



Advancing Long-Term Outdoor Air Pollution Exposure Assessment for Health Studies: Lessons from a Multinational Research Initiative

Hanna Boogaard¹ · Dan L. Crouse¹ · Allison P. Patton¹ · Ellen Mantus¹ · Susanne Breitner-Busch^{2,3} · Jaime E. Hart^{4,5} · Anna Oudin^{6,7} · Shu Yang⁸ · Jeffrey D. Yanosky⁹ · John Volckens^{10,11} · Jana B. Milford¹²

Received: 17 March 2026 / Accepted: 24 August 2026
© The Author(s) 2026

Abstract

Purpose of review The Health Effects Institute recently completed an extensive research initiative to investigate novel approaches to improve assessment of long-term exposure to outdoor air pollution for health studies in North America and Europe. This overview synthesizes the key results, strengths, and limitations, along with other key research and future research needs.

Recent findings Exposures predicted from the various models, including those accounting for daily mobility patterns and outdoor-indoor infiltration rates, generally resulted in moderately to highly correlated exposure estimates (correlation > 0.7). Health effect estimates associated with the alternative exposures were similar, albeit with some notable differences in the magnitude – but not the direction — of the associations.

Summary This research initiative can inform decisions about estimating air pollution exposure in epidemiological studies, as investigators balance concerns about bias, efficiency, and resource needs. Future work is needed to replicate findings in other studies and regions. Other areas of future work include generating representative validation data for long-term exposure estimates, propagating exposure measurement error more fully into health effect estimates and improving understanding of the influence of indoor environments.

Keywords Air pollution · Long-term exposure · Exposure assessment · Health · Epidemiology

The editors would like to thank Douglas O. Johns for assisting with the review of this manuscript.

✉ Hanna Boogaard
jboogaard@healtheffects.org

¹ Health Effects Institute, Boston, MA, USA

² Institute of Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany

³ Institute for Medical Information Processing, Biometry, and Epidemiology, Faculty of Medicine, Ludwig-Maximilians-Universität München, Munich, Germany

⁴ Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA

⁵ Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, MA, USA

⁶ Division of Occupational and Environmental Medicine, Department of Laboratory Medicine, Lund University, Lund, Sweden

⁷ Department of Epidemiology and Global Health, Umeå University, Umeå, Sweden

⁸ Department of Statistics, North Carolina State University, Raleigh, NC, USA

⁹ Department of Public Health Sciences, College of Medicine, The Pennsylvania State University, Hershey, PA, USA

¹⁰ Department of Environmental and Radiological Health Sciences, Colorado State University, Fort Collins, CO, USA

¹¹ Department of Mechanical Engineering, Colorado State University, Fort Collins, CO, USA

¹² Department of Mechanical Engineering, University of Colorado, Boulder, CO, USA

Introduction

Air pollution is widely acknowledged as the most important environmental public health risk factor worldwide [1]. There is a broad scientific consensus that exposure to air pollution significantly increases mortality and morbidity from cardiovascular and respiratory diseases, diabetes, and lung cancer, and reduces life expectancy [2, 3, 4, 5].

The assessment of long-term exposure to ambient air pollution for epidemiological studies, however, remains challenging. Early cohort studies characterized exposure by assigning the average concentration measured at one or a few central sites within a city to each participant within that city [6, 7]. Fixed-site regulatory networks — even those in North America and Western Europe — continue to have relatively limited spatial coverage in many areas, particularly in suburban and rural locations, and insufficient density to capture small-scale (within-city) variations of air pollution [8].

Novel measurement technologies and modeling approaches for long-term exposure to air pollution are increasingly being used to estimate air pollution at finer spatial scales for epidemiological studies of large populations. Advances include novel air pollution sensors, mobile monitoring, satellite data, hybrid models, and machine learning approaches [9]. Even with those advances, important limitations and challenges remain when predicting long-term air pollution exposure, particularly for pollutants that vary widely in space and time, such as ultrafine particles (UFPs), black carbon (BC), and nitrogen dioxide (NO₂).

In 2019, the Health Effects Institute (HEI) issued a request for research applications titled *Applying Novel Approaches to Improve Long-Term Exposure Assessment of Outdoor Air Pollution for Health Studies* [10]. HEI solicited proposals for studies that would develop and apply scalable, novel approaches to improve assessment of long-term exposure to outdoor air pollutants, evaluate exposure measurement error, and determine how different exposure assessment approaches might influence epidemiological health effect estimates. Five studies were funded and have recently been completed. This paper highlights the key findings and lessons learned through this research program.

Methods

We provide an overview and synthesis of the results from studies conducted under the HEI initiative, drawing on evidence from other key research that addresses similar questions. We discuss strengths and limitations of the HEI studies and areas for future research. The evaluation of the five studies included here is based on an independent

review by the HEI Improved Exposure Assessment Studies Review Panel (referred to as “Review Panel”). The Review Panel members are co-authors of this paper. A key goal is to provide an overview of the extensive body of work produced by the funded investigators, make the findings readily accessible, and emphasize future research directions.

Summary of the Studies’ Approaches and Key Results

HEI published the final reports of the studies in 2024 and 2025 [11, 12, 13, 14, 15]. We provide a summary of the studies’ approaches, key results, and their interpretation (Fig. 1; Table 1). As shown in the table, air pollutants covered included pollutants that typically vary widely in space and time, such as UFPs, BC, and NO₂. Fine particulate matter (PM_{2.5}) was also included as a reference pollutant, although PM_{2.5} is typically more homogeneously distributed. HEI did not impose a common set of pollutants across studies. Our review highlights findings on UFPs and BC, because these were the most addressed pollutants in the initiative and are “emerging” pollutants of health concern according to the World Health Organization (WHO) [5]. The evidence base for adverse health effects of PM_{2.5} and NO₂ is much more established [2, 3, 4, 5].

