Environment International 146 (2021) 106306

Contents lists available at ScienceDirect

Environment International

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

Comparison of associations between mortality and air pollution exposure estimated with a hybrid, a land-use regression and a dispersion model

Jochem O. Klompmaker [a,b,*], Nicole Janssen [a], Zorana J. Andersen [c], Richard Atkinson [d], Mariska Bauwelinck [e], Jie Chen [b], Kees de Hoogh [f,g], Danny Houthuijs [a], Klea Katsouyanni [h,i], Marten Marra [a], Bente Oftedal [j], Sophia Rodopoulou [h], Evangelia Samoli [h], Massimo Stafoggia [k,l], Maciej Strak [a,b], Wim Swart [a], Joost Wesseling [a], Danielle Vienneau [f,g], Bert Brunekreef [b,m], Gerard Hoek [b]

[a] *National Institute for Public Health and the Environment (RIVM), Bilthoven, the Netherlands*
[b] *Institute for Risk Assessment Sciences, Utrecht University, Netherlands*
[c] *University of Copenhagen, Copenhagen, Denmark*
[d] *St George's Hospital, University of London, London, UK*
[e] *Interface Demography – Department of Sociology, Vrije Universiteit Brussel, Brussels, Belgium*
[f] *Swiss Tropical and Public Health Institute, Basel, Switzerland*
[g] *University of Basel, Basel, Switzerland*
[h] *Dept. of Hygiene, Epidemiology and Medical Statistics, Medical School, National and Kapodistrian University of Athens, Athens, Greece*
[i] *NIHR HPRU Health Impact of Environmental Hazards & MRC Centre for Environment and Health Environmental Research Group, School of Public Health, Imperial College London, UK*
[j] *Department of Environmental Health, Norwegian Institute of Public Health, Oslo, Norway*
[k] *Department of Epidemiology, Lazio Region Health Service / ASL Roma 1, Rome, Italy*
[l] *Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden*
[m] *Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, the Netherlands*

ARTICLE INFO

Handling Editor: Xavier Querol

Keywords:
$PM_{2.5}$
NO_2
Exposure assessment
LUR model
Dispersion model
Mortality

ABSTRACT

Introduction: To characterize air pollution exposure at a fine spatial scale, different exposure assessment methods have been applied. Comparison of associations with health from different exposure methods are scarce. The aim of this study was to evaluate associations of air pollution based on hybrid, land-use regression (LUR) and dispersion models with natural cause and cause-specific mortality.
Methods: We followed a Dutch national cohort of approximately 10.5 million adults aged 29+ years from 2008 until 2012. We used Cox proportional hazard models with age as underlying time scale and adjusted for several potential individual and area-level socio-economic status confounders to evaluate associations of annual average residential NO_2, $PM_{2.5}$ and BC exposure estimates based on two stochastic models (Dutch LUR, European-wide hybrid) and deterministic Dutch dispersion models.
Results: Spatial variability of $PM_{2.5}$ and BC exposure was smaller for LUR compared to hybrid and dispersion models. NO_2 exposure variability was similar for the three methods. Pearson correlations between hybrid, LUR and dispersion modeled NO_2 and BC ranged from 0.72 to 0.83; correlations for $PM_{2.5}$ were slightly lower (0.61–0.72). In general, all three models showed stronger associations of air pollutants with respiratory disease and lung cancer mortality than with natural cause and cardiovascular disease mortality. The strength of the associations differed between the three exposure models. Associations of air pollutants estimated by LUR were generally weaker compared to associations of air pollutants estimated by hybrid and dispersion models. For natural cause mortality, we found a hazard ratio (HR) of 1.030 (95% confidence interval (CI): 1.019, 1.041) per 10 $\mu g/m^3$ for hybrid modeled NO_2, a HR of 1.003 (95% CI: 0.993, 1.013) per 10 $\mu g/m^3$ for LUR modeled NO_2 and a HR of 1.015 (95% CI: 1.005, 1.024) per 10 $\mu g/m^3$ for dispersion modeled NO_2.

* Corresponding author at: National Institute for Public Health and the Environment (RIVM), Antonie van Leeuwenhoeklaan 9, 3721 MA Bilthoven, the Netherlands.
E-mail address: jklompmaker@hsph.harvard.edu (J.O. Klompmaker).

https://doi.org/10.1016/j.envint.2020.106306
Received 27 August 2020; Received in revised form 4 November 2020; Accepted 26 November 2020
Available online 14 December 2020

Conclusion: Air pollution was positively associated with natural cause and cause-specific mortality, but the strength of the associations differed between the three exposure models. Our study documents that the selected exposure model may contribute to heterogeneity in effect estimates of associations between air pollution and health.

1. Introduction

A large number of epidemiological studies have shown associations of long-term exposure to ambient air pollution with mortality (Atkinson et al., 2018; Hoek et al., 2013). One of the main challenges of these studies is to assess residential long-term air pollution exposure at a fine spatial scale. Several exposure assessment methods have been developed and are now commonly applied (Jerrett et al., 2005; Hoek, 2017). Typically, studies have used land-use regression (LUR) or dispersion models to estimate long-term air pollution exposures (Jerrett et al., 2005; Hoek, 2017).

LUR and dispersion models are based on distinctly different methodological principles. LUR models are empirical models which use regression techniques to develop predictions based on air pollution measurements at a large number of sites and predictor data from geographic information systems (GIS). These empirical prediction models are then applied to non-measured locations. Dispersion models rely on deterministic equations and use data on emissions, source characteristics, chemical and physical properties of the pollutants, topography and meteorology modeling transport of pollutants through the atmosphere to estimate outdoor ground level air pollution concentrations (Jerrett et al., 2005; Hoek, 2017). Increased recognition of the limitations of both approaches has led to the development of hybrid models. These models combine data from dispersion models, land-use and surface monitoring data (Hoek, 2017) and are generally based on linear regression techniques, although Bayesian and machine learning methods (e.g. random forest) are increasingly used (Chen et al., 2019; Di et al., 2019; Cowie et al., 2019; Hanigan et al., 2017).

The use of different exposure assessment methods has been hypothesized to contribute to differences in effect estimates of associations between air pollution and health between epidemiological studies. Several studies showed moderate to good agreements between LUR and dispersion modeled air pollution concentrations (Dijkema et al., 2010; Marshall et al., 2008; Cyrys et al., 2005; Gulliver et al., 2011; Sellier et al., 2014; Hennig et al., 2016). De Hoogh et al. for example, reported a median correlation between LUR and dispersion model estimated nitrogen dioxide (NO_2) concentrations of 0.75 and between LUR and dispersion Particulate Matter <2.5 µg ($PM_{2.5}$) concentrations of 0.29 in 13 ESCAPE study areas (de Hoogh et al., 2014). However, at present, comparisons of effect estimates of different exposure assessment methods are scarce (Sellier et al., 2014; Wang et al., 2015; Jerrett et al., 2016).

This study is part of the ELAPSE project (*Effects of Low-Level Air Pollution: A Study in Europe*). Within the ELAPSE project, we evaluated associations of annual average air pollution concentrations based on a Europe-wide hybrid model including satellite observations, land-use predictors and dispersion model estimates in 11 cohorts with in-depth individual data and 7 large administrative/national cohorts (http://www.elapseproject.eu/). The aim of this study was to evaluate associations of long-term air pollution exposure based on hybrid, LUR and dispersion models with natural cause and cause-specific mortality in the Dutch national cohort (n ~ 10.5 million).

2. Methods

2.1. Study population and mortality outcomes

We created an administrative cohort that includes the full Dutch population aged 29+, on January 1, 2008, resulting in a study population of approximately 10.5 million adults. The cohort was compiled based on data from several databases from Statistics Netherlands (CBS), including mortality and individual characteristics (such as sex, marital status, region of origin and standardized household income), as described elsewhere (Fischer et al., 2015; Fischer et al., 2020). We followed the cohort from 1 January 2008 until 31 December 2012.

We linked area-level socio-economic status (SES) indicators to the cohort to adjust for potential confounding possibly not accounted for by the available individual (SES) indicators. As SES has multiple dimensions (e.g. income, occupation) and correlations between air pollutants and each of these dimension may be different, we included several SES indicators. The following indicators were linked: mean income (mean income per income recipient), percentage of non-western immigrants and unemployment rate (number of people with income support per 1000 inhabitants aged 15–64 years) in 2006 at both regional (NUTS 3, n = 40) and neighborhood level (n ~ 2600, representing on average approximately 2900 addresses). NUTS (Nomenclature des Unités Territoriales Statistique) is a geocode standard for referencing the subdivisions of countries for statistical purposes and is developed and regulated by Eurostat, the statistical office of the European Union. Further, we used a composite SES score that represents the education, occupational and economic status at regional level and at four digit postal code level (PC4, n ~ 4000, representing on average approximately 1800 addresses). This composite score was only available at PC4 level; hence we do not have the composite score at neighborhood level.

