** | **0.041 (0.028 to 0.054)** | **0.034 (0.017 to 0.051)** | **0.055 (0.042 to 0.068)** | **0.033 (0.020 to 0.047)** | **0.061 (0.039 to 0.083)** |
| **$PM_{10}$** | **0.023 (0.013 to 0.032)** | **0.014 (0.001 to 0.026)** | **0.035 (0.026 to 0.044)** | **0.023 (0.014 to 0.033)** | **0.048 (0.031 to 0.065)** |
| **$PM_{2.5}$** | **0.006 (0.001 to 0.010)** | 0.001 (-0.004 to 0.006) | **0.007 (0.003 to 0.011)** | **0.007 (0.003 to 0.011)** | **0.011 (0.004 to 0.019)** |
| **PMcoarse** | **0.024 (0.015 to 0.033)** | **0.016 (0.005 to 0.028)** | **0.039 (0.030 to 0.048)** | **0.022 (0.013 to 0.032)** | **0.047 (0.030 to 0.064)** |
| **$PM_{2.5absorbance}$** | **0.024 (0.013 to 0.035)** | 0.011 (-0.004 to 0.026) | **0.037 (0.026 to 0.047)** | **0.025 (0.014 to 0.036)** | **0.048 (0.029 to 0.067)** |
| **HbA1c (N=43537)** | | | | | |
| **$NO_2$** | -0.060 (-0.142 to 0.021) | -0.091 (-0.194 to 0.012) | -0.002 (-0.083 to 0.078) | -0.078 (-0.159 to 0.005) | -0.077 (-0.218 to 0.065) |
| **$PM_{10}$** | 0.014 (-0.044 to 0.072) | 0.029 (-0.048 to 0.105) | **0.065 (0.007 to 0.122)** | 0.024 (-0.035 to 0.083) | -0.002 (-0.112 to 0.109) |
| **$PM_{2.5}$** | **0.028** | **0.037** | **0.034** | **0.032** | 0.032 |

| | | | | | |
|---|---|---|---|---|---|
| | **(0.001 to 0.055)** | **(0.006 to 0.068)** | **(0.007 to 0.061)** | **(0.005 to 0.060)** | (-0.015 to 0.079) |
| **PMcoarse** | -0.027<br>(-0.084 to 0.031) | -0.037<br>(-0.107 to 0.034) | 0.036<br>(-0.021 to 0.093) | -0.021<br>(-0.079 to 0.037) | -0.069<br>(-0.178 to 0.040) |
| **$PM_{2.5absorbance}$** | 0.006<br>(-0.062 to 0.074) | 0.018<br>(-0.075 to 0.111) | 0.057<br>(-0.010 to 0.124) | 0.016<br>(-0.053 to 0.084) | 0.004<br>(-0.119 to 0.128) |

Effect estimates were calculated for each IQR increase of $NO_2$ (7.42 µg/$m^3$), $PM_{10}$ (0.95 µg/$m^3$), $PM_{2.5}$ (0.24µg/$m^3$) and PMcoarse (0.63µg/$m^3$) and $PM_{2.5absorbance}$ (0.22µg/$m^3$). M3: Model 3 adjusted for age, sex, education, employment status, smoking status, smoking pack-years, and alcohol consumption. Lday: day-time noise (07:00-19:00). Bold indicated where significant level<0.05

**Appendix-5.6 Categorical noise estimates in the main analyses**

| **hsCRP, mg/L** | | **M3** | **M3+$NO_2$** | **M3+$PM_{10}$** |
|---|---|---|---|---|
| **Lday, dB(A)** | N | ES (95%CI) | ES (95%CI) | ES (95%CI) |
| **<55** | 40577 | 1 | 1 | 1 |
| **55-60** | 10972 | 0.1%<br>(-2.2% to 2.5%) | -0.7%<br>(-3.3% to 2.0%) | -0.4%<br>(-3.0% to 2.3%) |
| **>=60** | 4381 | **4.2%**<br>**(0.6% to 7.8%)** | 2.6%<br>(-1.3% to 6.7%) | 3.4%<br>(-0.5% to 7.5%) |

