

## University of Groningen

### (Un)Healthy in the city

Zijlema, Wilma

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*

Publisher's PDF, also known as Version of record

*Publication date:*

2016

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Zijlema, W. (2016). *(Un)Healthy in the city: Adverse health effects of traffic-related noise and air pollution*. University of Groningen.

**Copyright**

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

**Take-down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*

# 2

## The LifeLines Cohort Study: a resource providing new opportunities for environmental epidemiology

Wilma L. Zijlema, Nynke Smidt, Bart Klijs, David W. Morley, John Gulliver, Kees de Hoogh, Salome Scholtens, Judith G.M. Rosmalen, Ronald P. Stolck

*Submitted*

## **ABSTRACT**

### **Background**

Lifelines is a population-based cohort study investigating healthy ageing among 167,729 participants from the Netherlands. We describe the reasons for choosing particular assessments of environmental exposures and consider the implications for future investigations.

### **Methods**

Exposure to ambient air pollution was estimated using two types of land use regression models. Road traffic noise exposure was assessed using a model based on Common Noise Assessment Methods in Europe. Data on perceived exposures and neighborhood characteristics are also available. A comprehensive medical assessment and questionnaires were completed.

### **Results**

Mean age was 45 years (standard deviation (SD) 13 years), and 59% were female. Median levels of NO<sub>2</sub> and PM<sub>10</sub> were 15.7 (interquartile range (IQR) 4.9) µg/m<sup>3</sup> and 24.0 (IQR 0.6) µg/m<sup>3</sup> respectively. Median level of daytime road traffic noise was 54.0 (IQR 4.2) dB(A).

### **Conclusion**

The combination of harmonized environmental exposures and extensive assessment of a variety of health outcomes offers opportunities for environmental epidemiology. LifeLines aims to be a resource for the international scientific community.

## BACKGROUND

Environmental epidemiology has contributed to public health by focusing on the reduction of harmful exposures from the environment. Smoke-free legislation [1], control of particulate air pollution [2], and the introduction of proper sanitation [3] are examples of policies that resulted in substantial benefits for public health. Although large efforts have been made, a lot more is to gain in environmental epidemiology, both in terms of public health policy and research. Prevention of harmful environmental exposures is important because many of these exposures are often involuntary and outside the immediate control of the individual. Moreover, many individuals may be affected by the same pollution source.

Studying health effects of environmental exposures can be challenging. Environmental exposures often occur in low concentrations and in complex mixtures. Most health outcomes of interest have many underlying risk factors whose effects may be much stronger than those of the environmental exposure. Risks associated with environmental exposures are generally small [4], making them often difficult to detect. Although effect sizes are small, impacts on health are still substantial because of the large number of exposed persons. Population based cohort studies with large sample sizes are needed to investigate the complex interaction between genetic, behavioral, and environmental factors in the development of multifactorial chronic diseases. Large population based studies can make important contributions to environmental epidemiology. Moreover, collaboration between cohort studies could help understand environmental causes of disease. The advantages of such an approach are widespread. Collaboration between studies will lead to very large sample sizes and wide exposure ranges. Furthermore, results can be compared across regions, in which not only the exposures, but also genetic, social and cultural factors vary. This makes the generalizability of study results larger. Being able to generalize study results across large regions in Europe is important for establishing European exposure norms and guidelines. Therefore, population studies must be harmonized and standardized, enabling the assembling of data in valid and effective ways. To this end, LifeLines aims to implement standardized and harmonized data collection methods. LifeLines is one of eight studies participating in the BioSHaRE (Biobank Standardisation and Harmonisation for Research Excellence in the European Union, [www.bioshare.eu](http://www.bioshare.eu)) project. Within the framework of this European harmonization initiative, tools were developed for data harmonization, database integration and federated data analyses [5]. Other participating biobanks are located in Germany, Norway, the UK, Finland, Italy, Estonia, and Ireland. As a result, standardized and validated methods for measurement were used wherever possible, and a part of the dataset

(including e.g. educational level, alcohol and tobacco use) has been retrospectively harmonized with other European cohorts [6]. With its large sample size, and the combination of harmonized environmental exposures and extensive assessment of a variety of health outcomes, LifeLines aims to be a resource for environmental epidemiology.

Our aim is to inform researchers on the choice of data collection methods and methodology of data on environmental exposures in LifeLines. Descriptive information of selected exposure and cohort data will also be reported. By describing these features of LifeLines we aim to highlight methodological choices made in the context of large population based cohort studies exposed to multiple environmental stressors.

## **METHODS**

### **Study design and participants**

The LifeLines Cohort Study was established to facilitate research on complex interactions between environmental, phenotypic and genetic factors in the development of chronic diseases [7]. The cohort profile of LifeLines has been described elsewhere [8]. Briefly, the recruitment of study participants took place between 2006-2013. In the Netherlands, all inhabitants are registered with a general practitioner. A large number of general practitioners from the Northern provinces of the Netherlands participated in the recruitment and invited all their patients between ages 25 and 50 years. Individuals who agreed to participate were asked to indicate whether their family members would also be willing to participate. In addition, individuals could also register themselves via the LifeLines website. The sample was recruited from the three Northern provinces of the Netherlands, but this was not a requirement for inclusion of family members [8]. Characteristics of adult LifeLines participants are broadly representative for the adult population of the north of the Netherlands [9].

Baseline data were collected from 167,729 participants, aged 6 months to 93 years. Follow-up is planned for at least 30 years, with questionnaires administered every 1.5 years, and a renewed physical examination scheduled every five years. Participants visited one of the LifeLines research sites for a physical examination, including spirometry, electrocardiogram (ECG), blood pressure measurements, anthropometry, cognition tests, and a psychiatric interview. Fasting blood and 24-hour urine samples were collected from all participants, and genome wide association data are currently available for a subsample of 15,638 participants. An extensive baseline questionnaire was completed at home, including questions

on history of illness, health related quality of life, lifestyle, socioeconomic status, psychosocial stress, work (profession, working hours), psychosocial characteristics, and medication use. Linkage will be established with records from general practitioners and health registries. A comprehensive and detailed overview of the available data is presented in the online LifeLines Data Catalogue ([www.lifelines.net](http://www.lifelines.net)).