As mortality outcomes, we selected natural cause (International Classification of Diseases, 10th Revision (ICD-10) codes: A00-R99), cardiovascular disease (I10-I70), respiratory disease (J00-J99) and lung cancer mortality (C34). Secondary analyses were conducted with more specific mortality outcomes: ischemic heart disease mortality (I20-I25; IHD), cerebrovascular mortality (I60-I69, CBV) and COPD mortality (J40-J44, J47).

2.2. Exposure assessment

We used annual average residential air pollution exposure estimates based on LUR models (Eeftens et al., 2012; Beelen et al., 2013), dispersion models (Keuken et al., 2013; Velders and Diederen, 2009) and hybrid models (De Hoogh et al., 2018), referred to as LUR, dispersion and hybrid models, respectively. Annual average residential NO_2 and $PM_{2.5}$ exposures were estimated by all three models. Black carbon (BC), measured as $PM_{2.5}$ absorbance based on reflectance measurement of the filters, was estimated by the LUR and hybrid model. Elemental carbon (EC) was estimated by the dispersion model. EC is often used as a proxy for BC and on average 1 unit $PM_{2.5}$ absorbance corresponds to 1.1 µg/m^3 elemental carbon (Janssen et al., 2011). A description of the different exposure assessment models is given below. Table 1 presents key features of the three methods, which are discussed in more detail below. The spatial variation of hybrid, LUR and dispersion modeled NO_2, $PM_{2.5}$ and BC is shown in Supplemental Fig. S1.

2.2.1. LUR model

LUR models were developed within the ESCAPE (European Study of Cohorts for Air Pollution Effects) project, using air pollution measurements collected during 2009 and 2010 (Eeftens et al., 2012; Beelen et al., 2013). Measurements were conducted at regional background, urban background and traffic sites throughout the Netherlands. Three 2-week measurements were conducted at 40 ($PM_{2.5}$ and BC) and 80 (NO_2) sites [of which 6 ($PM_{2.5}$ and BC) and 12 (NO_2) were located in Belgium]. For each measurement site, results from the three 2-week measurements were averaged to estimate the annual average, adjusting for temporal variation (Eeftens et al., 2012; Beelen et al., 2013). Next, road length (Eurostreets digital road network, version 3.1 derived from the TeleAtlas MultiNet data set), traffic intensity (Nationaal WegenBestand) and European CORINE land cover predictor variables were calculated for each site, using the site coordinates and data within a GIS.

For NO_2, $PM_{2.5}$ and BC, linear LUR models were developed using a supervised stepwise selection procedure (Eeftens et al., 2012; Beelen et al., 2013). For all available potential predictor variables, univariate linear regressions with annual average air pollution concentrations were evaluated first. The predictor variable with the highest adjusted explained variance (adjusted R^2) was included in the LUR model if the direction of effect was as defined *a priori*. Next, the effect of the remaining predictor variables on the model adjusted R^2 was evaluated and the predictor variable giving the highest gain in adjusted R^2 and the right direction of effect was included in the model. This process continued until there were no more variables with the right direction of effect, which added at least 0.01 (1%) to the adjusted R^2 of the previous model. Details of the measurements and LUR modelling have been published elsewhere (Eeftens et al., 2012; Beelen, 2013).

The NO_2, $PM_{2.5}$ and BC LUR models included several predictors, mainly traffic related indicators. Model performance was moderate to high; the explained variance was 86% for NO_2, 67% for $PM_{2.5}$ and 92% for $PM_{2.5}$abs (Eeftens et al., 2012; Beelen, 2013). LUR models are shown in Supplemental Table S1. The LUR models were used to estimate air pollution concentrations at each address in the Netherlands. LUR modeled NO_2 concentrations higher than 80 μg/m^3 (n ~ 300) were set to 80 μg/m^3 as these values are probably due to an unrealistic combination of explanatory variables (the maximum annual average NO_2 concentrations measured within the ESCAPE study was 61.5 μg/m^3).

2.2.2. Dispersion model

The Dutch Operational Priority Substances (OPS) dispersion models are a combination of a Lagrangian trajectory model (for long-distance transport) and a Gaussian plume model (for the local scale road traffic contribution) (Van Jaarsveld, 2004). The Lagrangian trajectory model simulates atmospheric processes and estimates annual background concentration for NO_2 and $PM_{2.5}$ at a spatial resolution of 1 * 1 km based on emission inventory data and meteorological parameters (Velders and Diederen, 2009). EC emissions are not included in the emission inventory for the Netherlands and European sources. Hence, fractions of EC in primary $PM_{2.5}$ emissions for all relevant sources, as developed in the EUCAARI (European Integrated project on Aerosol, Cloud, Climate, and Air Quality Interactions) European research project (http://www.atm.helsinki.fi/eucaari/), were used to estimate EC emission factors.

To estimate the contribution of road-traffic emissions to the 1*1 km annual average background concentrations predicted by the Lagrangian trajectory model, two standard Dutch models were used (Keuken et al., 2013). For the contribution of road-traffic emissions inner urban roads, the street canyon model SRM1 (*Standaardrekenmethode 1*), is used. This model originates from the CAR model (Calculation of Air Pollution from Road Traffic) and specifies a source–receptor as a function of the distance to the street axis for five different road types (Eerens et al., 1993). The contribution of traffic emissions to annual average concentrations depends on the emission rate, annual average wind speed, road type and distance to the road. For the contribution of road-traffic emissions from motorways, the line-source model SRM2 is used. This model is based on a Gaussian plume model which assumes that the contribution to ambient air pollution concentrations downwind of the motorway is dependent of the emission rate and the wind speed. Several factors, such as vehicle-induced turbulence, the upwind roughness of the terrain, the presence of noise screens near the motorway are taken into account in SRM2. Both models estimate the contribution of road-traffic air pollution on address level.

We used the average 2007–2009 OPS dispersion model background estimates. For the local scale road-traffic contribution, input data for the years 2007–2009 was not complete and reliable. Therefore, input data for the year 2014 was rescaled to the year 2008. More information about the dispersion models can be found elsewhere (Keuken et al., 2013; Velders and Diederen, 2009; Wesseling and Visser, 2003; Velders et al., 2020; Wesseling et al., 2011).

2.2.3. Hybrid model

Hybrid models were developed using measured NO_2 and $PM_{2.5}$ daily concentrations (aggregated to annual mean) for 2010, derived from the AirBase v8 dataset (EEA, 2015). BC concentrations were not available through AirBase, therefore, the ESCAPE annual mean BC concentrations based on measurements in Western-Europe, in the period 2009–2010, was used (De Hoogh et al., 2018). Potential predictor variables were prepared as a series of 100 * 100 m rasters. Satellite derived NO_2 and $PM_{2.5}$ (~10 km resolution) were offered to the model (De Hoogh et al., 2018). In addition, concentrations estimates for 2010 from long-range chemical transport models (CTM) were offered as predictors (De Hoogh et al., 2018). Road length (Eurostreets digital road network, version 3.1 derived from the TeleAtlas MultiNet data set) and European CORINE land cover predictor variables were calculated for each raster. The integration of large scale satellite and CTM predictors distinguish the hybrid model from the classical LUR model.

Hybrid models were developed in a two-stage statistical procedure. First, LUR models were developed according to the ESCAPE protocol (used to develop the LUR models used in this paper), involving supervised stepwise linear regression (De Hoogh et al., 2018). Second, ordinary kriging was applied to the residuals of the model (De Hoogh et al., 2018). If kriging was not successful longitude and/or latitude were offered as additional predictors. The rationale for offering longitude and latitude was to explain large scale spatial trends in air pollution across Europe, that were not accounted for by the land use, traffic, CTM and satellite predictor variables. Annual average air pollution concentrations were estimated at a 100 * 100 m spatial resolution.

The NO_2 hybrid model includes CTM NO_2, roads, major roads, natural and residential predictor variables (De Hoogh et al., 2018). For NO_2, no model was possible with both CTM and satellite estimates, as CTM and satellite estimates were moderately strongly correlated. We previously suggested that CTM's were better developed for NO_2 than for

Table 1
Exposure assessment methods for long-term air pollution exposures.

Exposure assessment method	Monitoring approach	Modelling Approach	Area	Spatial scale	Year
LUR model	3 × 14-day average per year, 40–80 sites	Empirical (LUR)	Netherlands + Belgium	residential address	2009
Dispersion model	Dutch routine monitoring (45–60 sites)	Deterministic model	Netherlands	residential address	2008[a]
Hybrid model	Annual average 436–2399 routine monitoring sites	Empirical (hybrid LUR with CTM, SAT[b])	Western Europe	100 × 100 m grid	2010

[a] Based on average 2007–2009 Lagrangian trajectory model background estimates and local scale road-traffic contribution rescaled from 2014 to 2008.
[b] CTM = chemical transport model; SAT = satellite.

$PM_{2.5}$ when discussing the contribution of CTM and satellite estimates to $PM_{2.5}$ and NO_2 hybrid models (De Hoogh et al., 2018). The $PM_{2.5}$ hybrid model includes satellite and CTM $PM_{2.5}$ estimates, altitude, all roads, natural areas, ports and residential area. Kriging was applied to explain the left over residual variation (De Hoogh et al., 2018). The BC hybrid model includes CTM $PM_{2.5}$ estimates, roads, $PM_{2.5}$ satellite estimates, urban green, residential and natural land variables and latitude. The explained variance was 59% for NO_2, 72% for $PM_{2.5}$ and 54% for BC (De Hoogh et al., 2018).