Canadian Study Incorporating Mobile Monitoring Data

Weichenthal and colleagues estimated associations between long-term exposures to UFPs and BC and mortality in Toronto and Montreal, Canada, using several exposure modeling approaches [11]. They conducted mobile monitoring campaigns in both cities in 2020 and 2021, thus during the COVID-19 pandemic, covering various times of day from 7 a.m. to 11 p.m., weekdays and weekends, and all four seasons. In total, 14 and 20 routes were selected in Montreal and Toronto, respectively, that were sampled in a random order to obtain a representative annual average. They measured UFP number concentrations, UFP size, and BC in real-time using research-grade instruments, then calculated median values over 100-meter road segments and all sampling days. In total, mobile monitoring data were aggregated to 5,819 and 7,051 road segments in Toronto and Montreal, respectively. On average, road segments were visited on 10 different days. The mobile monitoring data were randomly split into subsets to train (70%), validate (15%), and test (15%) the subsequent exposure models.

The investigators developed three new exposure models for each city separately: (a) land-use regression models based on the mobile monitoring data combined with

HEI's Research Initiative to Improve Long-term Exposure Assessment of Outdoor Air Pollution for Health Studies

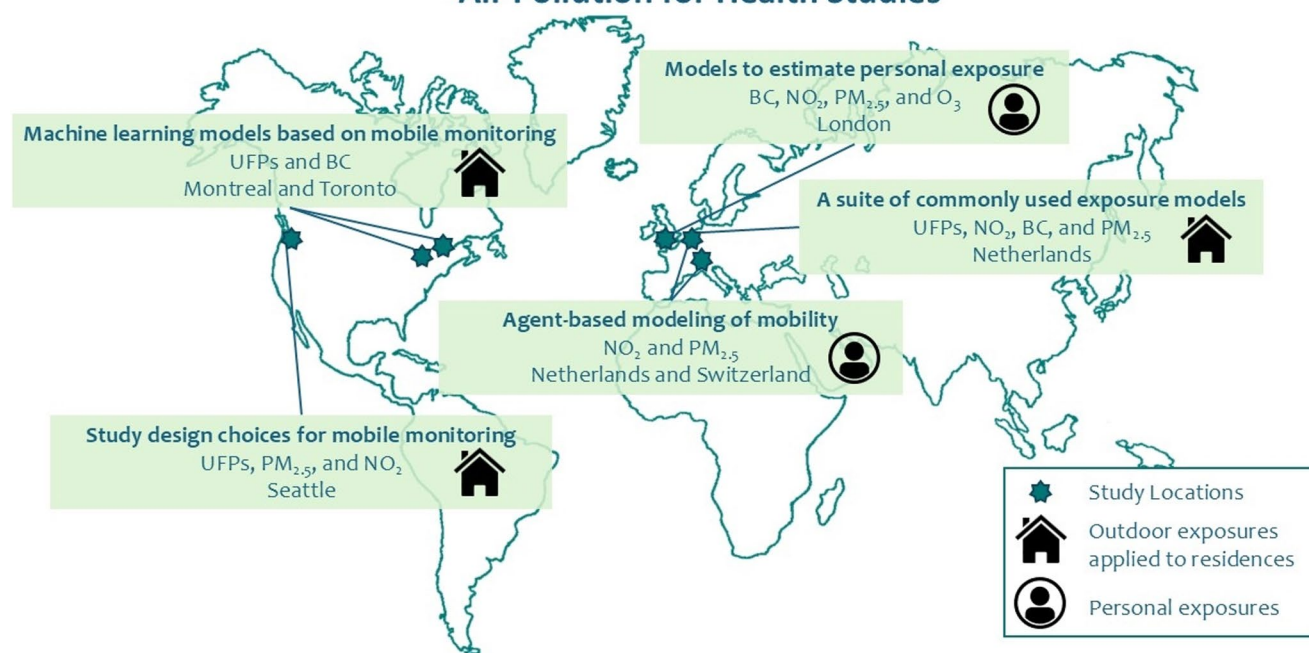


Fig. 1 Geographical Location and Key Features of HEI's Initiative to Improve Long-term Exposure Assessment of Outdoor Air Pollution for Health Studies

detailed meteorological, land use, and traffic information; (b) machine learning models, specifically convolutional neural network models using mobile monitoring data and aerial images from Google Maps; and (c) a combination of these two models. The models accounted for weather-related variations in air pollution during the monitoring campaign despite the random order of the measurement routes. Estimates from each model were developed using the training dataset and compared to observed values in the validation and test datasets. Historical trends in traffic information and nitrogen oxide emissions were used to project model predictions back in time (i.e., backcast).

The investigators then applied those models to a large representative sample of Canadian adults (1.5 million) from the Canadian Census Health and Environment Cohort residing in Toronto or Montreal. The study population included adults aged 25 years and older from multiple Census years (1991–2006), with mortality follow-up from 2001 to 2016. Exposures were assigned to participants based on home postal code (about the size of a city block) and accounted for address changes over time. The main analyses used exposure estimates developed with backcasting. The investigators used both single and multipollutant Cox proportional hazard models to assess the associations between exposures to UFP number concentrations and BC, and nonaccidental and cause-specific mortality. They adjusted the health analyses for important confounders, such as sociodemographic

factors and co-pollutants, including fine particulate matter ($PM_{2.5}$) and UFP size.

The land-use regression models performed better than the machine learning models, and the exposure models that combined land use and machine learning model predictions performed only slightly better than models that used land use regression alone. The combined model explained half or more (R^2 between 0.49 and 0.73) of the observed spatial variation in UFPs and BC and was considered the main exposure model in the health analyses. Long-term exposures to UFPs and BC were positively associated with mortality in single-pollutant models. The effect estimates were sensitive to adjustment for co-pollutants, including UFP size. Associations between UFP number concentrations and mortality increased after adjusting for UFP size, whereas associations between BC and mortality became generally weaker or null when UFP size was included. Generally, similar findings were reported for UFPs and BC across various alternative exposure assessment approaches, including without backcasting, albeit the association's magnitude — but not its direction — differed sometimes substantially.