### **Geocoding**

All participants' home addresses were geocoded. In addition to the most recent home address, address history is available from the Municipal Personal Record Database. This governmental registry contains personal data of all individuals who live or have lived in the Netherlands. Data on address history provide insights in residential mobility and length of exposure to different environments. This allows assessment of long-term exposures, which is relevant for life course epidemiology. Work addresses were also collected and will be geocoded as well, allowing for outdoor exposure estimation of air pollution and noise at both home and work location. Most previous studies estimate exposure to road traffic noise and air pollution at the home addresses of participants (e.g. [10–12]), assuming this to be a good indicator of personal exposure. Individuals spend a large amount of their time at the work address. Combining exposures from both home and work locations will result in better estimations of an individual's exposure [13]. A detailed description of the environmental exposures assessed in LifeLines is described below, and is summarized in Table 1.

### **Environmental exposures**

#### *Ambient air pollution*

Air pollution has been related to various health outcomes, including respiratory diseases [14] and cardiovascular disease [15]. In LifeLines, exposure to ambient air pollution was estimated using land use regression (LUR) models developed for the European Study of Cohorts for Air Pollution Effects (ESCAPE) [16,17] and using European wide LUR models enhanced with satellite derived air pollution estimates [18]. It was chosen to implement these air pollution models in LifeLines because both models are advantageous regarding comparability across studies in other European regions.

Within the ESCAPE project, LUR models for various European study areas were developed using a standardized approach [16,17]. ESCAPE LUR models were developed for NO<sub>2</sub> (nitrogen dioxide), NO<sub>2</sub> background, PM<sub>2.5</sub> (particulate matter with a diameter ≤2.5 μm), PM<sub>2.5</sub> absorbance (reflectance on PM<sub>2.5</sub> filters, i.e. a

**Table 1.** Overview of environmental exposures in the LifeLines Cohort Study

<b>Ambient air pollution</b>	
<i>ESCAPE models</i>	
NO <sub>2</sub>	nitrogen dioxide
NO <sub>2</sub> background	background level of nitrogen dioxide
PM <sub>2.5</sub>	particulate matter with diameter ≤2.5 μm
PM <sub>2.5</sub> absorbance	reflectance on PM <sub>2.5</sub> filters, i.e. marker of black carbon
PM <sub>10</sub>	particulate matter with diameter ≤10 μm
<i>EU-wide models</i>	
NO <sub>2</sub>	nitrogen dioxide
PM <sub>10</sub>	particulate matter with diameter ≤10 μm
<b>Road traffic noise (CNOSSOS-EU model)</b>	
L <sub>day</sub>	12-hour day time period from 07:00 to 19:00 hour
L <sub>evening</sub>	4-hour evening time period from 19:00 to 23:00 hours
L <sub>night</sub>	8-hour night time period from 23:00 to 07:00 hours
L <sub>aeq16</sub>	16-hour day and evening time period from 07:00 to 23:00 hours
L <sub>den</sub>	day-evening-night time period of 24 hours
L <sub>aeq</sub> 0-23 hours	Hourly noise estimates
<b>Questionnaire data</b>	
Noise annoyance	
Perceived exposures to power lines, mobile phone masts	
Mobile phone use	
Exposure to secondhand smoke	
Perceived living environment	
<b>Database linkage</b>	
Neighborhood characteristics (Statistics Netherlands)	Neighborhood level demographic and socioeconomic figures
LISA employment register	Location, type of establishment, number of employees

ESCAPE = European Study of Cohorts for Air Pollution Effects; EU-wide = European-wide; CNOSSOS-EU = Common Noise Assessment Methods in Europe

marker of black carbon), and PM<sub>10</sub> (particulate matter with a diameter ≤10 μm), and were based on annual average concentrations from an intensive monitoring campaign and GIS (geographic information system) derived predictor variables (e.g. distance to the nearest major road, traffic intensity, built-up land, population density, altitude). LUR models were developed using measurements carried out in 2009-2010 and predictor variable data for the same years. Model performance was evaluated by leave-one-out cross validation [16,17]. The adjusted explained variability in measured concentrations (R<sup>2</sup>) was 0.85 for NO<sub>2</sub>, 0.83 for NO<sub>2</sub> background, 0.66 for PM<sub>10</sub>, 0.64 for PM<sub>2.5</sub>, and 0.91 for PM<sub>2.5</sub> absorbance.

In addition, exposure to ambient PM<sub>10</sub> and NO<sub>2</sub> is available from European (EU)-wide models. These EU-wide models incorporate GIS-derived land use, road network, and topographic data, as well as satellite-derived estimates of ground level concentrations for PM<sub>2.5</sub> (as an indicator of PM<sub>10</sub>) and NO<sub>2</sub> [18]. Model development follows the ESCAPE procedure to construct the multiple linear regression equations, and are applicable for years 2005, 2006 (NO<sub>2</sub>) and 2007 (NO<sub>2</sub> and PM<sub>10</sub>). Models were evaluated against measured PM<sub>10</sub> and NO<sub>2</sub> concentrations at an independent subset of sites reserved for this purpose. The adjusted explained variability in measured concentrations (R<sup>2</sup>) was 0.48–0.58 for NO<sub>2</sub>, and 0.22–0.50 for PM<sub>10</sub>.

The main difference between the ESCAPE and EU-wide models is that the ESCAPE models are region specific, while EU-wide models are developed for a much larger area. ESCAPE models are developed for 20 (PM) to 36 (NO<sub>2</sub>) European regions and EU-wide models for 17 countries in Western Europe. In addition, monitoring data used in ESCAPE models originated from a monitoring campaign specifically conducted for the ESCAPE-project with monitoring sites selected for this purpose, whereas monitoring data for the EU-wide models were obtained from regulatory monitoring networks. ESCAPE models may perform better on a regional scale, while EU-wide models may make comparisons across multiple study areas easier, since the model is available for a larger region. A study including Lifelines and other cohorts falling in ESCAPE study areas (e.g. the British cohort EPIC-Oxford, also involved in BioSHaRE) should use air pollution exposures estimates from the relevant ESCAPE models. However, the EU-wide model can also provide air pollution exposure estimates for areas falling outside the ESCAPE study areas, and will therefore allow research including Lifelines plus non-ESCAPE cohorts (e.g. the Norwegian cohort HUNT, also involved in BioSHaRE).