2.3. Statistical analyses

To evaluate associations of the air pollutants with natural cause and cause-specific mortality, we used Cox proportional hazard models. We specified *a priori* Cox models stratified by sex, with age as underlying time scale and increasing degrees of covariate adjustment. All models applied a correction of the standard error for clustering in neighborhoods. Model 1 included no covariates. Model 2 included individual-level covariate data on standardized household income, region of origin and marital status. Model 3 (main model) additionally included area level data on mean income per income recipient, unemployment rate, percentage non-western immigrants and the socio-economic composite score (the educational, occupational and economical status). In addition, the difference between neighborhood and the region of mean income, unemployment rate percentage, non-western immigrants and the socio-economic composite score were included.

We explored exposure-response curves for all exposure-mortality associations (main model) using natural splines with 3 degrees of freedom. Exposure-response curves were consistent with linearity along the most commonly observed range of the concentration or indicated supra-linear shapes with steeper slopes at the lower end (Fig. S2a–d). Hence, we decided to report results for linear exposure terms. Associations were expressed per 10 $\mu g/m^3$ for NO_2, per 5 $\mu g/m^3$ for $PM_{2.5}$ and per 0.5 * 10^{-5}/m for BC (0.5 $\mu g/m^3$ for EC), based upon increments used within the ESCAPE project.

To evaluate potential mutual confounding of NO_2, $PM_{2.5}$ and BC assessed by the same method, we specified two-pollutant models. Further, we evaluated a joint hazard ratio (JHR) of air pollutants to assess the joint risk of exposure to a mixture of these pollutants. The JHR can be assessed using the Cumulative Risk Index (CRI) method (Jerrett, 2013; Crouse et al., 2015). The JHR represents the hazard for an interquartile range (IQR) increase in both air pollutants relative to the odds for no increase in any of the pollutants.

We denote the JHR based on the combination of the P pollutants evaluated at *x* as the Cumulative Risk Index (CRI) which was defined as:

$$CRI = exp\left\{\sum_{p=1}^{p}\widehat{\beta}_p x_p\right\} = \exp(\widehat{\beta}'x) = \prod_{p=1}^{p} JHR_p$$

where $\widehat{\beta}' = (\widehat{\beta}_1, \cdots, \widehat{\beta}_p)$ are the estimates of the log hazard ratio for the P exposures estimated in a Cox proportional hazard model consisting of all P exposures together, $x' = (x_1, \cdots, x_p)$ are the levels at which each exposure-specific HR is evaluated and $JHR_p = \exp(\widehat{\beta}_p x_p)$ denotes the JHR for the p^{th} exposure in a multi-exposure model. JHRs were estimated assuming additive effect estimates (log hazard ratios) of the exposures. The 95% confidence interval of CRI is defined by: $\exp\{\widehat{\beta}' x \pm 1.96 \times \sqrt{x' \times Cov(\widehat{\beta}) \times x}\}$. This definition of the confidence interval is similar to that described elsewhere (Jerrett, 2013; Crouse et al., 2015).

All statistical analyses were conducted in R (https://www.R-project.org/), version 3.4.0, following centrally developed analysis scripts.

3. Results

3.1. Study population and mortality

Our cohort consisted of 10,532,360 subjects aged 29 year or older who contributed to 50,707,159 person-years follow-up. The mean age at baseline was 52.5 years, the majority of the subjects were married and from Dutch origin (Table 2). We observed 606,527 natural cause deaths, 165,601 cardiovascular disease deaths, 63,285 respiratory disease deaths and 49,488 lung cancer deaths. Of all cardiovascular deaths, ~30% died from IHD and ~25% died from CBV. Of all respiratory deaths, ~47% died because of COPD.

3.2. Exposure distribution and mutual correlations

Hybrid modeled NO_2 and BC concentrations were higher than LUR and dispersion modeled concentrations (Fig. 1). The spatial variation in NO_2 was quite similar across all three exposure models. The exposure range (5th – 95th percentile) was larger for dispersion modeled BC than for hybrid and LUR modeled BC. Mean $PM_{2.5}$ concentrations were quite similar between the different models, but the variation differed. The IQR of $PM_{2.5}$ estimated by the LUR model (0.87 $\mu g/m^3$) was substantially lower than the IQR of $PM_{2.5}$ estimated by the hybrid model (1.90 $\mu g/m^3$) and the dispersion model (2.15 $\mu g/m^3$).

Pearson correlations between hybrid, LUR and dispersion modeled NO_2 at the residential addresses ranged from $r = 0.78$ to 0.83 (Table 3). Hybrid, LUR and dispersion modeled BC were slightly weaker correlated (Pearson $r = 0.72$–0.80). We found the lowest correlation between hybrid and LUR modeled $PM_{2.5}$ (Pearson $r = 0.61$). In general, NO_2 and BC from the same method were strongly correlated, while correlations of both exposures with $PM_{2.5}$ were lower (Table 3). Correlations between dispersion modeled NO_2, $PM_{2.5}$ and BC were substantially stronger than correlations based on hybrid or LUR models.

3.3. Associations of air pollution with mortality

3.3.1. Associations in single pollutants models

We found significant associations of air pollution with mortality outcomes in our main model (Table 4). For all three exposure models, the strongest associations were found with lung cancer mortality and the weakest associations with cardiovascular disease mortality, however the strength of the associations differed between the models. Across outcomes, effect estimates and statistical significance were more comparable between the hybrid and dispersion models. The LUR model generally showed weaker associations, except for lung cancer.

For natural cause mortality, we found associations for all three pollutants with the hybrid and dispersion model and no association with the LUR model. HRs were larger for the hybrid model for NO_2 and BC, but smaller for $PM_{2.5}$ compared to the dispersion model. For cardiovascular disease mortality, we found positive associations with the hybrid modeled pollutants and $PM_{2.5}$ from the dispersion model and no association with the LUR modeled pollutants. For respiratory disease mortality, we found the strongest associations with the hybrid and dispersion models for all three pollutants and no association with the LUR modeled pollutants. For lung cancer mortality, we found strong associations with all three models with similar HRs for all three pollutants.

In models without adjustments for individual and area-level SES indicators (Model 1), we found significant associations for all hybrid, LUR and dispersion modeled exposures and all outcomes, except for cardiovascular disease mortality (Fig. 2). The difference in strength of the associations in the minimally adjusted models was generally smaller between hybrid, LUR and dispersion modeled air pollution than in our main (fully adjusted) models, except for lung cancer mortality. For example, for natural cause mortality, the HR for LUR modeled $PM_{2.5}$ in the minimally adjusted model 1 was mildly larger compared to

Table 2
Population characteristics (n = 10,532,360).

Covariate	Category	N (%) or mean (sd)
Individual covariates		
Age		52.5 (15.1)
Sex	male	5,129,824 (48.7)
	female	5,402,536 (51.3)
Marital status	married	6,633,882 (63.0)
	widowed	836,538 (7.9)
	divorced	1,058,624 (10.1)
	single	2,003,316 (19.0)
Region of origin	Morocco	138,078 (1.3)
	Turkey	171,837 (1.6)
	Suriname	183,153 (1.7)
	Antilles Netherlands	58,971 (0.6)
	non-western	265,125 (2.5)
	western	1,009,761 (9.6)
	Dutch	8,705,435 (82.7)
Standardized household income	<1%	71,732 (0.7)
	1–5%	156,883 (1.5)
	5–10%	334,400 (3.2)
	10–25%	1,301,197 (12.4)
	25–50%	2,570,225 (24.5)
	50–75%	2,870,947 (27.3)
	75–90%	1,881,113 (17.9)
	90–95%	653,661 (6.2)
	95–99%	529,543 (5.0)
	>99%	133,457 (1.3)
Area-level covariates		
Composite SES 4 digit postal code	Based on education, income and paid occupation (year = 2007–2010)	0.02 (0.98)
Mean income neighborhood	Mean income per income recipient *€ 1000 (year = 2006)	18.23 (2.50)
Unemployment rate neighborhood	Number of people with income support per 1000 inhabitants of 15–64 years (year = 2006)	27.15 (8.78)
Percentage non-western immigrants neighborhood	Percentage non-western immigrants (year = 2006)	10.09 (11.87)
Composite SES region	Based on education, income and paid occupation (year = 2007–2010)	0.01 (0.27)
Mean income region	Mean income per income recipient *€ 1000 (year = 2006)	18.17 (1.16)
Unemployment rate region	Number of people with income support per 1000 inhabitants of 15–64 years (year = 2006)	27.35 (5.83)
Percentage non-western immigrants region	Percentage non-western immigrants (year = 2006)	10.47 (7.00)
Mortality outcomes		
natural cause mortality		606,527
cardiovascular mortality		165,601
ischemic heart disease mortality		49,248
cerebrovascular disease mortality		40,597
respiratory disease mortality		63,285
COPD mortality		29,882
Lung cancer mortality		49,488

dispersion modeled $PM_{2.5}$, whereas the HR in the fully adjusted model 3 was lower than for dispersion modeled $PM_{2.5}$. For NO_2 and BC the same pattern was found for natural cause mortality. Further, we note that for all mortality outcomes, associations in the minimally adjusted model were strongest for hybrid modeled NO_2 and BC. Upon adjustment for potential confounders, hybrid modeled pollutants showed similar decreases in HRs compared to the LUR modeled pollutants, but the HRs in model 1 were generally larger for hybrid modeled pollutants than for LUR modeled pollutants. As a result, in the main model, associations of hybrid modeled pollutants with most outcomes remained significant.