United States Study Incorporating Mobile Monitoring Data

Sheppard and colleagues evaluated different approaches to mobile monitoring study design in Seattle, Washington,

Table 1 Key Characteristics of HEI's Initiative to Improve Long-term Exposure Assessment of Outdoor Air Pollution for Health Studies

Principal Investigator	Cohorts	Location	Study Period	Study Population	Sample Size	Outcomes	Air Pollutants	Exposure Assessment	Air Pollutant Instruments	Monitoring Data
Weichenthal [11]	CanCHEC	Montreal and Toronto, Canada	1991–2016	Adults	1.5 million	Mortality	UFPs and BC	Machine learning and land-use regression models	Naneos Partector 2, DiSCmini and MicroAeth MA350	Mobile
Sheppard [12]	ACT Air pollution study	Seattle, United States	1994–2020	Older adults	5,400	Cognitive function	UFPs, PM _{2.5} , and NO ₂	Universal kriging, spatiotemporal, and machine learning models	NanoScan, P-Track, DiSCmini, Plan-tower sensor PMS A003, Alphasense sensor B43F, and others	Mobile, outdoor low-cost sensors, regulatory, and research-grade monitors
de Hoogh [13]	EPIC-NL, SAPALDIA, SNC	Netherlands, Switzerland	1991–2018	Adults	3.5 million	Mortality, stroke, coronary events, lung function, and blood pressure	NO ₂ and PM _{2.5}	Agent-based, land-use regression, and machine learning models	Reference monitors	Regulatory monitors, location tracking
Hoek [14]	DUELS, EPIC-NL, PIAMA	Netherlands	1993–2019	Children and adults	10.8 million	Mortality, stroke, coronary events, lung function, and asthma	UFPs, NO ₂ , BC, and PM _{2.5}	Land-use regression, dispersion, machine learning, and hybrid models	CPC3007, AE51, Dust-Trak, Sense-ion SPS30, Alphasense sensor B43F, Harvard Impactors, Ogawa, and others	Mobile, outdoor low-cost sensors, regulatory, and research-grade monitors
Katsouyanni [15]	BLW, COPE, DEMiSt, PASTA, London segment of UK Biobank	London, UK	2006–2024	Adults	62,000 ^a	Mortality, asthma, COPD, myocardial infarction, stroke, and dementia	BC, NO ₂ , PM _{2.5} , and O ₃	Land-use regression, dispersion, machine learning, and hybrid models	MicroAeth AE51, MicroAeth MA300/350, Alphasense OPC-N2, Senserion SPS30, A4 sensors	Personal measurements, regulatory monitors

^aRefers to the London segment of UK Biobank

ACT=Adult Changes in Thought; BLW=Breathe London Wearables; CanCHEC=Canadian Census Health and Environment Cohort; COPD=Chronic Obstructive Pulmonary Disease; COPE=Characterisation of COPD Exacerbations using Environmental Exposure Modelling; DEMiSt=Driver Diesel Exposure Mitigation Study; DUELS=Dutch Environmental Longitudinal Study; EPIC-NL=European Prospective Investigation on Cancer and Nutrition–Netherlands; PIAMA=Prevention and Incidence of Asthma and Mite Allergy; PASTA=Physical Activity through Sustainable Transport Approaches; SAPALDIA=Swiss Study on Air Pollution and Lung Disease in Adults; SNC=Swiss National Cohort

United States [12]. They leveraged detailed air pollution data, including a mobile monitoring campaign of UFPs, and cognitive function data from about 5,000 participants in the Adult Changes in Thought (ACT) Air Pollution study — a cohort study of older adults.

Most analyses focused on UFP data from a previous mobile roadside monitoring data campaign conducted in 2019–2020, thus before the COVID-19 pandemic. UFP were measured using three different instruments: a P-Trak, a NanoScan, and a DiSCmini. The campaign covered various

times of day from 5 a.m. to 11 p.m., weekdays and weekends, and all four seasons using a temporally balanced sampling design. In that campaign, short-term measurements were made from a parked vehicle at 309 roadside locations, with about 30 visits made to each site. These sites were intended to represent cohort participants' residential locations, which were on average 611 m away from their closest roadside measurement site. The investigators also conducted analyses using on-road measurements of UFPs between roadside sites along nine predefined routes for a total of 5,878 road

segments. Each 100-meter segment was visited an average of 28 times. The Nanoscan was used as the primary instrument for most analyses of mobile roadside data, whereas the P-Trak was used for analyses of mobile on-road data.

The investigators subsampled UFP measurement data to evaluate various exposure assessment study design features, including fewer visits per site, fewer days of the week, restricted hours of the day, and fewer seasons. The investigators used either the full dataset or subsets of measurements to develop annual average exposure estimates using a suite of models. The main models were developed using universal kriging with dimension reduction using partial least squares (for brevity: universal kriging models). The universal kriging approach was compared to spatial random forest and other machine learning algorithms. A suite of geographical covariates (about 200) was available for inclusion in the models, including indicators of land use, roadway proximity, and population density. They assessed the performance of each model using cross-validation (mobile monitoring data) and calculated several measures, including the mean squared error-based explained variance (R^2 for simplicity). The investigators also explored the tradeoffs between universal kriging exposure model performance and logistical features (both cost and time) to identify optimal monitoring designs.

Each model was used to predict the 5-year average UFP exposures prior to the cognitive function measurement obtained at baseline (1994 or later). The investigators applied standard linear regression to assess the association between exposure to ultrafine particles from each exposure model and baseline cognitive function, adjusted for participant age, sex, education, and calendar year (confounder model 1). Some results were also adjusted for participant race and socioeconomic status (confounder model 2). In the comparisons, the models using all UFP data from the mobile roadside campaign were taken as the reference model. The investigators quantified the exposure measurement error using their earlier developed bootstrap methods [16, 17, 18] and corrected the health effect estimates accordingly in the reference model.