### *Road traffic noise*

Environmental noise has been related to a variety of adverse outcomes, including hearing loss, annoyance, sleep disturbance, cognitive impairment, and cardiovascular disease [19]. In LifeLines, road traffic noise was estimated using an implementation of the Common Noise Assessment Methods in Europe (CNOSSOS-EU) noise modelling framework [20,21], which was developed as a common methodology for noise modelling across Europe. This noise model was preferred since it allows comparison of results from different countries. The CNOSSOS-EU noise model implemented within LifeLines is using lower resolution source data because the highest resolution input data at national or large regional level is either unavailable, expensive or would be too computationally intensive to process. The performance of CNOSSOS-EU using the lower resolution inputs has been shown



to be reasonable for application in epidemiological studies. The model's exposure ranking, i.e. prediction of noisier and quieter sites, was adequate (Spearman's rank = 0.75;  $p < 0.001$ ), but the predicted noise levels have relatively large errors (root mean square error (RMSE) = 4.46 dB(A)) [22].

The CNOSSOS-EU framework contains empirically derived equations to determine the initial noise level based on traffic flow and sound attenuation (i.e. sound reduction or damping) based on known environmental factors and physical processes. To estimate source noise on road segments in the Netherlands, traffic information was obtained including hourly flow of passenger cars, heavy goods vehicles and their average speeds. The sound propagation model was based on the CORINE (Coordination of information on the environment) land cover dataset that has a European wide coverage accurate to 100 meters for major land cover types [23]. In particular, the distinction between urban fabric and areas of vegetation was made. Traffic data originated from year 2009 and landcover data from 2006. Full details of the noise model used for the LifeLines study are provided elsewhere [22].

### *Perceived exposures*

Besides the actual exposure to environmental factors, also perceived exposure and concerns about the health risks associated with the exposure might influence health outcomes [24]. When studying particular exposures, for example electromagnetic fields, the public's perception of the health risks is as relevant to health as the exposure itself. Moreover, sometimes perceived exposures have stronger associations with diminished health than the actual exposures [25–28]. In addition to the modelled estimates for air pollution and noise exposures, perceived exposures are measured in LifeLines. Noise annoyance from eight different sources was measured using a standardized self-report questionnaire. The sources of noise annoyance include for example air, road and rail traffic. This questionnaire originates from the International Organization for Standardization (ISO) guideline, which provides specifications for socio-acoustic surveys and social surveys that include questions on noise effects [29]. Similar noise annoyance questions were used in the HYENA study on Hypertension and Exposure to Noise near Airports [30]. Perceived exposure to electromagnetic fields is measured using a questionnaire on perceived exposure to power lines and mobile phone masts. Participants were asked to what extent they think they are exposed to radiation from power lines and mobile phone masts, and whether they perceive this as bad for their health (adapted from [31]). In addition, a number of questions on the use of mobile phones (e.g. average time per week using mobile phone; on which side of the head) were included. These questions were adapted from the UKBiobank

questionnaire [32], which is one of the cohorts involved in the BioSHaRE project. Exposure to second-hand smoke is assessed with questions about the duration and place (in the household, at the workplace) of exposure, and originate from the European Community Respiratory Health Survey (ECRHS, [33]).

In addition to work addresses, information was collected on working hours and type of profession, which may be relevant for occupational exposure studies. In the next follow-up questionnaire, LifeLines will also measure how participants perceive their living environment. The questionnaire comprises of nine items investigating characteristics of the physical and social living environment as perceived by the respondent (e.g. neighborhood satisfaction, social interaction with neighbors), and was based on the 2010 health survey of the Dutch Community Health Services (Dutch name: GGD Gezondheidsenquête 2010) [34] and the 2006 WoON questionnaire (Dutch name: WoonOnderzoek Nederland) [35].

### *Neighborhood characteristics*

Various neighborhood characteristics have been associated with health, ranging from cardiometabolic risk factors [36] to life expectancy [37]. Data on neighborhood characteristics from Statistics Netherlands are available for linkage to the LifeLines database. Statistics Netherlands publishes demographic and socio-economic figures for municipalities, districts and neighborhoods [38]. These figures cover various themes, for example housing, education, income, and land use. Such information enables to investigate the impact of neighborhood conditions on health [39].

Furthermore, the LISA employment register ([www.lisa.nl](http://www.lisa.nl)) is linked to the LifeLines database. This register contains nationwide information on locations (geocoded at the address level) of establishments where paid work is done. The database also contains information about the type of establishment (i.e. restaurants, hospitals, shops) and the number of employees. Using the LISA employment register, it is possible to investigate the density and distance to specific facilities in relation to various health outcomes.

## **RESULTS**

The first data release of the baseline sample consisted of 95,432 participants, of which 58.7% were female. Mean age was 45.2 years (standard deviation (SD) 12.6 years), and men were more often current smokers than women (23.8% and 20.4%, respectively). Overall, mean body mass index (BMI) was 26.1 kg/m<sup>2</sup> (SD 4.3 kg/m<sup>2</sup>), and mean systolic blood pressure was 125.6 mmHg (SD 15.3 mmHg).

Forced vital capacity (FVC), a measure for lung function assessed with spirometry, was on average 4.5 (SD 1.1) liters. Most participants live in rural (i.e. <500 addresses per km<sup>2</sup>) neighborhoods (41.2%) (Table 2).

**Table 2.** Sample characteristics by sex of the LifeLines Cohort Study

	Women	Men	Total
N (%)	56 053 (58.7)	39 379 (41.3)	95 432
Age (years)	44.9 (12.6)	45.7 (12.7)	45.2 (12.6)
Education (%)			
No or primary	3.2	3.0	3.1
Lower or preparatory vocational	12.4	16.2	14.0
Lower general secondary	16.0	11.7	14.2
Intermediate vocational or apprenticeship	30.6	31.0	30.8
Higher general secondary or pre-university secondary	9.9	7.0	8.7
Higher vocational or university	28.1	31.0	29.3
Current smokers (%)	20.4	23.8	21.8
BMI (kg/m <sup>2</sup> )	25.9 (4.7)	26.4 (3.7)	26.1 (4.3)
SBP (mmHg)	122.1 (15.3)	130.7 (14.0)	125.6 (15.3)
DBP (mmHg)	71.8 (8.7)	76.6 (9.3)	73.8 (9.3)
FVC (L)	3.9 (0.6)	5.4 (0.9)	4.5 (1.1)
FEV1 <sup>e</sup> (L)	3.0 (0.6)	4.1 (0.8)	3.5 (0.8)
Urbanity <sup>b</sup> (%)			
Rural	40.8	41.9	41.2
Semi-rural	24.3	25.4	24.8
Intermediate urban-rural	17.5	16.7	17.1
Semi-urban	10.6	10.1	10.4
Urban	6.8	6.0	6.5

Means (SD) are presented for continuous variables, and percentages are presented for categorical variables.