| **Total cholesterol, mmol/L** | | **M3** | **M3+$NO_2$** | **M3+$PM_{10}$** |
|---|---|---|---|---|
| **Lday, dB(A)** | N | ES (95%CI) | ES (95%CI) | ES (95%CI) |
| **<55** | 56437 | 1 | 1 | 1 |
| **55-60** | 17976 | 0.001<br>(-0.015 to 0.018) | 0.005<br>(-0.014 to 0.024) | 0.007<br>(-0.012 to 0.025) |
| **>=60** | 7177 | -0.004<br>(-0.028 to 0.020) | 0.001<br>(-0.026 to 0.029) | 0.005<br>(-0.022 to 0.032) |

| **HDL cholesterol, mmol/L** | | **M3** | **M3+$NO_2$** | **M3+$PM_{10}$** |
|---|---|---|---|---|
| **Lday, dB(A)** | N | ES (95%CI) | ES (95%CI) | ES (95%CI) |
| **<55** | 56437 | 1 | 1 | 1 |
| **55-60** | 17976 | **0.012**<br>**(0.006 to 0.018)** | **0.008**<br>**(0.001 to 0.014)** | **0.010**<br>**(0.004 to 0.017)** |
| **>=60** | 7177 | **0.020**<br>**(0.012 to 0.029)** | **0.014**<br>**(0.004 to 0.024)** | **0.019**<br>**(0.009 to 0.029)** |

| **Triglycerides, mmol/L** | | **M3** | **M3+$NO_2$** | **M3+$PM_{10}$** |
|---|---|---|---|---|
| **Lday, dB(A)** | N | ES (95%CI) | ES (95%CI) | ES (95%CI) |
| **<55** | 56630 | 1 | 1 | 1 |
| **55-60** | 17990 | 0.004<br>(-0.010 to 0.018) | -0.002<br>(-0.018 to 0.013) | -0.001<br>(-0.016 to 0.014) |
| **>=60** | 7179 | 0.017<br>(-0.003 to 0.037) | -0.002<br>(-0.025 to 0.021) | 0.0001<br>(-0.022 to 0.023) |

| **Fasting glucose, mmol/L** | | **M3** | **M3+$NO_2$** | **M3+$PM_{10}$** |
|---|---|---|---|---|
| **Lday, dB(A)** | N | ES (95%CI) | ES (95%CI) | ES (95%CI) |
| **<55** | 38984 | 1 | 1 | 1 |
| **55-60** | 16903 | 0.004<br>(-0.010 to 0.017) | -0.001<br>(-0.016 to 0.015) | 0.002<br>(-0.013 to 0.017) |
| **>=60** | 6878 | **0.033**<br>**(0.013 to 0.053)** | 0.018<br>(-0.004 to 0.040) | **0.022**<br>**(0.0001 to 0.044)** |

| **HbA1c, mmol/mol** | | **M3** | **M3+$NO_2$** | **M3+$PM_{10}$** |
|---|---|---|---|---|
| **Lday, dB(A)** | N | ES (95%CI) | ES (95%CI) | ES (95%CI) |
| **<55** | 31464 | 1 | 1 | 1 |
| **55-60** | 13741 | -0.019<br>(-0.106 to 0.068) | -0.005<br>(-0.101 to 0.091) | -0.018<br>(-0.113 to 0.078) |
| **>=60** | 5709 | -0.019<br>(-0.142 to 0.105) | 0.044<br>(-0.095 to 0.183) | 0.017<br>(-0.119 to 0.154) |