The universal kriging reference model using all ultrafine particle data from the mobile roadside campaign had a cross-validated R^2 of 0.65 (for NanoScan measurements) and 0.77 (for P-Trak measurements). The universal kriging model typically performed slightly better than the various machine learning models.

Universal kriging models with restricted mobile roadside sampling of UFPs consistently produced lower-performing exposure models than the reference model developed using the full dataset. The investigators found that a mobile monitoring study with roadside UFP sampling of at least 12 visits per location optimized exposure model performance while

limiting costs. Furthermore, the investigators noted that exposure sampling in mobile monitoring campaigns should cover all days of the week, most hours of the day (including early morning and late evening), and at least two seasons.

Exposure to UFPs estimated using the reference model was negatively (adversely) associated with cognitive function at baseline after adjustment for participant age, sex, education, and calendar year (confounder model 1). The reduced sampling designs yielded similar findings regarding negative (adverse) associations between UFP exposure and cognitive function at baseline, albeit with sometimes notable differences in association magnitude (up to 60%). In the reference exposure model for UFP, the investigators reported that the observed associations using confounder model 1 were more strongly affected by the features of the mobile monitoring design than by exposure measurement error (up to a 6% increase in association magnitude and a 13% increase in width of the confidence interval). The observed negative association in the reference model disappeared in health models that further adjusted for race and socioeconomic status (confounder model 2), hampering further assessments of the influence of sampling design on health estimates using different exposure estimates for those health models.

Using mobile on-road data for UFPs instead of roadside data, the investigators found that most comparisons identified the same design features or elements as important, although with a few notable differences. Strikingly, on-road modeling results were generally similar when road segments were tested 4 times versus 12 times, although the latter produced more stable results.

A small part of the study evaluated low-cost sensors for $PM_{2.5}$ and NO_2 and is summarized in Supplementary Materials.

European Study Accounting for People's Mobility

De Hoogh and colleagues assessed whether exposure estimates that accounted for people's mobility would improve exposure assessment for NO_2 and $PM_{2.5}$ and potentially reduce exposure measurement error in three cohort studies in Switzerland and the Netherlands [13]. The study addressed an important outstanding issue, as most health studies assess long-term air pollution exposure using outdoor concentrations estimated at participants' residential locations only.

De Hoogh and colleagues simulated participants' mobility patterns using agent-based modeling – a computer simulation of actions and interactions among, for example, people, places, built environment, transportation, and time that allows for the integration of individual and population behaviors [19]. Their simulations were informed by existing

travel survey and census data from 2010 to 2019 in Switzerland and the Netherlands. The investigators developed 13 activity profiles to represent different time–activity patterns (e.g., commuting) using characteristics available (e.g., employment status) from the health studies described below. Except for a subset of participants in one cohort, workplace locations were unknown and were estimated. They repeated the simulations between 50 and 1,000 times and evaluated the model using GPS location data collected over 2 weeks in tracking campaigns that included nearly 700 participants in the two countries in 2022 (after the COVID-19 pandemic).

The investigators developed nationwide hourly concentration maps of outdoor NO₂ and PM_{2.5} and linked those to the agent-based model output (i.e., individual time–activity data) to estimate exposure. In Switzerland, the investigators used annual average concentration maps for 2016, derived from previously published spatiotemporal models [20, 21], and applied a temporal adjustment approach to obtain hourly estimates [22]. In the Netherlands, hourly land-use regression models were developed specifically for this study using regulatory monitoring data from 2016 to 2019, stratified by season and weekday type (weekday or weekend day). The investigators developed and compared supervised linear regression models and random forest models in the Netherlands. They assessed model performance in terms of explained variance (R^2) using fivefold cross-validation. The newly developed hourly land-use regression models in the Netherlands performed slightly better for NO₂ (cross-validation R^2 =0.6–0.7) than for PM_{2.5} (cross-validation R^2 =0.4–0.5). The random forest models performed slightly (up to 4%) better than the supervised linear regression models for both pollutants and were subsequently used to combine outdoor concentrations with time–activity data in the Dutch health studies.

The investigators then conducted epidemiological analyses using three adult population-based cohort studies: the Swiss National Cohort (SNC), the European Prospective Investigation into Cancer, Netherlands (EPIC-NL), and the Swiss Cohort Study on Air Pollution and Lung and Heart Diseases in Adults (SAPALDIA). They compared health effects estimates based on residential versus mobility-enhanced exposure estimates. For all cohort participants, the annual average residential exposure was assigned at the baseline home address (1993–2011, depending on the cohort). The mobility-enhanced exposure was assigned based on the most appropriate activity profile for each participant, taking into account their sex, employment status, and socioeconomic status. The investigators used single-pollutant Cox proportional hazards models for mortality and morbidity outcomes in the SNC and EPIC-NL analyses and repeated cross-sectional linear regression analyses for lung function and blood pressure outcomes in the SAPALDIA

analyses. All results were adjusted for age, sex, and individual and area-level socioeconomic status. Results from the analyses of EPIC-NL and SAPALDIA were further adjusted for various lifestyle factors.

The mobility-enhanced exposure estimates were highly correlated (Pearson correlation > 0.95) with residential-only estimates for NO₂ and PM_{2.5} among the cohort participants. In all cohorts, mean air pollution exposures and contrasts (interquartile range) in individuals' exposures were slightly lower for mobility-enhanced exposures than for residential exposures. For both exposures, the contrast was higher for NO₂ than for PM_{2.5}, consistent with the larger spatial variation of NO₂ [23].

The residential and mobility-enhanced exposure estimates showed similar associations with various health outcomes within the cohorts. Hence, the results suggest that the bias in epidemiological studies based on outdoor concentrations at residential locations might be small and that accounting for mobility might not be important. Positive (adverse) associations were observed most clearly in the SNC. The investigators found that exposure to NO₂ and PM_{2.5} was associated with an increased risk of nonaccidental mortality and that exposure to NO₂, but not exposure to PM_{2.5}, was associated with an increased risk of cardiovascular mortality. They reported generally non-statistically significant and, in some cases, negative associations for the health outcomes in the EPIC-NL and SAPALDIA cohorts. No consistent difference in the association was found between exposures using known versus estimated workplace locations in the subset of SAPALDIA participants for whom workplace locations were available.