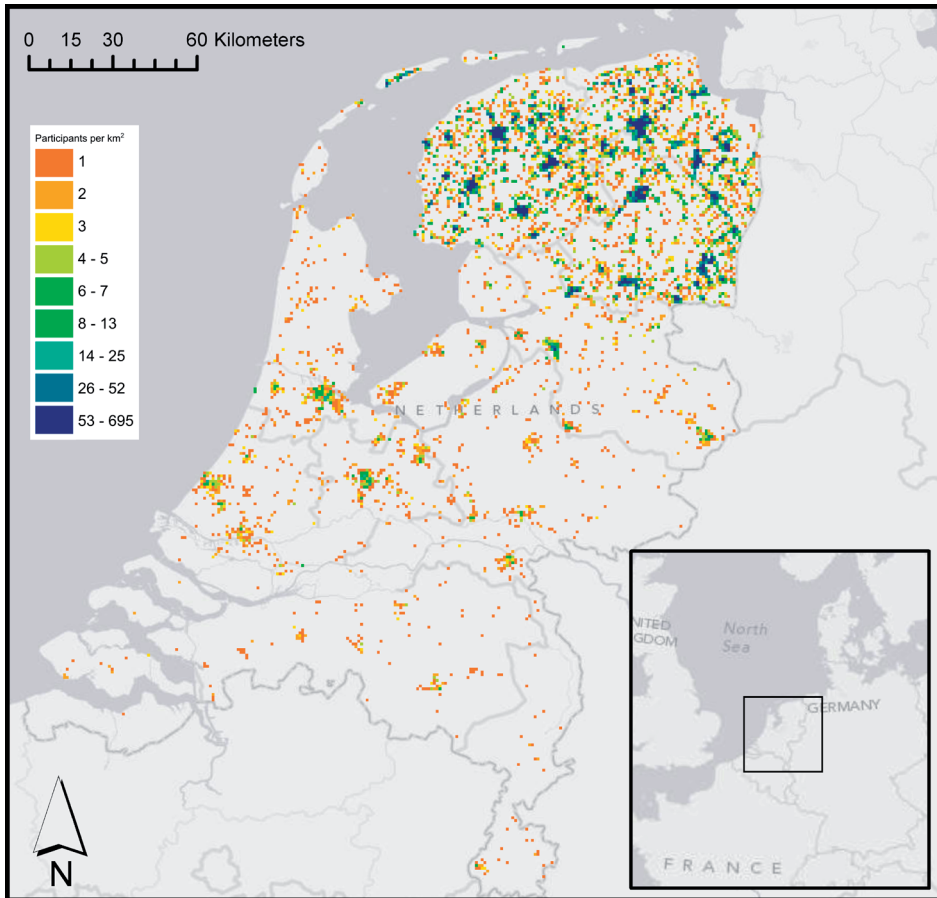
<sup>a</sup>Data are based on the first data release of n=95,432.

<sup>b</sup>Average number of addresses per km<sup>2</sup> within a range of 1 kilometer, categorized into five levels ranging from rural (<500 addresses per km<sup>2</sup>) to urban (≥2500 addresses per km<sup>2</sup>).

Abbreviations: BMI=body mass index; SBP=systolic blood pressure; DBP=diastolic blood pressure; FVC=forced vital capacity; FEV1= forced expiratory volume in 1 second

Most participants live in the three Northern provinces of the Netherlands, but part of the participants (approximately 3%) live elsewhere in the country (Figure 1). Median levels of NO<sub>2</sub> were 15.7 (interquartile range (IQR) 4.9) µg/m<sup>3</sup> (ESCAPE) and 20.6 (IQR 7.9) µg/m<sup>3</sup> (EU-wide), and 24.0 (IQR 0.6) µg/m<sup>3</sup> (ESCAPE) and 23.6 (IQR 2.4) µg/m<sup>3</sup> (EU-wide) for PM<sub>10</sub>.

Distributions of estimated concentrations of NO<sub>2</sub> and PM<sub>10</sub> are shown in Figure 2. Correlation between ESCAPE-LUR modelled NO<sub>2</sub> and satellite-enhanced LUR modelled NO<sub>2</sub> was high (Spearman's rho = 0.86), while correlation for PM<sub>10</sub> from both models was moderate (Spearman's rho = 0.54).

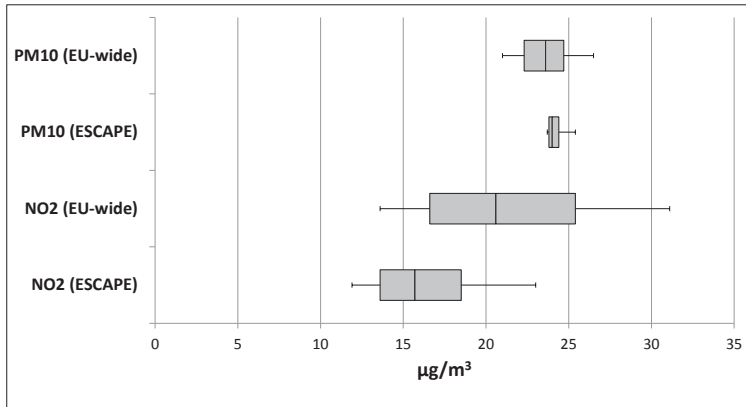


**Figure 1.** LifeLines study area and number of participants per square kilometer <sup>a</sup>

<sup>a</sup> Participants were aggregated and plotted at the center of each 1 kilometer grid cell.