We found weak associations of hybrid and dispersion modeled air pollution with CBV mortality and no associations with IHD mortality (Table S2). Hybrid and dispersion modeled air pollution were both positively associated with COPD mortality. LUR modeled pollutants were not significantly associated with IHD, CBV and COPD mortality.

3.3.2. Associations in multi pollutants models

In two pollutant models with combinations of hybrid modeled pollutants, associations of NO_2 and BC were barely affected by adjustment for $PM_{2.5}$, except for lung cancer mortality (Table 5 for main mortality outcomes and Table S3 for secondary mortality outcomes). Associations of hybrid $PM_{2.5}$ with natural cause and respiratory disease mortality on the other hand attenuated and lost significance. In two pollutant models with combinations of dispersion modeled pollutants, associations of $PM_{2.5}$ with all outcomes remained (borderline) significant after adjustment for NO_2 or BC. Associations of dispersion modeled NO_2 or BC with all outcomes attenuated and lost significance after adjustment for $PM_{2.5}$. For example, for respiratory disease mortality, the HR of dispersion modeled NO_2 changed from 1.036 (95% CI: 1.015, 1.058) to 0.928 (95% CI: 0.894, 0.963) after adjustment for $PM_{2.5}$ and the HR of dispersion modeled $PM_{2.5}$ changed from 1.126 (95% CI: 1.087, 1.167) to 1.250 (95% CI: 1.174, 1.331) after adjustment for NO_2.

For all three models, JHR of combinations of pollutants, expressed per IQR increase (not per fixed increment), were similar or slightly larger compared to HRs of single pollutant models expressed per IQR (Fig. S3). For example, for lung cancer mortality, the HR of hybrid modeled NO_2 was 1.084 (95% CI: 1.063, 1.105) per IQR increase and the JHR of hybrid modeled NO_2 and $PM_{2.5}$ was 1.090 (95% CI: 1.069, 1.112) per IQR increase and the HR of dispersion modeled $PM_{2.5}$ was 1.066 (95% CI: 1.047, 1.085) per IQR increase and the JHR of dispersion modeled $PM_{2.5}$ and NO_2 was 1.068 (95% CI: 1.049, 1.088). JHRs based on hybrid modeled pollutants were fairly similar in strength. JHR based on dispersion modeled pollutants were generally largest for a combination of $PM_{2.5}$ and BC.

4. Discussion

Exposure estimates from the hybrid, LUR and dispersion models were moderately to strongly correlated. We found generally positive associations of air pollution with natural cause, cardiovascular disease, respiratory disease and lung cancer mortality, but the strength of the associations differed between the three exposure models. For all three models, the strongest associations were found with lung cancer mortality and the weakest associations with cardiovascular disease

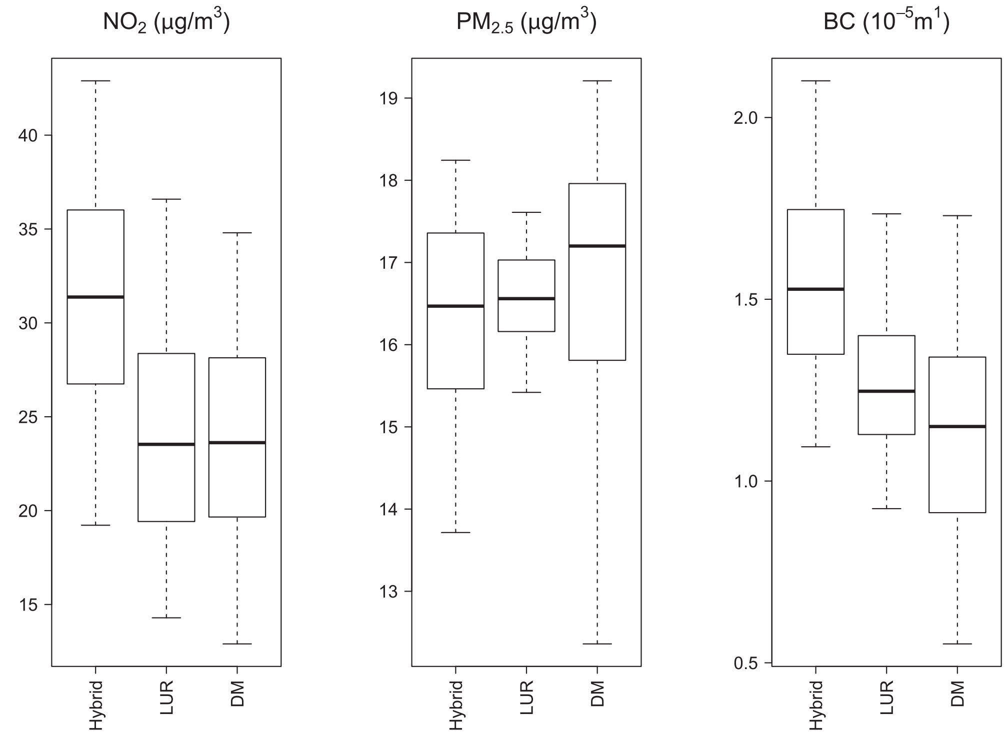

Fig. 1. Boxplots of NO_2, $PM_{2.5}$ and BC concentrations based on a hybrid, LUR and dispersion (DM) model.[a,b] ([a]The middle box represents the middle 50% of the concentration, the line that divides the box in two is the median. The upper and lower end of the whiskers represent the 5th and 95th percentile. [b] DM BC is modeled per μg/m^3).

Table 3
Pearson correlations between NO_2, $PM_{2.5}$ and BC based on a hybrid, LUR and dispersion (DM) model and mean concentrations (standard deviation = sd) of each pollutant.[a]

Pearson correlation		Hybrid			LUR			DM		
		NO_2	$PM_{2.5}$	BC	NO_2	$PM_{2.5}$	BC	NO_2	$PM_{2.5}$	BC
Hybrid	NO_2	1.00								
	$PM_{2.5}$	0.58	1.00							
	BC	0.88	0.52	1.00						
LUR	NO_2	0.83	0.43	0.77	1.00					
	$PM_{2.5}$	0.52	0.61	0.56	0.45	1.00				
	BC	0.78	0.44	0.80	0.86	0.74	1.00			
DM	NO_2	0.79	0.56	0.74	0.78	0.63	0.79	1.00		
	$PM_{2.5}$	0.71	0.69	0.66	0.66	0.72	0.70	0.86	1.00	
	BC	0.75	0.51	0.72	0.72	0.60	0.75	0.94	0.85	1.00
Mean (sd)		31.3 μg/m^3 (7.1)	16.3 μg/m^3 (1.4)	1.6 * 10^{-5}/m (0.3)	24.3 μg/m^3 (6.9)	16.6 μg/m^3 (0.7)	1.3 * 10^{-5}/m (0.2)	23.9 μg/m^3 (6.8)	16.7 μg/m^3 (2.0)	1.1 μg/m^3 (0.4)

[a] Mutual correlations are given in gray cells.

mortality. Air pollution modeled with the hybrid and dispersion models were generally significantly associated with all mortality outcomes, whereas air pollutants modeled with LUR were only significantly associated with lung cancer mortality. Differences between the three models were smaller for the minimally adjusted confounder model than the main model, suggesting that sensitivity to confounding differed between the three exposure models. Two pollutant models suggested more robust associations with NO_2 for the hybrid model and with $PM_{2.5}$ for the

Table 4
Associations of air pollution based on a hybrid, LUR and dispersion (DM) model on natural cause, cardiovascular, respiratory and lung cancer mortality in single-pollutant models.[a,b]

Outcome	Pollutant	Hybrid	LUR	DM
		HR (95% CI)	HR (95% CI)	HR (95% CI)
Natural cause mortality	NO_2	1.030 (1.019, 1.041)	1.003 (0.993, 1.013)	1.015 (1.005, 1.024)
	$PM_{2.5}$	1.021 (0.999, 1.044)	1.006 (0.973, 1.040)	1.035 (1.018, 1.052)
	BC	1.030 (1.019, 1.041)	1.005 (0.993, 1.017)	1.018 (1.009, 1.027)
Cardiovascular disease mortality	NO_2	1.017 (1.003, 1.031)	0.984 (0.970, 0.999)	1.000 (0.988, 1.013)
	$PM_{2.5}$	1.015 (0.988, 1.042)	0.994 (0.952, 1.039)	1.021 (0.998, 1.044)
	BC	1.018 (1.003, 1.033)	0.991 (0.975, 1.007)	1.006 (0.994, 1.018)
Respiratory disease mortality	NO_2	1.038 (1.015, 1.062)	0.996 (0.972, 1.019)	1.036 (1.015, 1.058)
	$PM_{2.5}$	1.058 (1.009, 1.111)	1.059 (0.989, 1.133)	1.126 (1.087, 1.167)
	BC	1.052 (1.028, 1.076)	1.005 (0.981, 1.031)	1.043 (1.024, 1.062)
Lung cancer mortality	NO_2	1.091 (1.069, 1.114)	1.089 (1.065, 1.113)	1.079 (1.056, 1.102)
	$PM_{2.5}$	1.169 (1.121, 1.218)	1.243 (1.155, 1.337)	1.155 (1.112, 1.198)
	BC	1.085 (1.062, 1.108)	1.087 (1.061, 1.114)	1.071 (1.051, 1.091)

[a] Associations are expressed per 10 μg/m^3 for NO_2, per 5 μg/m^3 for $PM_{2.5}$ and per 0.5 * 10^{-5}/m for BC (0.5 μg/m^3 for EC).
[b] Associations of main model are adjusted for age, strata(sex), random (neighborhood), standardized household income, region of origin, marital status, socio-economic composite score region, mean income per income recipient region, unemployment rate region, percentage non-western immigrants region, and the difference between neighborhood and region of mean income, unemployment rate, non-western immigrants and the composite SES score (4 digit postal code).

dispersion model.