Bold indicated where significant level<0.05

## 2. Chapter 6

**Appendix-6.1 Subgroup analyses based on Model 2 (M2) by incident CVD outcome and by exposure**

| | N | Cerebrovascular disease | | | | Ischaemic heart disease | | | | Total CVD | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **$NO_2$** | | HR | 95%ci | | p-interaction | HR | 95%ci | | p-interaction | HR | 95%ci | | p-interaction |
| **men** | 123692 | 1.13 | 0.97 | 1.32 | 0.00 | 0.95 | 0.86 | 1.05 | 0.00 | 1.03 | 0.98 | 1.07 | 0.00 |
| **women** | 159186 | 0.99 | 0.85 | 1.15 | | 1.02 | 0.90 | 1.15 | | 0.98 | 0.94 | 1.02 | |
| **age<60years** | 194421 | 1.07 | 0.92 | 1.24 | 0.00 | 1.01 | 0.92 | 1.12 | 0.00 | 1.00 | 0.96 | 1.03 | 0.00 |
| **age>=60 years** | 88457 | 1.07 | 0.92 | 1.25 | | 0.97 | 0.86 | 1.09 | | 1.04 | 0.99 | 1.10 | |
| **BMI<25kg/$m^2$** | 123362 | 0.99 | 0.86 | 1.15 | 0.70 | 1.00 | 0.88 | 1.12 | 0.00 | 1.00 | 0.97 | 1.04 | 0.00 |
| **BMI: 25-30 kg/$m^2$** | 115157 | 1.12 | 0.93 | 1.35 | | 0.95 | 0.85 | 1.07 | | 0.98 | 0.93 | 1.02 | |
| **BMI>30 kg/$m^2$** | 43295 | 1.27 | 0.94 | 1.72 | | 1.08 | 0.90 | 1.29 | | 1.08 | 1.00 | 1.17 | |
| **never-smoker** | 156407 | 1.09 | 0.94 | 1.27 | 0.00 | 1.09 | 0.98 | 1.23 | 0.00 | 1.02 | 0.98 | 1.06 | 0.00 |
| **ex-smoker** | 91317 | 0.98 | 0.80 | 1.20 | | 0.97 | 0.85 | 1.10 | | 1.00 | 0.95 | 1.05 | |
| **current-smoker** | 35154 | 1.07 | 0.84 | 1.35 | | 0.81 | 0.69 | 0.96 | | 0.93 | 0.87 | 1.00 | |
| **low- education** | 38437 | 1.20 | 0.93 | 1.54 | 0.02 | 0.94 | 0.78 | 1.13 | 0.00 | 1.02 | 0.94 | 1.11 | 0.00 |
| **Medium-education** | 77357 | 1.06 | 0.88 | 1.28 | | 0.92 | 0.80 | 1.06 | | 1.01 | 0.96 | 1.05 | |
| **High-education** | 167084 | 0.99 | 0.84 | 1.16 | | 1.04 | 0.93 | 1.15 | | 0.99 | 0.96 | 1.03 | |
| **had diabetes** | 5184 | 1.38 | 0.80 | 2.38 | 0.10 | 1.04 | 0.74 | 1.47 | 0.00 | 0.99 | 0.84 | 1.17 | 0.00 |
| **not had diabetes** | 277236 | 1.05 | 0.94 | 1.17 | | 0.97 | 0.90 | 1.05 | | 1.00 | 0.97 | 1.02 | |
| | | | | | | | | | | | | | |
| **$PM_{10}$** | | HR | 95%ci | | | HR | 95%ci | | | HR | 95%ci | | |
| **men** | 123425 | 1.12 | 0.93 | 1.34 | 0.00 | 0.99 | 0.89 | 1.11 | 0.00 | **1.06** | **1.01** | **1.12** | 0.00 |
| **women** | 158800 | 1.04 | 0.88 | 1.24 | | 1.01 | 0.88 | 1.16 | | 0.99 | 0.95 | 1.04 | |
| **age<60years** | 193966 | 1.10 | 0.92 | 1.31 | 0.00 | 1.05 | 0.94 | 1.18 | 0.00 | 1.01 | 0.97 | 1.05 | 0.00 |
| **age>=60 years** | 88259 | 1.09 | 0.91 | 1.31 | | 0.97 | 0.85 | 1.11 | | **1.10** | **1.04** | **1.16** | |