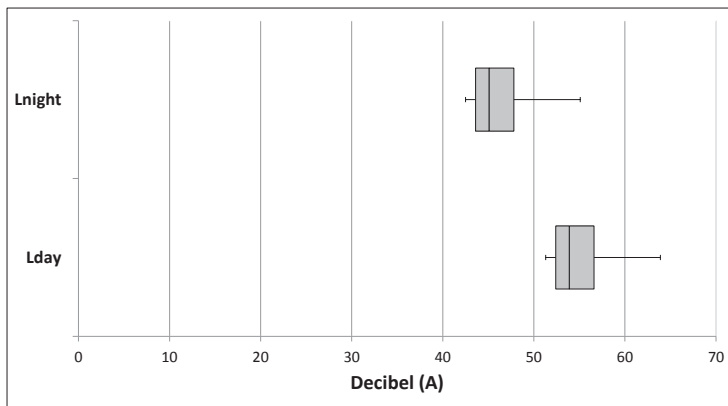
Five A-weighted indicators of road traffic noise were estimated from the CNOS-SOS-EU model. The noise indicators are  $L_{\text{day}}$  (12-hour day time period from 07:00 to 19:00 hour);  $L_{\text{evening}}$  (4-hour evening time period from 19:00 to 23:00 hours);  $L_{\text{night}}$  (8-hour night time period from 23:00 to 07:00 hours);  $L_{\text{aeq16}}$  (16-hour day and evening time period from 07:00 to 23:00 hours); and  $L_{\text{den}}$  (day-evening-night time period of 24 hours). Median levels of road traffic noise were 54.0 (IQR 4.2) dB(A) for  $L_{\text{day}}$  and 45.1 (IQR 4.2) dB(A) for  $L_{\text{night}}$ . Figure 3 shows distributions of daytime and nighttime road traffic noise estimates for LifeLines.

LifeLines participants living in urban neighborhoods had highest exposure to air pollution (nitrogen dioxide; Figure 4) and 24-hours road traffic noise (Figure 5), compared to participants in neighborhoods of lower degree of urbanity.



**Figure 2.** Distribution of nitrogen dioxide and particulate matter ( $\leq 10 \mu\text{m}$ ) exposure in the LifeLines Cohort Study with estimates based on ESCAPE land use regression model and EU-wide land use regression model. Medians, 25<sup>th</sup>, and 75<sup>th</sup> percentiles are shown in the box, whiskers indicate 5<sup>th</sup> and 95<sup>th</sup> percentiles.

ESCAPE = European Study of Cohorts for Air Pollution Effects; EU-wide = European-wide.

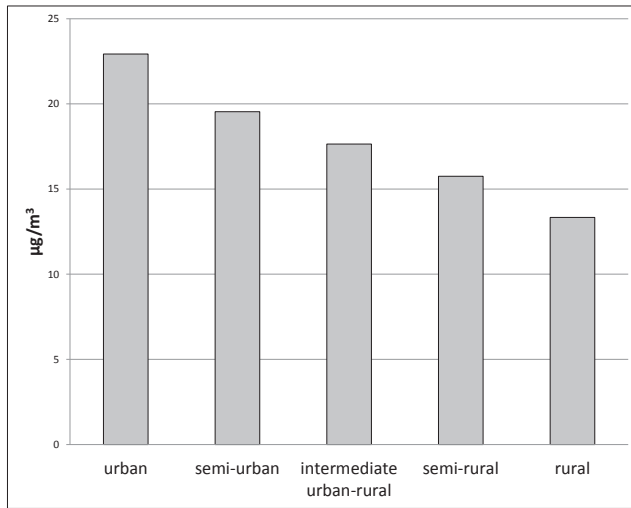


**Figure 3.** Distribution of road traffic noise exposure  $L_{\text{day}}$  and  $L_{\text{night}}$  in the LifeLines Cohort Study. Medians, 25<sup>th</sup>, and 75<sup>th</sup> percentiles are shown in the box, whiskers indicate 5<sup>th</sup> and 95<sup>th</sup> percentiles.

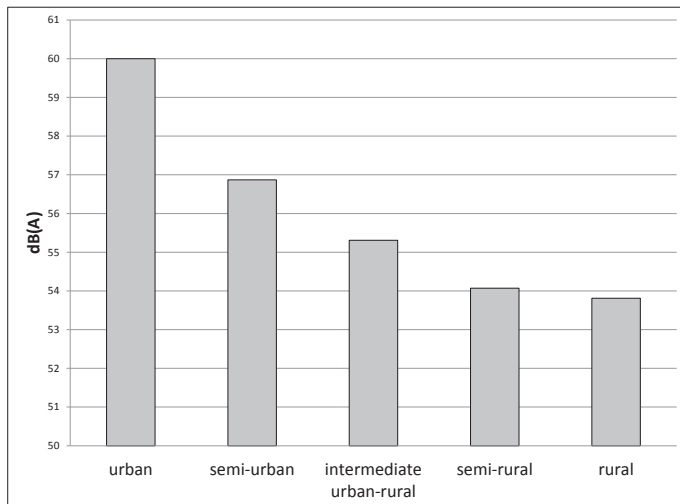
dB(A) = decibel (A);  $L_{\text{day}}$  = A-weighted equivalent noise level over the 12-hour day time period from 07:00 to 19:00 hours;  $L_{\text{night}}$  = A-weighted equivalent noise level over the 8-hour night time period from 23:00 to 07:00 hours.

## DATA ACCESS AND LINKAGE

LifeLines has adopted an open protocol, meaning that within the standing infrastructure additional data and biomaterial collection, and linkage with other (environmental) data sources can be implemented by submitting a research proposal to LifeLines. Environmental biomonitoring of exposure and response could



**Figure 4.** Median exposure to ambient nitrogen dioxide ( $\text{NO}_2$ ; based on ESCAPE model) according to degree of urbanity within the LifeLines Cohort Study.