4.1. Hybrid, LUR and dispersion modeled air pollution exposure patterns

Correlations between hybrid, LUR and dispersion modeled air pollution were strongest for NO_2. This is likely due to the strong contribution of traffic to the small-scale variation of NO_2 within the Netherlands. The LUR and hybrid model differ from the dispersion model in that they do not rely on dispersion processes and assumptions about (traffic) emissions, but only use land use data (LUR model), such as traffic intensity, or a combination of land use data, satellite observations and dispersion model estimates (hybrid model). However, land use data, such as traffic intensity and population density, is used in the calculations of emissions and dispersions of air pollution in the dispersion models. Correlations between hybrid, LUR and dispersion modeled BC were slightly lower than for NO_2, despite the fact that BC is also strongly determined by traffic. This might be due to the smaller number of BC measurements and a lack of emissions data. BC measurements are not available through AirBase and BC emissions are not included in the emission inventory for the Netherlands. Instead, $PM_{2.5}$ satellite and CTM are included in the hybrid BC model, as indicators of BC emissions, which may have resulted in lower model performance. Because of the lack of BC emission data, assumed fractions of EC in primary $PM_{2.5}$ emissions for all relevant sources are used to estimate EC by the dispersion model. Correlations between hybrid, LUR and dispersion modeled $PM_{2.5}$ were lowest. This could be due to the low spatial variation of $PM_{2.5}$ within the Netherlands. The influence of traffic sources on $PM_{2.5}$ concentrations is lower and the influence of other sources such as industry is higher compared to NO_2 and BC concentrations (Eeftens et al., 2012).

De Hoogh et al. previously reported a strong correlation for LUR and dispersion modeled NO_2 (median Pearson correlation of 0.75) and a weak correlation for LUR and dispersion modeled $PM_{2.5}$ (median Pearson correlation of 0.29) within 13 ESCAPE study areas (de Hoogh et al., 2014). Correlations between dispersion modeled air pollution and *measured* air pollution at the Dutch ESCAPE measurement sites was 0.85 for NO_2 and 0.54 for $PM_{2.5}$ (de Hoogh et al., 2014). However, the correlations between the dispersion and the LUR model of approximately 1000 Dutch children of the Prevention and Incidence of Asthma and Mite Allergy birth cohort were strong for NO_2 and $PM_{2.5}$ (Pearson R = 0.89 for NO_2 and 0.81 for $PM_{2.5}$) (Wang et al., 2015). Sellier et al. and Wang et al. also reported strong correlations between LUR and dispersion modeled NO_2 (Pearson R > 0.85) (Sellier et al., 2014; Wang et al., 2015). Wang et al. also reported a strong correlation between LUR and dispersion modeled $PM_{2.5}$ (Pearson R = 0.86) (Wang et al., 2015).

Hybrid modeled NO_2 concentrations were higher compared to LUR and dispersion modeled NO_2 concentrations. De Hoogh et al. (De Hoogh et al., 2018) reported an overestimation (fractional bias) of 13% for hybrid modeled NO_2 compared to ESCAPE NO_2 measurements in the overall ELAPSE area and a 26% overestimation for the Dutch ESCAPE NO_2 measurements. This overestimation is in line with the differences between hybrid and LUR modeled concentrations reported in this study. In the ESCAPE project, NO_2 was measured with Ogawa badges, which resulted in lower concentrations compared to chemiluminescence measurements on which the hybrid model was based (Cyrys et al., 2012). In the Netherlands, the Ogawa measurements were about 20% lower than the concurrent chemiluminescence measurements (Cyrys et al., 2012). We do not have an explanation why the hybrid model predicts higher concentrations than the dispersion model, which fits the Dutch monitoring data well. This may be a limitation of applying a European model in a single country. The models also differed in locations of monitoring sites, which may have influenced the concentrations levels. Monitoring sites used for LUR models were selected to represent residential exposure, i.e. sites were located near building facades representative for residential addresses. Monitoring stations used for dispersion and hybrid modelling on the other hand, were used for regulatory purposes and not all located near residential addresses. The higher hybrid modeled NO_2 concentrations (year = 2010) compared to LUR (year = 2009) and dispersion modeled NO_2 (year = 2008) are likely not due to the different year. NO_2 concentrations measured in the Dutch monitoring network showed that the average NO_2 concentrations in the Netherlands were somewhat lower in 2010 compared to 2009 and 2008 (RIVM, DCMR, and GGD Amsterdam, 2018). However, we note that LUR model estimates are based on 3 two-week measurements and recalibrated to annual averages, while hybrid (NO_2 and $PM_{2.5}$) and dispersion model estimates are based on/validated with continuous measurements from monitoring stations. $PM_{2.5}$ concentrations were fairly similar between the three methods. However, LUR modeled $PM_{2.5}$ has a substantially lower spatial variation than hybrid and dispersion modeled $PM_{2.5}$. This is likely due to the limited contrast in $PM_{2.5}$ concentrations measured at urban background and regional background in the Netherlands within the ESCAPE project (Eeftens et al., 2012). Further, large scale satellite, CTM predictors (hybrid) and the Lagrangian trajectory model (dispersion) may better capture long-range transport than