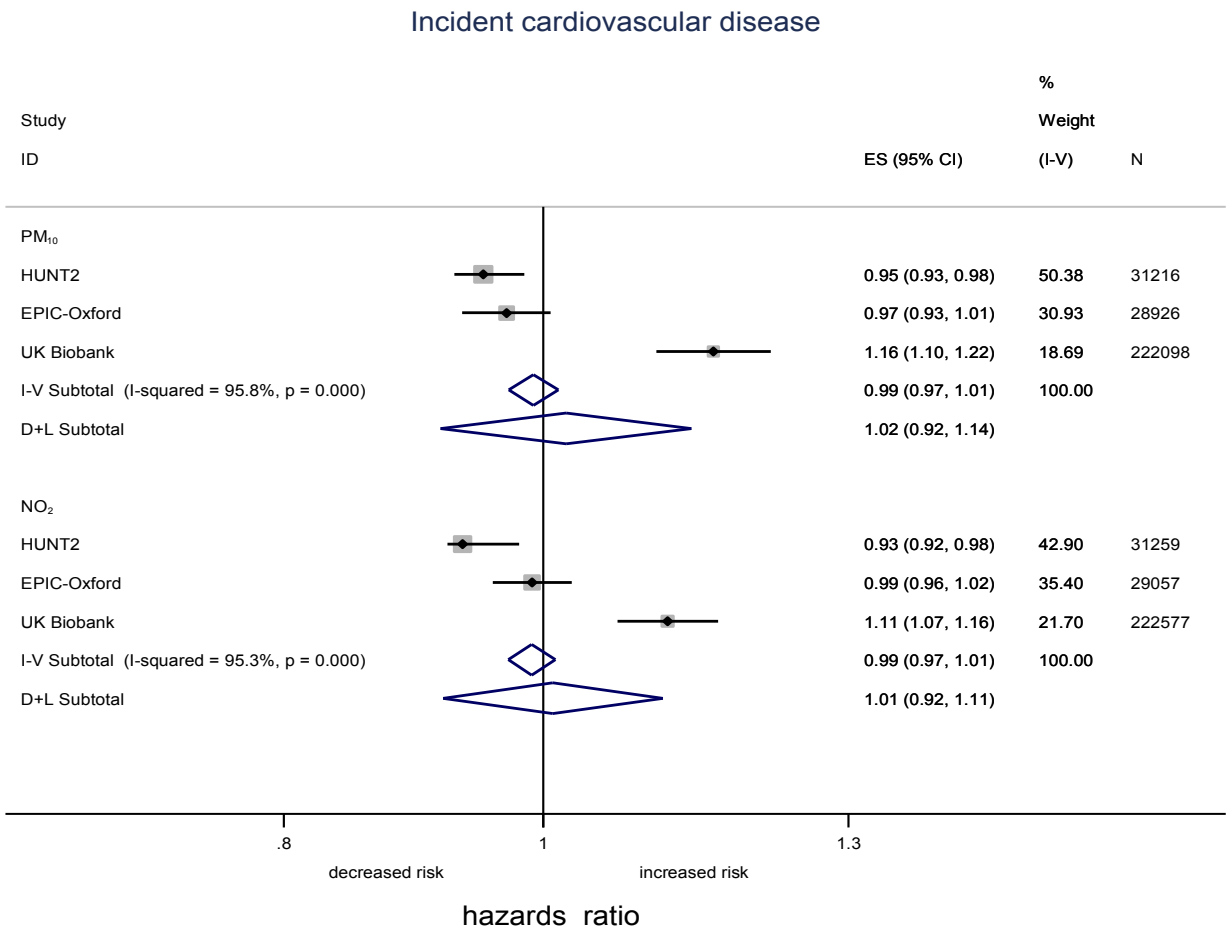
| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **BMI<25kg/m²** | 123075 | 1.05 | 0.88 | 1.24 | 0.60 | 1.00 | 0.87 | 1.15 | 0.00 | 1.01 | 0.97 | 1.06 | 0.00 |
| **BMI: 25-30 kg/m²** | 114899 | 1.10 | 0.89 | 1.37 | | 0.99 | 0.87 | 1.13 | | 1.02 | 0.97 | 1.08 | |
| **BMI>30 kg/m²** | 43188 | 1.23 | 0.85 | 1.77 | | 1.13 | 0.92 | 1.38 | | **1.09** | **1.00** | **1.19** | |
| **never-smoker** | 156048 | 1.03 | 0.86 | 1.23 | 0.00 | 1.13 | 0.99 | 1.29 | 0.00 | 1.03 | 0.98 | 1.08 | 0.00 |
| **ex-smoker** | 91095 | 1.09 | 0.87 | 1.37 | | 1.00 | 0.87 | 1.15 | | 1.03 | 0.97 | 1.09 | |
| **current-smoker** | 35082 | 1.18 | 0.89 | 1.55 | | 0.82 | 0.68 | 0.98 | | 0.98 | 0.90 | 1.06 | |
| **low- education** | 38354 | 1.22 | 0.93 | 1.60 | 0.05 | 0.95 | 0.78 | 1.15 | 0.00 | 1.04 | 0.95 | 1.13 | 0.00 |
| **Medium-education** | 77157 | 1.19 | 0.95 | 1.48 | | 0.90 | 0.77 | 1.06 | | 1.02 | 0.97 | 1.08 | |
| **High-education** | 166714 | 0.94 | 0.78 | 1.14 | | 1.10 | 0.98 | 1.25 | | 1.02 | 0.98 | 1.07 | |
| **had diabetes** | 5173 | 1.44 | 0.72 | 2.90 | 0.15 | 1.10 | 0.73 | 1.66 | 0.00 | 1.06 | 0.87 | 1.29 | 0.00 |
| **not had diabetes** | 276595 | 1.07 | 0.94 | 1.21 | | 0.99 | 0.91 | 1.09 | | 1.02 | 0.99 | 1.05 | |