**Figure 5.** Median exposure to road traffic noise ( $L_{den}$ ; average A-weighted noise level estimated over a 24 hour period) according to degree of urbanity within the LifeLines Cohort Study.

also be implemented. One example is LifeLines DEEP, an add-on study where in a subsample of participants additional biological samples (feces, exhaled air) were collected, additional blood analyses were undertaken, and additional questionnaires were filled out [40]. Other exposures types, such as domestic radon, electromagnetic fields, harmful chemicals (e.g. pesticides) are of interest for research

in LifeLines, and proposals for additional environmental exposures assessments relevant for 'Healthy Ageing' are warmly welcomed. Furthermore, biological samples are stored for future analyses, enabling for example measurement of exposure biomarkers. Data and biomaterials are provided on a fee-for-service basis and may be used for scientific research only. Public and private researchers, from inside and outside the Netherlands are invited to submit a research proposal to the LifeLines Research Office. Quality control of the data of hard copy questionnaires consist of various stages, including preparation for scanning the questionnaire, scanning the questionnaires, verification of the data on missing values, inconsistencies, and extreme values and a final data quality control is conducted in the database. Quality control of the data is done by trained medical students and data managers, using Standard Operating Procedures (SOPs). Data is released within a remote system (LifeLines workspace) running on a high performance computer cluster, which ensures data quality and security. The LifeLines research website ([www.lifelines.net](http://www.lifelines.net)) provides the details of the application process, the data collection, and an overview of publications with LifeLines data.

## **DISCUSSION**

The combination of harmonized environmental exposures, relevant mediators and modifiers, and extensive assessment of multiple health outcomes makes LifeLines a great resource for environmental epidemiology. This paper provides an overview of assessments in the LifeLines Cohort Study that are relevant to environmental epidemiology.

### **Limitations and strengths**

The results in this paper are based on the first 95,432 participants that were included in LifeLines. Geocoding and exposure estimation using the noise and air pollution models of the full cohort is currently still ongoing. Since the inclusion of participants was independent of their place of residence, we have no reason to suspect geographical differences between participants in the first data release and the complete sample. Our study area is relatively rural compared to other parts of the Netherlands [41]. As a consequence, levels of air pollution and noise exposures are lower than other parts of the Netherlands, due to for example lower population densities, and less extensive road networks. For example, exposure to NO<sub>2</sub> in the EPIC-PROSPECT cohort located in the city of Utrecht and surrounding areas was on average 26.7 µg/m<sup>3</sup> [42], compared to 15.7 µg/m<sup>3</sup> in LifeLines. Conclusions based on research undertaken with LifeLines data will therefore be

limited to exposure levels in that particular range. Nevertheless, participants live in both urban and rural areas and are spread across a large part of the Netherlands, resulting in a great spatial variation. A major challenge in environmental epidemiology includes accurate exposure assessment [24,43,44]. Actual measurement of individual-level exposures in a cohort as large as LifeLines would be impossible. Therefore other approaches to estimate exposures are used, such as land use regression modelling. Use of these models introduces misclassification of exposure to a varying degree; for example, due to daily mobility and long-term residential mobility [45]. In LifeLines, misclassification due to residential mobility can be tackled because data is available on address history, which is for some participants available from periods as early as year 1943. Future research should focus on characterizing exposures in earlier years, allowing for assessment of long-term exposures which is relevant for life course epidemiology. Data about the number of hours a person works and work location could provide insights into daily mobility. In addition, the collection of work addresses allows for outdoor exposure estimation of noise and air pollution at the work location. Combining exposures at both home and work location will result in better estimation of an individual's exposure [13]. Another factor contributing to exposure misclassification is for example the type of housing (e.g. insulation quality, bedroom location). We did not collect information with regard to type of housing, but this is planned for the follow up questionnaires.

Ethical and legal challenges involved with the collection, storage and sharing of biobank data are addressed with procedures and facilities to ensure privacy protection. Data is released via a remote desktop from which pseudonymized data can be accessed. This workspace runs on a high-performance computer cluster, ensuring data quality and security. High quality standards apply to the entire process of data and biomaterial collection, storage and release. Privacy, security and traceability are ensured in all aspects, and in 2014, LifeLines received the ISO-9001 quality certificate [8]. For example, geocoded addresses combined with individual data are not available for data release to researchers. In addition to the risk of re-identification of participants, publications of local health hazards could expose (non-)participants to risks to their social rights and private economic interests. As population biobanks have a legal obligation to inform their participants about the risks of their participation, it has been recommended that they also acknowledge the risks posed by the publication of exposure studies to the legitimate interests of their participants and the public [46]. Effectively, this means that LifeLines researchers are not allowed to publish detailed exposure maps for the LifeLines area.



With 167,729 participants, LifeLines is one of the largest population based cohort studies of the world. A large amount of data is collected, biological samples are stored for future analyses (e.g. measurement of exposure biomarkers), and for a subsample of 15,638 participants genome wide genotype data are available. These numbers are large enough for studying effects of environmental exposures in vulnerable subgroups. This is important, because many factors (genetics, individual disease states, psychosocial stress, and socioeconomic status) have the potential to interact with environmental exposures [47]. The major strength of this study is the use of harmonized exposure models for ambient air pollution and for road traffic noise, and the use of validated questionnaires. This, and also data harmonization facilitates comparability and combination with data from other cohorts and regions. This is a great advantage especially for environmental epidemiology, where large sample sizes and broad exposure ranges are needed. Existing and future collaborations with other biobanks and international consortia hold the promise of answering complex questions in environmental epidemiology. Furthermore, the prospective nature of LifeLines allows research into long-term effects of environmental exposures. In addition, both objective (modelled) and subjective (questionnaire based) exposures were assessed. This enables studying effects of the exposure itself, and of perception of the exposure. Key research questions that will be investigated in LifeLines are about effects of noise and air pollution on healthy ageing.

## **CONCLUSIONS**

LifeLines aims to be a resource for the national and international scientific community. Requests for data and biomaterials can be submitted to the LifeLines Research Office ([LLscience@umcg.nl](mailto:LLscience@umcg.nl)).