European Study Comparing a Suite of Exposure Models

Hoek and colleagues compared the performance of several long-term exposure assessment methods for UFPs, BC, NO₂ and PM_{2.5} in the Netherlands [14]. They used existing data and models from previous collaborative projects where available, including the European Study of Cohorts for Air Pollution Effects (ESCAPE), Effects of Low-Level Air Pollution: A Study in Europe (ELAPSE), and EXposome Powered tools for healthy living in urbAN Settings (EXPANSE). The investigators evaluated annual average outdoor air pollution concentrations using a suite of exposure assessment methods that differ in their underlying monitoring data (low-cost sensors, mobile monitoring, and fixed-site routine or fixed-site study-specific monitoring using research-grade instruments) and modeling approaches (land-use regression models or air pollution dispersion models). The comparison included exposure models derived from four previous mobile monitoring campaigns with mobile and/or

road-side sampling and coverage of different time periods. Furthermore, to develop the land-use regression models, the investigators tested supervised linear regression and two machine learning methods (random forest and Least Absolute Shrinkage and Selection Operator). Low-cost sensor data were used primarily in hybrid models, where they were combined with dispersion model estimates and data from the routine national monitoring network to assess the added value of low-cost sensor data (i.e., Alphasense sensor B43F) for NO₂.

The predictions of the exposure models were compared at 20,000 random residential addresses in the Netherlands and tested on existing and new validation data spanning a 20-year period. The new validation dataset entailed 1-week measurements in two different seasons at 31 residential outdoor locations across the Netherlands from 2021 to 2023. Data from a reference site in the middle of the country were used to account for temporal differences between measurement weeks.

They applied the various exposure models to three major cohorts in the Netherlands — an administrative cohort of 10.8 million adults (DUELS), the European Prospective Investigation into Cancer and Nutrition Netherlands (EPIC-NL), and the Prevention and Incidence of Asthma and Mite Allergy (PIAMA) birth cohort — to evaluate how these exposure models influence health effect estimates in epidemiological studies.

For all cohort participants, the various exposure estimates were assigned to the residential address at the time of cohort recruitment. The investigators used single-pollutant Cox proportional hazard models for mortality and morbidity outcomes in DUELS and EPIC-NL, and linear regression and discrete-time hazard models for asthma incidence and lung function in PIAMA. All results were adjusted for age, sex, and individual and area-level socioeconomic status. EPIC-NL and PIAMA results were further adjusted for various lifestyle factors. The investigators expressed the effect estimates using fixed increments across methods and method-specific interquartile ranges of exposure. In further exploratory analyses, the investigators examined whether and why effect estimates differed across exposure assessment methods.

The various exposure models generally resulted in moderately to highly correlated exposure predictions for BC, NO₂, and UFPs at residential sites across the Netherlands (Pearson correlation > 0.7). In some cases, predicted exposure levels and exposure contrasts varied widely across methods, with the mobile monitoring models predicting the higher levels and contrasts, probably related to the on-road monitoring. For PM_{2.5}, the correlations among the models were generally lower (Pearson correlation < 0.4) and exhibited low exposure contrast. PM_{2.5} models based on mobile

monitoring demonstrated poor internal model performance. In most comparisons, the investigators observed only small differences among the three land-use regression algorithms. Exposure models explained a low-to-moderate amount of spatial variance in UFP and NO₂ measurements at the new validation sites, and a positive bias (overestimation) was reported across all models. Most models explained a moderate amount of spatial variance in past exposures at existing validation sites for BC and NO₂ and explained less variance for UFPs. Most models predicted past exposures of PM_{2.5} poorly, possibly due to small variability. The addition of low-cost sensor data did not improve NO₂ estimates in models that combined dispersion model estimates and data from the routine national monitoring network.

The application of various exposure assessment approaches in health studies yielded similar findings regarding associations with mortality and morbidity outcomes, albeit with sometimes notable differences in association magnitude — but not the direction. Positive (adverse) associations were observed most clearly in DUELS with positive associations found across models for all four pollutants for nonaccidental mortality and to a lesser extent in PIAMA, which found statistically significant positive associations with asthma incidence and reduced lung function in most models. Null or sometimes even negative associations with nonaccidental mortality and incidence of coronary events and stroke were reported in the EPIC-NL cohort. There were no consistent differences in associations between exposure assessment approaches based on mobile monitoring, fixed-site monitoring, or dispersion models. Factors that did explain some of the heterogeneity of the associations included the performance of the model at external validation sites and the predicted exposure contrast. The year of the exposure model did not explain the heterogeneity, and the investigators documented stable spatial contrasts over a 10-year period in a subset of models with multiyear predictions available for BC, NO₂, and PM_{2.5}.

European Study Incorporating Personal Exposure Data

Katsouyanni and colleagues evaluated whether estimates of long-term personal exposure to outdoor air pollution in London, United Kingdom, yield more accurate and precise estimates of the health effects than alternative approaches [15]. They leveraged personal exposure data from four earlier studies carried out between 2015 and 2019. These studies were conducted during different periods, covered different pollutants, used different instruments for a given pollutant, and featured participants of various ages including children and older adults (Supplementary Materials Table S1). The sample sizes ranged from 38 to 164 participants. The

duration of monitoring per participant ranged from an average of 4.5 days to 144 days.

The investigators used those data to estimate annual exposures to outdoor-generated PM_{2.5}, NO₂, ozone (O₃) and BC for participants in each study by estimating infiltration of outdoor pollution to indoor environments and accounting for daily mobility patterns. In the absence of paired indoor and outdoor measurements, infiltration efficiency was estimated using a dynamic mass balance model.