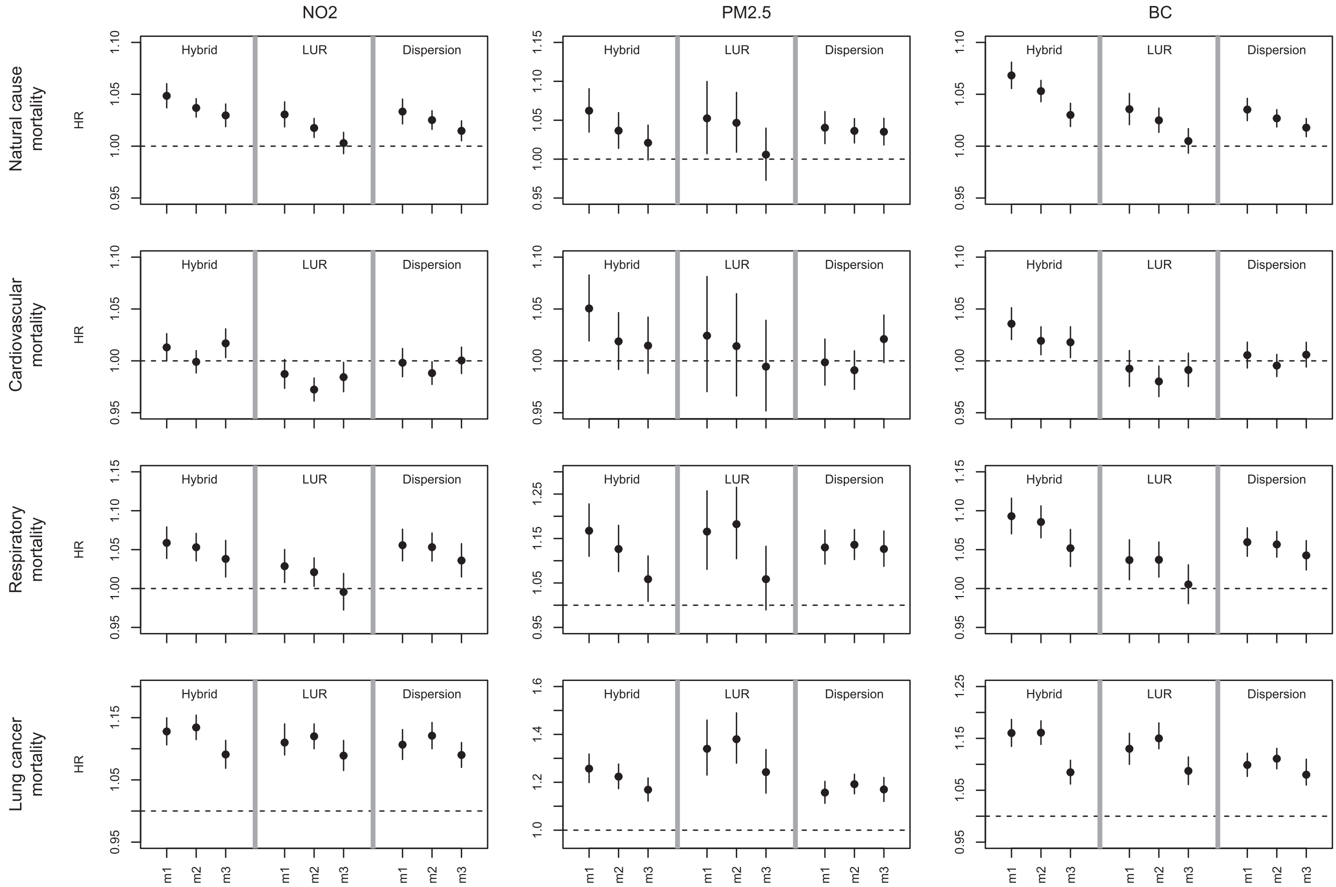

Fig. 2. Associations of air pollution based on a hybrid, LUR and dispersion model with natural cause, cardiovascular, respiratory and lung cancer mortality in models with increasing degree of adjustment for potential confounders.[a,b]. ([a]Associations are expressed per 10 μg/m^3 for NO_2, per 5 μg/m^3 for $PM_{2.5}$ and per 0.5 * 10^{-5}/m for BC (0.5 μg/m^3 for EC). Model 1 (m1) included the baseline hazard, a strata term for sex and a cluster for neighborhood. Model 2 (m2) additionally included standardized household income, region of origin and marital status. Model 3 (m3, main model) additionally included mean income per income recipient of the region, unemployment rate of the region, percentage non-western immigrants of the region and the socio-economic composite score (the educational, occupational and economical status) of the region, the difference between mean income per income recipient of the neighborhood and that of the region, the difference between unemployment rate of the neighborhood and that of the region, the difference between percentage non-western immigrants of the neighborhood and that of the region, the difference between the socio-economic composite score (based on the educational, occupational and economical status) at a four digit postal code level and that of the region).

by monitoring alone.

4.2. Differences in effect estimates between hybrid, LUR and dispersion modeled air pollution

Based on the moderate to strong correlations between exposure estimates of the three models, we had expected more consistency in effect estimates of an association between air pollution and mortality. We do not know which of the three exposure methods performs best. Differences in effect estimates in the fully adjusted models could be explained by differences in exposure measurement error, in predicted exposure contrasts and in sensitivity to adjustment for confounders.

Exposure measurement error effects can lead to an attenuation of effect estimates and an increase of the confidence intervals (Basagaña et al., 2012). The impact of exposure measurement error on the estimated health effects is complex, as it depends on the combination of errors (Berkson, classical) (Samoli et al., 2020; Butland et al., 2020). The combination of errors may differ between pollutants and between methods. Performance of exposure assessment models are generally evaluated by R^2 measures. However, published validation statistics of the hybrid and LUR models cannot be directly compared. Performance of hybrid models was based on (hold out validation of) measurement sites across Western-Europe, while validation for LUR models was based on leave-one-out cross validation of Dutch ESCAPE measurements sites. Performance of the hybrid models for the Netherlands only, showed that the explained variation of the ESCAPE measurements of the hybrid NO_2 model was high in the Netherlands (R2 = 76%, 80 sites) , in contrast to the explained variation of the hybrid $PM_{2.5}$ model (R^2 = 13%, 40 sites) (De Hoogh et al., 2018). Further, we note that 40 measurement sites were used to develop the Dutch $PM_{2.5}$ and BC LUR model (Beelen et al., 2013), while 543 measurement sites were used to develop the hybrid $PM_{2.5}$ model and 436 sites were used to develop the hybrid BC model (De Hoogh et al., 2018). Models based on a small number of measurement sites tend to give higher R^2 and leave-one out cross-validated R^2 than those based on more sites (Basagaña et al., 2012). A larger number of sites will produce more stable models (Basagaña et al., 2012; Wang et al., 2013), which was the rationale to develop the Europe-wide hybrid models in the ELAPSE project.

National exposure assessment models, such as the LUR and dispersion models, may better capture national-specific small-scale variation patterns than the Western-Europe wide hybrid model, as relations between air pollution and predictor variables may differ between countries (De Hoogh et al., 2018). The hybrid model tends to average intra-study area differences in air pollution - predictor variables relations over entire study area (De Hoogh et al., 2018). However, as the LUR models used in this study are based on a relatively small number of measurement sites, they may not capture small-scale air pollution variations better than dispersion or hybrid models. Further, differences in air

Table 5
Associations of air pollution based on a hybrid, LUR and dispersion model with natural cause, cardiovascular, respiratory and lung cancer mortality in multi-pollutant models.[a,b]

Outcome	Pollutant	Adj. for $PM_{2.5}$			Adj. for BC			Adj. for NO_2		
		HR (95% CI)			HR (95% CI)			HR (95% CI)		
		Hybrid	LUR	DM	Hybrid	LUR	DM	Hybrid	LUR	DM
Natural cause mortality	NO_2	1.039 (1.025, 1.054)	1.003 (0.991, 1.015)	0.992 (0.977, 1.008)	1.017 (0.997, 1.038)	0.999 (0.982, 1.016)	0.975 (0.950, 1.001)	.	.	.
	$PM_{2.5}$	.	.	.	0.988 (0.963, 1.015)	0.982 (0.927, 1.042)	1.023 (0.996, 1.05)	0.972 (0.944, 1.000)	1.001 (0.963, 1.041)	1.046 (1.017, 1.076)
	BC	1.033 (1.020, 1.047)	1.010 (0.989, 1.031)	1.008 (0.994, 1.022)	.	.	.	1.015 (0.994, 1.036)	1.006 (0.986, 1.026)	1.039 (1.014, 1.065)
Cardio- vascular disease mortality	NO_2	1.021 (1.002, 1.040)	0.980 (0.964, 0.997)	0.969 (0.948, 0.990)	1.007 (0.984, 1.032)	0.976 (0.954, 0.998)	0.960 (0.926, 0.995)	.	.	.
	$PM_{2.5}$	.	.	.	0.996 (0.965, 1.028)	1.047 (0.967, 1.133)	1.036 (1.000, 1.073)	0.988 (0.953, 1.025)	1.027 (0.977, 1.08)	1.067 (1.028, 1.108)
	BC	1.019 (1.001, 1.037)	0.978 (0.950, 1.007)	0.991 (0.972, 1.009)	.	.	.	1.011 (0.985, 1.038)	1.013 (0.988, 1.038)	1.041 (1.006, 1.076)
Respiratory disease mortality	NO_2	1.033 (1.001, 1.065)	0.982 (0.956, 1.010)	0.928 (0.894, 0.963)	0.983 (0.941, 1.027)	0.980 (0.943, 1.018)	0.946 (0.891, 1.004)	.	.	.
	$PM_{2.5}$	.	.	.	1.010 (0.953, 1.070)	1.145 (1.013, 1.294)	1.176 (1.108, 1.249)	1.016 (0.952, 1.085)	1.090 (1.007, 1.178)	1.250 (1.174, 1.332)
	BC	1.049 (1.021, 1.078)	0.967 (0.925, 1.011)	0.972 (0.943, 1.003)	.	.	.	1.067 (1.021, 1.116)	1.023 (0.983, 1.065)	1.092 (1.035, 1.152)
Lung cancer mortality	NO_2	1.063 (1.035, 1.092)	1.072 (1.046, 1.099)	1.018 (0.985, 1.052)	1.075 (1.035, 1.116)	1.074 (1.039, 1.110)	1.040 (0.980, 1.100)	.	.	.
	$PM_{2.5}$	.	.	.	1.110 (1.058, 1.163)	1.061 (0.942, 1.195)	1.120 (1.060, 1.190)	1.082 (1.026, 1.141)	1.106 (1.021, 1.200)	1.126 (1.062, 1.194)
	BC	1.054 (1.028, 1.081)	1.069 (1.027, 1.112)	1.020 (0.990, 1.050)	.	.	.	1.018 (0.979, 1.059)	1.020 (0.984, 1.059)	1.040 (0.990, 1.090)

[a] Associations are expressed per 10 $\mu g/m^3$ for NO_2, per 5 $\mu g/m^3$ for $PM_{2.5}$ and per 0.5 * 10^{-5}/m for BC (0.5 $\mu g/m^3$ for EC).
[b] Associations of main model are adjusted for age, strata(sex), random(neighborhood), standardized household income, region of origin, marital status, socio-economic composite score region, mean income per income recipient region, unemployment rate region, percentage non-western immigrants region, and the difference between neighborhood and region of mean income, unemployment rate, non-western immigrants and the composite SES score (4 digit postal code).

pollution - predictor variables relations in the Netherlands and in Western-Europe might be limited. Another notable difference is that the hybrid models estimated air pollution concentrations on a 100 * 100 m grid while the LUR and dispersion modeled pollutants are estimated on address level. Address level estimates might be more accurate, but they may also be more sensitive to geocoding errors. We accounted for major geocoding errors in the LUR model (e.g. unreasonably small distance between address geocode and nearest road) by truncating the predictor variables to the minimum or maximum observed at the monitoring sites. Because the LUR, dispersion and hybrid models were independently developed over a different area and using different input data, our ability to attribute the differences in effect estimates to specific components of the exposure modelling method was somewhat hindered.