## REFERENCES

1. Been J V, Nurmatov UB, Cox B, Nawrot TS, van Schayck CP, Sheikh A. Effect of smoke-free legislation on perinatal and child health: a systematic review and meta-analysis. *Lancet*. 2014;383: 1549–60. doi:10.1016/S0140-6736(14)60082-9
2. Clancy L, Goodman P, Sinclair H, Dockery DW. Effect of air-pollution control on death rates in Dublin, Ireland: an intervention study. *Lancet*. 2002;360: 1210–4. doi:10.1016/S0140-6736(02)11281-5
3. Hamlin C, Sheard S. Revolutions in public health: 1848, and 1998? *BMJ*. 1998;317: 587–591.
4. Pekkanen J, Pearce N. Environmental epidemiology: Challenges and opportunities. *Environ Health Perspect*. 2001;109: 1–5. doi:10.2307/3434913
5. Doiron D, Burton P, Marcon Y, Gaye A, Wolffenbuttel BH, Perola M, et al. Data harmonization and federated analysis of population-based studies: the BioSHARe project. *Emerg Themes Epidemiol*. 2013;10: 12. doi:10.1186/1742-7622-10-12
6. Van Vliet-Ostaptchouk J V, Nuotio ML, Slagter SN, Doiron D, Fischer K, Foco L, et al. The prevalence of metabolic syndrome and metabolically healthy obesity in Europe: a collaborative analysis of ten large cohort studies. *BMC Endocr Disord*. 2014;14: 9. doi:10.1186/1472-6823-14-9
7. Stolk RP, Rosmalen JGM, Postma DS, de Boer RA, Navis G, Slaets JJP, et al. Universal risk factors for multifactorial diseases: LifeLines: a three-generation population-based study. *Eur J Epidemiol*. 2008;23: 67–74. doi:10.1007/s10654-007-9204-4
8. Scholtens S, Smidt N, Swertz MA, Bakker SJ, Dotinga A, Vonk JM, et al. Cohort Profile: LifeLines, a three-generation cohort study and biobank. *Int J Epidemiol*. 2014; doi: 10.1093/ije/dyu229. doi:10.1093/ije/dyu229
9. Klijs B, Scholtens S, Mandemakers JJ, Snieder H, Stolk RP, Smidt N. Representativeness of the LifeLines Cohort Study. Ali RI, editor. *PLoS One*. 2015;10: e0137203. doi:10.1371/journal.pone.0137203
10. Foraster M, Basagaña X, Aguilera I, Rivera M, Agis D, Bouso L, et al. 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). *Environ Health Perspect*. 2014;122: 404–11. doi:10.1289/ehp.1306497
11. Fuks KB, Weinmayr G, Foraster M, Dratva J, Hampel R, Houthuijs D, et al. Arterial blood pressure and long-term exposure to traffic-related air pollution: an analysis in the European Study of Cohorts for Air Pollution Effects (ESCAPE). *Environ Health Perspect*. 2014;122: 896–905. doi:10.1289/ehp.1307725
12. Fischer PH, Marra M, Ameling CB, Hoek G, Beelen R, de Hoogh K, et al. Air Pollution and Mortality in Seven Million Adults: The Dutch Environmental Longitudinal Study (DUELS). *Environ Health Perspect*. 2015; doi:10.1289/ehp.1408254
13. Ragettli MS, Phuleria HC, Tsai M-Y, Schindler C, de Nazelle A, Ducret-Stich RE, et al. The relevance of commuter and work/school exposure in an epidemiological study on traffic-related air pollution. *J Expo Sci Environ Epidemiol*. 2014; doi:10.1038/jes.2014.83
14. Götschi T, Heinrich J, Sunyer J, Künzli N. Long-term effects of ambient air pollution on lung function: a review. *Epidemiology*. 2008;19: 690–701. doi:10.1097/EDE.0b013e318181650f
15. Newby DE, Mannucci PM, Tell GS, Baccarelli AA, Brook RD, Donaldson K, et al. Expert position paper on air pollution and cardiovascular disease. *Eur Heart J*. 2014;36: 83–93b. doi:10.1093/eurheartj/ehu458

16. Beelen R, Hoek G, Vienneau D, Eeftens M, Dimakopoulou K, Pedeli X, et al. Development of NO<sub>2</sub> and NO<sub>x</sub> land use regression models for estimating air pollution exposure in 36 study areas in Europe – The ESCAPE project. *Atmos Environ*. 2013;72: 10–23. doi:10.1016/j.atmosenv.2013.02.037
17. Eeftens M, Beelen R, de Hoogh K, Bellander T, Cesaroni G, Cirach M, et al. Development of Land Use Regression models for PM<sub>2.5</sub>, PM<sub>2.5</sub> absorbance, PM<sub>10</sub> and PM<sub>coarse</sub> in 20 European study areas; results of the ESCAPE project. *Environ Sci Technol*. 2012;46: 11195–11205. doi:10.1021/es301948k
18. Vienneau D, de Hoogh K, Bechle MJ, Beelen R, van Donkelaar A, Martin R V, et al. Western European land use regression incorporating satellite- and ground-based measurements of NO<sub>2</sub> and PM<sub>10</sub>. *Environ Sci Technol*. 2013;47: 13555–64. doi:10.1021/es403089q
19. Basner M, Babisch W, Davis A, Brink M, Clark C, Janssen S, et al. Auditory and non-auditory effects of noise on health. *Lancet*. 2013;383: 1325–32. doi:10.1016/S0140-6736(13)61613-X
20. Kefalopoulos S, Paviotti M, Anfosso-Lédée F, Van Maercke D, Shilton S, Jones N. Advances in the development of common noise assessment methods in Europe: The CNOSSOS-EU framework for strategic environmental noise mapping. *Sci Total Environ*. 2014;482-483: 400–10. doi:10.1016/j.scitotenv.2014.02.031
21. Kefalopoulos S, Paviotti M, Anfosso-Lédée F. *Common Noise Assessment Methods in Europe (CNOSSOS-EU)*. 2012.
22. Morley DW, de Hoogh K, Fecht D, Fabbri F, Bell M, Goodman PS, et al. International scale implementation of the CNOSSOS-EU road traffic noise prediction model for epidemiological studies. *Environ Pollut*. 2015;206: 332–341. doi:10.1016/j.envpol.2015.07.031
23. European Environment Agency. *Corine Land Cover 2000 Seamless Vector Data (V16)* [Internet]. 2015 [cited 1 Jul 2015]. Available: <http://www.eea.europa.eu/data-and-maps/data/corine-land-cover-2000-clc2000-seamless-vector-database-4>
24. Rothman KJ. Methodologic frontiers in environmental epidemiology. *Environ Health Perspect*. 1993;101 Suppl: 19–21.
25. Baliatsas C, Bolte J, Yzermans J, Kelfkens G, Hooiveld M, Lebret E, et al. Actual and perceived exposure to electromagnetic fields and non-specific physical symptoms: An epidemiological study based on self-reported data and electronic medical records. *Int J Hyg Environ Health*. 2015; doi:10.1016/j.ijheh.2015.02.001
26. Fyhri A, Klæboe R. Road traffic noise, sensitivity, annoyance and self-reported health—A structural equation model exercise. *Environ Int*. 2009;35: 91–97. doi:http://dx.doi.org/10.1016/j.envint.2008.08.006
27. Héritier H, Vienneau D, Frei P, Eze IC, Brink M, Probst-Hensch N, et al. The Association between Road Traffic Noise Exposure, Annoyance and Health-Related Quality of Life (HRQOL). *Int J Environ Res Public Health*. 2014;11: 12652–12667. doi:10.3390/ijerph111212652
28. Zijlema WL, Morley D, Stolk RP, Rosmalen JGM. Noise and somatic symptoms: A role for personality traits? *Int J Hyg Environ Health*. 2015; doi:10.1016/j.ijheh.2015.05.001
29. International Organization for Standardization (ISO). *Acoustics – Assessment of Noise Annoyance by Means of Social and Socio-acoustic Surveys; ISO/TS 15666*. Geneva, Switzerland; 2003.
30. Babisch W, Houthuijs D, Pershagen G, Cadum E, Katsouyanni K, Velonakis M, et al. Annoyance due to aircraft noise has increased over the years—results of the HYENA study. *Environ Int*. 2009;35: 1169–76. doi:10.1016/j.envint.2009.07.012