For comparison, they assigned alternative estimates of exposures to the participants using existing exposure models and datasets, namely, measurements at nearby government air quality monitoring sites, Spatio-Temporal Exposure Assessment Methods (STEAM) models of outdoor pollution levels that combined land-use regression models and air pollution dispersion estimates (and satellite data for PM_{2.5} only), and models that combined ambient concentration estimates with typical time-activity patterns and pollutant concentration adjustments for indoor and vehicle microenvironments (London Hybrid Exposure Model, LHEM). The investigators then estimated the magnitude of measurement errors between the exposure estimates derived from the personal exposure measurement studies (which they call “true” exposures) and the various alternative exposure estimates. In practice, however, personal exposure to air pollution strictly from outdoor sources cannot be observed directly and, thus, “true” exposure is challenging to estimate.

To explore potential biases, Katsouyanni and colleagues conducted simulation studies to assess the effect of estimated exposure measurement error on mortality risk estimates. They used the data and information obtained from the analyses described to derive and apply two advanced statistical correction methods: simulation extrapolation and regression calibration. To inform the measurement error correction, they used personal exposure estimates derived using data from the COPE study on older adults with chronic obstructive pulmonary disease (COPD), which had the longest duration of measurements. Finally, they examined associations between the alternative exposure estimates and various mortality and morbidity outcomes in about 62,000 participants from the UK Biobank cohort who lived in London. The exposure estimates were assigned at baseline (2006–2010), and participants were followed up until 2024. They used Cox proportional hazards models and adjusted for various confounders, including smoking and socioeconomic status, and measurement error.

The investigators reported that exposure to outdoor-generated air pollution accounted for a large proportion of participants’ total personal exposure in the four personal air pollution exposure measurement studies (ranging from 41 to 85%). They reported that the annual mean exposure estimates assigned by approaches relying on ambient

concentrations were much higher than those estimated from personal exposure measurement studies. Spearman correlations between the annual mean exposure estimates from the personal measurement-based studies and the alternative exposure estimates were generally low ranging from −0.19 to 0.21 across pollutants and personal exposure studies.

Across the simulation studies, the resulting biases in health effect estimates from using alternative exposure estimates in place of the “true” estimates from the personal exposure measurement studies were consistently large and consistently in a direction toward smaller estimated health effects. In analyses using the UK Biobank cohort, the investigators found little convincing evidence of associations between health outcomes and any of the estimated exposures in uncorrected models. Many of the hazard ratios approached 1.00, and only very few achieved statistical significance. Notably, however, effect estimates increased by two to four times in magnitude in many cases (but remained statistically nonsignificant) when adjusted for measurement error using either correction method.

Discussion

The simultaneous funding and collaboration among the investigators fostered synergies among the teams and facilitated methodological developments to improve exposure assessment. A particularly strong aspect of the studies was the leverage of a wealth of air pollution and health data and where needed, extensive new data collection. The extensive year-long mobile monitoring campaigns in two of the studies were an advance over prior work, covering multiple times of day, including early morning and late evening hours, weekdays, and weekends — thus including those times of day when people might be more likely to be at home — and all four seasons. Many other mobile monitoring campaigns have collected less data, sampled during more restricted periods, such as business hours only, or had short monitoring durations lasting only a few weeks or months [24].

The extensive modeling of air pollution exposure and rigorous evaluation of model performance are further strengths of the studies. The comparison of a large suite of exposure models commonly used in epidemiological studies was notable, making the results relevant and widely applicable. The studies included a variety of exposure models such as land-use regression models, universal kriging models, machine learning models, dispersion models, and hybrid models that combined monitoring data, land use data, satellite observations, and estimates from chemical transport models. The large number of geographical

covariates available for inclusion in the models was also notable. Further, the investigators reported several measures of model performance to test accuracy and possible bias — thereby providing an in-depth performance assessment.

The application of various exposure models to health outcomes was another strength. The investigators applied exposure estimates to diverse population-based cohorts in North America and Europe. These cohorts differed in size, population, location, health outcomes, and study period. In particular, the health analyses using very large administrative cohorts (up to 10.8 million in DUELS, essentially all Dutch adult citizens aged 30 years or older) formed by linking census data, population registries, and death registries were considered informative. The inclusion of both mortality and morbidity outcomes and the availability of lifestyle factors (e.g., smoking and diet) in smaller cohorts (e.g., SAPALDIA, EPIC-NL, PIAMA), were also noted as strengths.

The studies addressed critical research gaps in understanding the health effects of long-term exposure to UFPs and BC, for which the evidence base is still limited, partly due to the lack of comprehensive routine monitoring. Dozens of peer-reviewed papers have been published by the study teams, many in high-impact journals such as *The American Journal of Respiratory and Critical Care Medicine* [25] and *Environment International* [26]. The Sheppard study alone produced at least 13 papers, as documented in the report's annotated bibliography [12].

Across the studies, common limitations included the short duration and limited number of samples available for validation, and the temporal misalignment between health outcome data versus recently acquired air pollution measurements for exposure modeling. These and other aspects of the study designs, approaches, and interpretations of the findings are discussed in the following sections.

Challenges With Mobile Monitoring

The Weichenthal study featured an extensive, year-long, mobile monitoring campaign, even so only 100-m road segments were, on average, visited on 10 different days with about 6 s of observation per visit. That equates to a total sampling duration of about 60 s per year at each road segment. Longer monitoring times might have provided more stable estimates of annual average UFP and BC levels but would likely be cost-prohibitive. On the other hand, in a detailed comparison using Google Street View cars in Oakland, California, it was documented that only four to eight repeat visits per 30-m road segment produced robust long-term nitrogen oxide and BC exposure models [27, 28]. Note that UFPs were not measured in the Oakland study. Similar

evaluation studies reported the need for at least 12 visits for stable UFP models [29], in line with the findings from the Sheppard study [12]. How many times a street segment needs to be measured for use in health studies depends on the location including the temporal variability of local sources of air pollution, among other factors [30].