Predicted concentration ranges differed between hybrid, LUR and dispersion modeled pollutants. The IQR for LUR modeled BC and especially $PM_{2.5}$ were lower than for hybrid and dispersion modeled BC and $PM_{2.5}$. HRs are expressed per fixed increment to be able to compare effects of hybrid, LUR and dispersion modeled pollutants. However, HRs of $PM_{2.5}$ are expressed per 5 $\mu g/m^3$ which is more than five times as large as the IQR of LUR modeled $PM_{2.5}$ and more than twice as large as the difference between the 95th − 5th percentile. A lower spatial variation limits the ability to capture differences in event rates across exposure ranges. However, if differences in event rates are captured by LUR modeled $PM_{2.5}$, HR expressed per 5 $\mu g/m^3$ can be very large, such as the association of LUR modeled $PM_{2.5}$ with lung cancer mortality. Effect estimates expressed per IQR were weaker for LUR modeled $PM_{2.5}$ than for hybrid and dispersion modeled $PM_{2.5}$ (Fig. S3).

We observed fairly similar effect estimates in the minimally adjusted models. In the fully adjusted models, where we adjusted for individual, neighborhood- and regional-level SES variables, the differences in effect estimates were larger. By adjusting for area-level SES variables, some of the neighborhood and regional scale variation in mortality was removed from the air pollution estimates. As the exposure assessment models differ in structure (i.e. importance of large- and small-scale predictors), they may be differentially sensitive to relations with area-level variables. The higher sensitivity of the hybrid and LUR models to ***adjustment for confounders*** may be due to the use of more generic predictor variables such as population density compared to specific air pollution emissions in the dispersion model. Further, we note that we did not adjust for traffic noise and personal lifestyle factors, such as smoking status and BMI, that may have impacted effect estimates of the models differently.

Correlations between air pollutants were much stronger in the dispersion model compared to the hybrid and LUR model. This may be due to the inclusion of different predictors and zones of influence (e.g. size of circular buffers) in the hybrid and LUR models versus the use of the same fundamental dispersion processes and assumptions about emissions for all pollutants in the dispersion models (Fecht et al., 2016). The much stronger correlation between $PM_{2.5}$ and NO_2/BC in the dispersion model could be due to an overestimation of the contribution of traffic to the total $PM_{2.5}$ concentration in the dispersion model or an underestimation of the contribution of traffic to the total $PM_{2.5}$ concentration in the hybrid and LUR model. We found that dispersion modeled $PM_{2.5}$ was slightly stronger correlated with hybrid modeled NO_2 (Pearson r = 0.71) than with hybrid modeled $PM_{2.5}$ (Pearson r = 0.69). Compared to hybrid and LUR modeled pollutants, the high correlation between dispersion modeled pollutants resulted in unstable effect estimates. In two-pollutant models with combinations of dispersion modeled pollutants, $PM_{2.5}$ remained associated with natural cause, cardiovascular and respiratory disease mortality outcomes after adjustment for NO_2 or BC. The opposite pattern was seen for hybrid modeled pollutants. Despite the weaker correlations of hybrid and LUR pollutants, JHRs showed a similar pattern for all three methods. The JHRs of the air pollutant mixture were similar or only slightly higher compared to HRs from single pollutant models for each exposure assessment methods. This suggest that a single pollutant could be sufficient to characterize the toxicity of the air pollution mixture (of NO_2, $PM_{2.5}$ and BC).

4.3. Comparison with previous studies

There are only a few studies that compared effect estimates of LUR and dispersion models. Sellier et al. (2014) previously compared associations of LUR and dispersion modeled NO_2 with birth weight. Most associations were weak and not significant and differed only slightly between the exposure models (Sellier et al., 2014). Wang et al. (2015) reported similar significant associations of LUR and dispersion modeled NO_2 and BC with lung function (FEV_1, FVC) in Dutch children, while effect estimates (expressed per 5 μg/m^3) with LUR modeled $PM_{2.5}$ were two times higher than for dispersion modeled $PM_{2.5}$. This could be due to the difference in exposure contrast; the IQR for dispersion modeled $PM_{2.5}$ (3.7 μg/m^3) was three times larger than the IQR for LUR modeled $PM_{2.5}$ (1.1 μg/m^3) (Wang et al., 2015).

Associations of hybrid and dispersion NO_2 with natural cause, cardiovascular and respiratory disease mortality in our study were in line with recent HRs from meta-analysis (Atkinson et al., 2018). Associations of hybrid and dispersion $PM_{2.5}$ were slightly weaker than HRs from meta-analysis for cardiovascular disease mortality and stronger than HRs from meta-analysis for respiratory disease mortality (Hoek et al., 2013). We found the strongest associations with lung cancer mortality, but note that we lack information about smoking. However, a study by Fischer et al. (Fischer et al., 2015) suggested that it is unlikely that uncontrolled confounding from smoking or BMI substantially biased associations of air pollution with mortality in the Dutch national cohort. Klompmaker et al. (2020) found no associations of air pollutants, modeled by the same LUR model as used in this article, with natural cause and cause-specific mortality in a 5 year follow-up of a large national health survey in models with adjustment for individual SES indicators and lifestyle factors and in models with adjustment for individual, neighborhood and regional SES. The authors speculated that the null findings might be due to the short follow-up period (Klompmaker et al., 2020).

Two previously published studies reported positive associations of exposure to air pollution with mortality in a Dutch administrative cohort (including > 7 million individuals aged 30 years or older) (Fischer et al., 2015, 2020). Fischer et al. (2015) used different LUR models than we used. Associations of their LUR modeled NO_2 were different from associations of our LUR modeled NO_2, but similar to associations of hybrid and dispersion modeled NO_2. Fischer et al. (2020) used the same dispersion model as we used and reported slightly stronger associations for $PM_{2.5}$ and BC with natural cause mortality. We note some important differences, as both studies performed by Fischer and colleagues excluded all individuals who moved 5 years before the start of the follow-up period (Fischer et al., 2015; Fischer et al., 2020). This may have contributed to slightly stronger associations compared to the associations in this study. In addition, the follow-up period of Fischer et al. (2015) and of Fischer et al. (2020) was from 2004 to 2011 and from 2008 to 2015, respectively. Furthermore, both studies only adjusted for neighborhood SES composite score as area-level SES indicator, while we included several neighborhood and regional level SES indicators. Associations of NO_2 with natural cause mortality with limited adjustments for neighborhood and regional SES were similar to associations of our main model. For example, for natural cause mortality, for models with adjustments for only regional and neighborhood composite SES, we found a HR of 1.022 (95% CI: 1.014, 1.030) for hybrid modeled NO_2, a HR of 1.005 (95% CI: 0.994, 1.016) for LUR modeled NO_2 and a HR of 1.014 (95% CI: 1.006, 1.022) for dispersion modeled NO_2. For models with adjustment for neighborhood and regional mean income, percentage of non-western immigrants and unemployment rate, but not for composite SES, we found a HR of 1.032 (95% CI: 1.021, 1.043) for hybrid modeled NO_2, a HR of 1.003 (95% CI: 0.999, 1.007) for LUR modeled NO_2 and a HR of 1.016 (95% CI: 1.007, 1.026) for dispersion modeled NO_2. We acknowledge that some over-adjustment is possible, but preferred this to inadequate adjustment for SES.

5. Conclusion

Air pollution exposure estimates from a hybrid model, LUR and dispersion model were moderately to strongly correlated. We found generally positive associations of air pollution with natural cause, cardiovascular disease, cerebrovascular disease, respiratory disease, COPD and lung cancer mortality, but not with ischemic heart disease mortality. Despite the strong mutual correlations, the strength of the associations differed between the three exposure models. For all three models, the strongest associations were found with lung cancer mortality and the weakest associations with cardiovascular disease mortality. Air pollution modeled by the hybrid and dispersion models were generally more strongly associated with mortality than air pollution modeled by the LUR. Two pollutant models suggested more robust associations with NO_2 for the hybrid model and with $PM_{2.5}$ for the dispersion model. The difference between effect estimates depended on the mortality outcome. Differences in effect estimates between models are likely due to different measurement error, different sensitivity to confounding and different predicted exposure contrasts. Overall, our study documents that the selected exposure model may contribute to heterogeneity in effect estimates from cohort studies of long-term exposure to outdoor air pollution and mortality.

Funding

This study was funded by the Health Effects Institute, United States (4954-RFA14-3/16-5) and the RIVM, the Netherlands (RIVM Strategic Program (SPR); S/121004 HERACLES). Gerard Hoek was also supported by the CLAIRE project funded by the Health Effects Institute, United States, RESEARCH AGREEMENT # 4973-RFA19-1/20-7.