| **LDAY** | | HR | 95%ci | | | HR | 95%ci | | | HR | 95%ci | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **men** | 121783 | 1.01 | 0.95 | 1.08 | 0.00 | 0.99 | 0.95 | 1.03 | 0.00 | 1.01 | 0.99 | 1.03 | 0.00 |
| **women** | 157084 | 0.98 | 0.91 | 1.04 | | 1.05 | 1.00 | 1.11 | | 1.00 | 0.98 | 1.02 | |
| **age<60years** | 191294 | 1.00 | 0.93 | 1.07 | 0.00 | 1.03 | 0.99 | 1.07 | 0.00 | 1.01 | 0.99 | 1.02 | 0.00 |
| **age>=60 years** | 87573 | 0.99 | 0.93 | 1.06 | | 0.99 | 0.94 | 1.04 | | 1.00 | 0.98 | 1.02 | |
| **BMI<25kg/m²** | 121684 | 0.99 | 0.92 | 1.06 | 0.90 | 0.99 | 0.93 | 1.04 | 0.00 | 1.00 | 0.98 | 1.03 | 0.00 |
| **BMI: 25-30 kg/m²** | 113472 | 0.99 | 0.92 | 1.06 | | 1.01 | 0.96 | 1.06 | | 1.00 | 0.98 | 1.02 | |
| **BMI>30 kg/m²** | 42698 | 1.01 | 0.90 | 1.15 | | 1.06 | 0.98 | 1.14 | | 1.02 | 0.99 | 1.05 | |
| **never-smoker** | 154687 | 0.99 | 0.92 | 1.06 | 0.00 | 1.00 | 0.95 | 1.05 | 0.00 | 0.99 | 0.97 | 1.01 | 0.00 |
| **ex-smoker** | 90227 | 1.00 | 0.92 | 1.09 | | 0.97 | 0.91 | 1.02 | | 1.00 | 0.98 | 1.02 | |
| **current-smoker** | 33953 | 0.99 | 0.90 | 1.08 | | **1.07** | **1.01** | **1.14** | | **1.04** | **1.01** | **1.06** | |
| **low- education** | 37227 | 0.97 | 0.90 | 1.04 | 0.40 | 1.01 | 0.96 | 1.07 | 0.00 | 1.01 | 0.99 | 1.03 | 0.00 |
| **Medium-education** | 75624 | 0.99 | 0.91 | 1.07 | | 1.02 | 0.96 | 1.08 | | 1.01 | 0.99 | 1.04 | |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **High-education** | 166016 | 1.06 | 0.97 | 1.16 | | 1.00 | 0.94 | 1.07 | | 0.99 | 0.97 | 1.02 | |
| **had diabetes** | 5103 | 1.03 | 0.80 | 1.34 | 0.50 | 0.99 | 0.84 | 1.17 | 0.00 | 1.04 | 0.96 | 1.13 | 0.00 |
| **not had diabetes** | 273311 | 0.99 | 0.95 | 1.04 | | 1.01 | 0.98 | 1.04 | | 1.00 | 0.99 | 1.02 | |

Model 2: adjusted for cohort, time segments of follow-up period, age, sex, education, employment status, smoking status and alcohol consumption (main model). HR: Hazard ratios. HR expressed as per 4.1 μg/$m^3$ increase $PM_{10}$, 13.2 μg/$m^3$ increase of $NO_2$ and 3.9 dB(A) increase of Lday. Bold indicated where significant level<0.05

**Appendix-6.2 Cohort-specific meta-analysis of air pollution effects on incident total CVD**



**Appendix-6.3 Cohort-specific meta-analysis of daytime road traffic noise effects on incident total CVD**