The investigators in the Sheppard study further emphasized the importance of including samples beyond business hours, including extended times of the day and weekends [12]. Most campaigns have sampled during typical business hours because a technician is required to operate the monitoring platform, potentially introducing bias in characterizing the long-term average [24, 30]. Opportunistic mobile monitoring, for example, on public transportation might be promising to extend this time window to weekend days, holidays or night hours [30]. Notably, the study by Hoek and colleagues did not find differences in effect estimates when comparing models based on weekday mobile monitoring between the hours of 9 a.m. and 4 p.m. versus monitoring between 8 a.m. and 10 p.m. [14]. However, Hoek did not have access to a mobile monitoring study that covered 24 h a day and lacked weekend days, limiting the comparison. A limitation across all studies is that the studies did not capture short-term high air pollution exposure because exposures were averaged over one or more years. It is unclear how individual or repeated short-term high-intensity air pollution exposure events influence long-term exposure and health estimates.

A fundamental challenge with mobile monitoring is that it is difficult to separate temporal and spatial variation, as the platform moves in both time and space. To enable better comparison of mobile measurements sampled at different times of day, some mobile monitoring studies have used fixed-site monitor data for the temporal adjustment, although questions remain about whether one or a few fixed-site monitors can sufficiently represent UFP and BC temporal patterns over space [24]. The absence of fixed-site monitor data in both new mobile monitoring studies (Sheppard and Weichenthal) prevented such a temporal adjustment. Instead, the Weichenthal study relied on meteorological data for the temporal adjustment, and the Sheppard study explored the use of simulated UFP data at an urban background site. Moreover, the absence of fixed-site monitor data in both mobile monitoring studies (Sheppard and Weichenthal) hampered an evaluation of how well on-road measurements represent outdoor concentrations at off-road residential locations. Typically, on-road measurements are higher than the off-road air pollution values immediately outside residences in the range of 20 to 30% between on-road and off-road measurements, but the amount of overestimation depends on the setting and pollutant [30].

Mobile monitoring studies need to pay particular attention to extreme values and remove the influence of nearby polluting vehicles, while taking precautions to avoid self-contamination. Hence a rigorous process is needed to reduce the influence of potential extreme values and increase data evenness. The Weichenthal mobile monitoring campaigns positioned the sampling inlet to minimize self-contamination and computed median values to avoid outliers. The Sheppard study took three steps to reduce the influence of possible extreme values: they calculated medians for each visit (or segment), winsorized across all visits for each site (which replaces values at the tails of the distribution with a fixed percentile value), and then log-transformed the data for subsequent exposure modeling. A sensitivity analysis investigating how removing possible extreme values affected the subsequent exposure models was included in a paper resulting from this work [31].

Limitations in the Validation Data

Validation of long-term exposure models involves spatial comparison between modeled and measured concentrations. Temporal comparisons using a single or a few routine monitoring sites cannot inform how well a model performs spatially. Hoek and Katsouyanni's studies leveraged numerous datasets and conducted particularly extensive validation efforts but nevertheless faced some limitations. The utility of new validation data collected in the Hoek study was limited by the short duration of the measurements, non-simultaneous sampling necessitated by limited equipment, and the small number of repeated measurements for different seasons. Some existing validation datasets used in the Hoek study were hampered by similar issues and compounded by differences in equipment and monitoring strategies. Similar limitations were noted in the Katsouyanni study, as several of the personal exposure datasets they used included measurements collected over only a few days. Those issues limited the ability to reliably extrapolate the measurements to annual mean exposures and then compare them with predictions from annual average air pollution exposure models.

Indeed, in the Hoek study, large temporal variation was evident for all four pollutants in the new measurement data, even after accounting for temporal differences between measurement weeks using a reference site. Moreover, the investigators did not interpret the data for BC and PM_{2.5} because of the low spatial correlation between the first and second week of measurements. The comparisons using existing external validation data were limited by similar temporal coverage issues when testing the first Europe-wide UFP model [32]. There is a clear need for more research and investment in validation datasets for long-term exposure studies.

Long-term Temporal Stability of UFP Spatial Contrasts

A key challenge, especially for UFP, is predicting historical exposure contrasts relevant to epidemiological studies, which is difficult due to the lack of historical time-series measurements across many locations. The temporal misalignment between health outcome data versus recently acquired UFP data for exposure modeling was apparent in the approaches using mobile monitoring and low-cost sensor data, both approaches made possible by new measurement technologies. In Sheppard's application to the ACT study, the UFP models used were based on mobile measurements that were conducted up to 25 years after the end of the health outcome follow-up. In the Weichenthal and Hoek studies, shorter mismatches were noted for the models based on mobile monitoring, up to 5 years.

The Weichenthal study with CanCHEC applied a backcasting procedure using supplemental nitrogen oxide emissions and traffic data to partially address the lack of UFP and BC data for earlier years [11]. This approach represents an advance over most other studies. However, because data were lacking to evaluate the backcast surfaces, this procedure could introduce uncertainty that could affect the exposure and health estimates in unpredictable ways, depending on the quality of the data and modeling techniques used and how well nitrogen oxide emissions and traffic volume correlate with UFPs and BC.

Without a backcasting procedure, the assumption is that air pollution surfaces remain constant over time. This assumption is questionable because air pollution concentrations have been declining over the past few decades in many high-income countries, due largely to successful air quality regulation and subsequent reductions in emissions from major air pollution sources, including transportation and power generation [2, 3, 4, 5, 33].

This temporal misalignment between the period captured by the mobile measurements and the exposure window most relevant for epidemiological purposes is also apparent in some other cohort studies investigating UFPs [34, 35, 36, 37, 38]. The temporal mismatch in other studies is typically up to 10 years after the end of follow-up. The ability of the UFP model to assign historical exposure was evaluated in Amsterdam, the Netherlands, where an R^2 value of 0.36 was reported when using models to predict measurements collected about 10 years earlier [39]. A similar evaluation using various historical datasets was conducted for the Europe-wide UFP model, which was based on mobile monitoring data from 2018 to 2024 data. R^2 values ranged from 0.25 to 0.70 when applying the model to predict measurements collected between 3 and 20 years earlier, with no apparent temporal transferability

pattern [32]. For example, the lowest R^2 values were reported for the model to predict measurements collected 20 or 4 years earlier in Amsterdam or Rome, respectively, and the highest R^2 value was observed for data collected 6 years ago in Switzerland [32]. UFP exposure models appear to have moderate temporal transferability across a 5 to 10-year time span, but this needs verification in other settings.