CRediT authorship contribution statement

Jochem O. Klompmaker: Formal analysis, Methodology, Software, Visualization, Writing - original draft. **Nicole Janssen:** Data curation, Software, Supervision, Writing - review & editing. **Zorana J. Andersen:** Writing - review & editing. **Richard Atkinson:** Writing - review & editing. **Mariska Bauwelinck:** Writing - review & editing. **Jie Chen:** Project administration, Writing - review & editing. **Kees de Hoogh:** Methodology, Software, Writing - review & editing. **Danny Houthuijs:** Writing - review & editing. **Klea Katsouyanni:** Methodology, Writing -

review & editing. **Marten Marra:** Data curation, Software, Writing - review & editing. **Bente Oftedal:** Writing - review & editing. **Sophia Rodopoulou:** Methodology, Software, Writing - review & editing. **Evangelia Samoli:** Methodology, Software, Writing - review & editing. **Massimo Stafoggia:** Methodology, Software, Writing - review & editing. **Maciej Strak:** Project administration, Writing - review & editing. **Wim Swart:** Software, Writing - review & editing. **Joost Wesseling:** Methodology, Data curation, Writing - review & editing. **Danielle Vienneau:** Writing - review & editing. **Bert Brunekreef:** Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Software, Supervision, Validation, Writing - review & editing. **Gerard Hoek:** Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Software, Supervision, Validation, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envint.2020.106306.

References

Atkinson, R.W., et al., 2018. Long-term concentrations of nitrogen dioxide and mortality: a meta-analysis of cohort studies. Epidemiol. (Cambridge, Mass.) 29 (4), 460.
Basagaña, X., et al., 2012. Effect of the number of measurement sites on land use regression models in estimating local air pollution. Atmos. Environ. 54, 634–642.
Beelen, R., et al., 2013. Development of NO 2 and NO x land use regression models for estimating air pollution exposure in 36 study areas in Europe–the ESCAPE project. Atmos. Environ. 72, 10–23.
Butland, B.K., et al., 2020. Comparing the performance of air pollution models for nitrogen dioxide and ozone in the context of a multilevel epidemiological analysis. Environ. Epidemiol. 4 (3), e093.
Chen, J., et al., 2019. A comparison of linear regression, regularization, and machine learning algorithms to develop Europe-wide spatial models of fine particles and nitrogen dioxide. Environ. Int. 130, 104934.
Cowie, C.T., et al., 2019. Comparison of model estimates from an intra-city land use regression model with a national satellite-LUR and a regional Bayesian Maximum Entropy model, in estimating NO2 for a birth cohort in Sydney, Australia. Environ. Res. 174, 24–34.
Crouse, D.L., et al., 2015. Ambient PM2. 5, O3, and NO2 exposures and associations with mortality over 16 years of follow-up in the Canadian Census Health and Environment Cohort (CanCHEC). Environ. Health Perspect. 123 (11), 1180.
Cyrys, J., et al., 2005. GIS-based estimation of exposure to particulate matter and NO2 in an urban area: stochastic versus dispersion modeling. Environ. Health Perspect. 113 (8), 987–992.
Cyrys, J., et al., 2012. Variation of NO2 and NOx concentrations between and within 36 European study areas: results from the ESCAPE study. Atmos. Environ. 62, 374–390.
de Hoogh, K., et al., 2014. Comparing land use regression and dispersion modelling to assess residential exposure to ambient air pollution for epidemiological studies. Environ. Int. 73, 382–392.
De Hoogh, K., et al., 2018. Spatial PM2. 5, NO2, O3 and BC models for Western Europe-Evaluation of spatiotemporal stability. Environ. Int. 120, 81–92.
Di, Q., et al., 2019. Assessing NO2 Concentration and Model Uncertainty with High Spatiotemporal Resolution across the Contiguous United States Using Ensemble Model Averaging. Environ. Sci. Technol. 54 (3), 1372–1384.
Dijkema, M.B., et al., 2010. A comparison of different approaches to estimate small-scale spatial variation in outdoor NO2 concentrations. Environ. Health Perspect. 119 (5), 670–675.
EEA, 2015. Airbase - The European Air Quality Database, Version 8. [cited 2019 16-10]; Available from: http://www.eea.europa.eu/data-and-maps/data/airbase-the-european-air-quality-database-8.
Eeftens, M., et al., 2012. 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. Atmos. Environ. 62, 303–317.
Eeftens, M., et al., 2012. Development of land use regression models for PM2. 5, PM2. 5 absorbance, PM10 and PMcoarse in 20 European study areas; results of the ESCAPE project. Environ. Sci. Technol. 46(20), 11195–11205.
Eerens, H., Sliggers, C., Van den Hout, K., 1993. The CAR model: the Dutch method to determine city street air quality. Atmos. Environ. Part B 27 (4), 389–399.
Fecht, D., et al., 2016. Spatial and temporal associations of road traffic noise and air pollution in London: Implications for epidemiological studies. Environ. Int. 88, 235–242.
Fischer, P.H., et al., 2015. Air pollution and mortality in seven million adults: the Dutch environmental longitudinal study (DUELS). Environ. Health Perspect. 123 (7), 697–704.
Fischer, P.H., et al., 2020. Particulate air pollution from different sources and mortality in 7.5 million adults—The Dutch Environmental Longitudinal Study (DUELS). Sci. Total Environ. 705, 135778.
Gulliver, J., et al., 2011. Comparative assessment of GIS-based methods and metrics for estimating long-term exposures to air pollution. Atmos. Environ. 45 (39), 7072–7080.
Hanigan, I.C., et al., 2017. Blending multiple nitrogen dioxide data sources for neighborhood estimates of long-term exposure for health research. Environ. Sci. Technol. 51 (21), 12473–12480.
Hennig, F., et al., 2016. Comparison of land-use regression modeling with dispersion and chemistry transport modeling to assign air pollution concentrations within the Ruhr area. 7(3), p. 48.
Hoek, G., et al., 2013. Long-term air pollution exposure and cardio-respiratory mortality: a review. Environ. Health 12 (1), 43.
Hoek, G., 2017. Methods for assessing long-term exposures to outdoor air pollutants. Curr. Environ. Health Reports 4 (4), 450–462.
Janssen, N.A., et al., 2011. Black carbon as an additional indicator of the adverse health effects of airborne particles compared with PM10 and PM2. 5. Environ. Health Perspect. 119 (12), 1691–1699.
Jerrett, M., et al., 2005. A review and evaluation of intraurban air pollution exposure models. J. Eposure Sci. Environ. Epidemiol. 15 (2), 185.
Jerrett, M., et al., 2013. Spatial analysis of air pollution and mortality in California. Am. J. Respir. Crit. Care Med. 188 (5), 593–599.
Jerrett, M., et al., 2016. Comparing the health effects of ambient particulate matter estimated using ground-based versus remote sensing exposure estimates. Environ. Health Perspect. 125 (4), 552–559.
Keuken, M., et al., 2013. Modelling elemental carbon at regional, urban and traffic locations in The Netherlands. Atmos. Environ. 73, 73–80.
Klompmaker, J.O., et al., 2020. Surrounding green, air pollution, traffic noise exposure and non-accidental and cause-specific mortality. Environ. Int. 134, 105341.
Marshall, J.D., Nethery, E., Brauer, M., 2008. Within-urban variability in ambient air pollution: comparison of estimation methods. Atmos. Environ. 42 (6), 1359–1369.
RIVM, DCMR, and GGD Amsterdam, 2018. Concentratie stikstofdioxide in lucht. [cited 2020 23-04]; Available from: https://www.clo.nl/indicatoren/nl0231-stikstofdioxide.
Samoli, E., et al., 2020. The impact of measurement error in modeled ambient particles exposures on health effect estimates in multilevel analysis: A simulation study. Environ. Epidemiol. 4 (3), e094.
Sellier, Y., et al., 2014. Health effects of ambient air pollution: do different methods for estimating exposure lead to different results? Environ. Int. 66, 165–173.
Van Jaarsveld, J., 2004. The operational priority substances model.
Velders, G.J., et al., 2020. Effects of European emission reductions on air quality in the Netherlands and the associated health effects. Atmos. Environ. 221, 117109.
Velders, G.J., Diederen, H.S., 2009. Likelihood of meeting the EU limit values for NO2 and PM10 concentrations in the Netherlands. Atmos. Environ. 43 (19), 3060–3069.
Wang, M., et al., 2013. Evaluation of land use regression models for NO2 and particulate matter in 20 European study areas: the ESCAPE project. Environ. Sci. Technol. 47 (9), 4357–4364.
Wang, M., et al., 2015. Air pollution and lung function in dutch children: A comparison of exposure estimates and associations based on land use regression and dispersion exposure modeling approaches. Environ. Health Perspect. 123 (8), 847–851.
Wesseling, J., Beijk, R., Bezemer, A., 2011. An efficient modeling system for nation-wide compliance testing. In: 14th Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, 2-6 October 2011.
Wesseling, J., Visser, G.T., 2003. An inter-comparison of the TNO Traffic Model, field data and wind tunnel measurements. TNO Report 207, 2003.