31. Porsius JT, Claassen L, Smid T, Woudenberg F, Timmermans DRM. Health responses to a new high-voltage power line route: design of a quasi-experimental prospective field study in the Netherlands. *BMC Public Health*. 2014;14: 237. doi:10.1186/1471-2458-14-237
32. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLoS Med*. 2015;12: e1001779. doi:10.1371/journal.pmed.1001779
33. Janson C, Chinn S, Jarvis D, Zock J-P, Torén K, Burney P. Effect of passive smoking on respiratory symptoms, bronchial responsiveness, lung function, and total serum IgE in the European Community Respiratory Health Survey: a cross-sectional study. *Lancet*. 2001;358: 2103–2109. doi:10.1016/S0140-6736(01)07214-2
34. GGD. Gezondheidsenquête. In: 2015 [Internet]. [cited 1 Sep 2015]. Available: <http://ggd.groningen.nl/onderzoek/gezondheidsprofiel>
35. Rijksoverheid. WoonOnderzoek Nederland (WoON). In: 2015 [Internet]. [cited 1 Sep 2015]. Available: <https://www.rijksoverheid.nl/onderwerpen/onderzoeken-over-bouwen-wonen-leefomgeving/inhoud/lopende-onderzoeken/woononderzoek-nederland-woon>
36. Leal C, Chaix B. The influence of geographic life environments on cardiometabolic risk factors: a systematic review, a methodological assessment and a research agenda. *Obes Rev*. 2011;12: 217–30. doi:10.1111/j.1467-789X.2010.00726.x
37. Moser A, Panczak R, Zwahlen M, Clough-Gorr KM, Spoerri A, Stuck AE, et al. What does your neighbourhood say about you? A study of life expectancy in 1.3 million Swiss neighbourhoods. *J Epidemiol Community Health*. 2014;68: 1125–32. doi:10.1136/jech-2014-204352
38. Statistics Netherlands. CBS - Kerncijfers wijken en buurten 2004-2013. In: Centraal Bureau voor de Statistiek, Den Haag/Heerlen [Internet]. 2013. Available: <http://www.cbs.nl/nl-NL/menu/themas/dossiers/nederland-regionaal/cijfers/incidenteel/maatwerk/wijk-buurt-statistieken/kwb-recent/default.htm>
39. Emaus MJ, Bakker MF, Beelen RMJ, Veldhuis WB, Peeters PHM, van Gils CH. Degree of urbanization and mammographic density in Dutch breast cancer screening participants: results from the EPIC-NL cohort. *Breast Cancer Res Treat*. 2014;148: 655–63. doi:10.1007/s10549-014-3205-2
40. Tigchelaar EF, Zhernakova A, Dekens JAM, Hermes G, Baranska A, Mujagic Z, et al. Cohort profile: LifeLines DEEP, a prospective, general population cohort study in the northern Netherlands: study design and baseline characteristics. *BMJ Open*. 2015;5: e006772. doi:10.1136/bmjopen-2014-006772
41. Maas J, Verheij RA, de Vries S, Spreuwenberg P, Schellevis FG, Groenewegen PP. Morbidity is related to a green living environment. *J Epidemiol Community Health*. 2009;63: 967–973. doi:10.1136/jech.2008.079038
42. Beelen R, Raaschou-Nielsen O, Stafoggia M, Andersen ZJ, Weinmayr G, Hoffmann B, et al. Effects of long-term exposure to air pollution on natural-cause mortality: an analysis of 22 European cohorts within the multicentre ESCAPE project. *Lancet*. 2014;383: 785–795. doi:10.1016/S0140-6736(13)62158-3
43. Brunekreef B. Exposure science, the exposome, and public health. *Environ Mol Mutagen*. 2013;54: 596–8. doi:10.1002/em.21767
44. Peters A, Hoek G, Katsouyanni K. Understanding the link between environmental exposures and health: does the exposome promise too much? *J Epidemiol Community Health*. 2011; doi:10.1136/jech-2011-200643

## Chapter 2

45. Chang ET, Adami H-O, Bailey WH, Boffetta P, Krieger RI, Moolgavkar SH, et al. Validity of geographically modeled environmental exposure estimates. *Crit Rev Toxicol.* 2014;44: 450–66. doi:10.3109/10408444.2014.902029
46. Bovenberg JA, Hansell A, de Hoogh K, Knoppers BM. Nature, nurture and exposure: Connecting biobank data with geographic data could yield public and individual health benefits, but risks to human rights need to be assessed. *EMBO Rep.* 2015;16: 404–6. doi:10.15252/embr.201439902
47. Schwartz J, Bellinger D, Glass T. Exploring potential sources of differential vulnerability and susceptibility in risk from environmental hazards to expand the scope of risk assessment. *Am J Public Health.* 2011;101 Suppl: S94–101. doi:10.2105/AJPH.2011.300272