Impact of the COVID-19 Pandemic

Several setbacks, including some related to the COVID-19 pandemic, led to substantial changes to the originally intended study scopes. The impact was particularly apparent in the De Hoogh study, which had to postpone tracking campaigns until COVID-19 restrictions were lifted so that eventually only a small convenience sample was obtained. Hence, the investigators decided to use the tracking data only to evaluate the agent-based modeling, rather than to inform the model directly. The evaluation of the agent-based modeling using tracking data was inherently limited because the convenience sample was not as representative as intended of the general population, cohort participants, and pre-pandemic activity patterns. In the Hoek study, the collection of new monitoring data to validate the models had to be postponed, resulting in fewer measurements than originally planned. In the Weichenthal study, the mobile monitoring campaign was conducted during the pandemic, thus caution is warranted in generalizing those findings, and the originally planned extensive fixed-site monitoring for model development or evaluation was not achieved. In the Katsouyanni study, the likely atypical patterns of exposure that might have occurred during the period of the COVID-19 pandemic in the application to the UK Biobank cohort were ignored. The findings in the Sheppard study were not affected by the COVID-19 pandemic in 2020 because monitoring was completed before that time and only historical health data were used.

The COVID-19 pandemic led to unprecedented restrictions that dramatically reduced global and local travel, shut down schools and businesses, and halted some industrial activity. The decrease in transportation and non-essential business activities during the COVID-19 lockdowns temporarily led to a substantial reduction in air pollution in many countries, particularly NO_2 , and, to a lesser degree, $\text{PM}_{2.5}$ [40, 41]. As has now been established, these changes were only temporary. As restrictions have lifted, emissions have risen, erasing improvements in air quality. Nonetheless, the profound impact of human activity levels on air pollution underscored what is possible and inspired renewed calls for cleaner air in the long term.

Use of Real-world Data Versus Simulations

The HEI-funded studies used either real-world health outcome data or a combination of real-world data with simulations to evaluate how the different exposure estimates influence health effect estimates. An advantage of using simulated data is that the underlying “true” health effects are known in that scenario, allowing one to test a single feature (e.g., exposure measurement error) systematically while holding all other conditions constant. The challenge with simulation studies is that they might not adequately represent the real world, which is complex, and that many different scenarios might be necessary. Both approaches have their pros and cons; the use of both approaches in the Katsouyanni study was considered a strength.

In Sheppard’s application to the ACT study, the null findings from the reference exposure model using confounder model 2 hampered the assessment of the influence of sampling design on health estimates using different exposure estimates. Some carefully designed simulations would have been informative to evaluate how the different exposure estimates influence health effect estimates. Likewise, in the Katsouyanni study, the largely null findings in the analysis of the UK Biobank cohort uncorrected for measurement error hampered the evaluation.

Lessons Learned and Remaining Areas for Future Research

HEI seeks to embark on the next stage of innovative and policy-relevant science, integrating valuable lessons learned from this research initiative into the new research. We describe below several areas of potential interest. We note that our suggestions for future research are largely focused on high-income countries and that research needs in low-and-middle income countries are generally quite different [42].

UFPs and BC: Emerging Evidence for Health Effects and Regulations

Current evidence on the long-term health effects of UFPs is limited, and existing studies have not established definitive evidence for independent health effects of UFPs from $\text{PM}_{2.5}$ [3, 4]. Hence, there are no specific ambient air quality standards or guidelines for UFPs and BC, and regulatory agencies do not commonly measure them. Although the WHO did not develop air quality guidelines for UFPs and BC, it did provide “good practice statements” for these pollutants geared toward additional monitoring, mitigation, and epidemiological research [5]. Following WHO’s guidance, the

new Ambient Air Quality Directive in Europe will improve the harmonization and availability of UFP measurements across member states, enabling consistent data collection for future Europe-wide models [43]. The new Directive includes requirements to monitor ambient air concentrations of UFP and particle number size distribution in urban supersites, as well as UFP concentrations in rural areas and pollution hotspots, including roadways with heavy traffic, airports, ports, industrial sites, among others [43].

The Directive defines UFP as particle number concentrations per cubic centimeter for a size range with a lower limit of 10 nm diameter and no upper limit [43]. This formal definition is important because, to date, no standard methods for characterizing UFP have been established, which hampers research and regulation. The lower size limit is critical because most UFPs are smaller than 20 nm, and even small differences in the low cut-off point in the 20 nm range can lead to substantial differences in particle number concentrations. The lower particle cut-offs of different instruments vary considerably, typically ranging from 2 nm to 20 nm [44]. The HEI-funded studies that included mobile monitoring for UFPs used multiple instruments with lower size cut-offs of 10–20 nm. Sheppard and colleagues used data from two different instruments (P-Trak and a NanoScan) with different size cutoffs for separate portions of their study and found some differences in resulting exposure estimates but similar spatial patterns.

Although no long-term study was identified in the HEI 2013 review [44], Ohlwein and colleagues identified 10 epidemiological studies that examined long-term UFP exposures [45], with only one study on mortality [46]. In a more recent review, Bergmann and colleagues identified 85 long-term studies published between 2011 and 2024, eight of which examined mortality [47]. See Supplementary Materials Table S2 for a summary table of the identified studies in the Bergmann review in general populations investigating total, cardiovascular or respiratory mortality [25, 37, 46, 48, 49]. Note that the analyses described in the Bauma [48] and Lloyd [25] papers were derived from the HEI studies of Hoek and Weichenthal, respectively. The Hoek study found positive associations between UFP and mortality in the DUELS cohort that were robust across exposure estimation approaches. The Weichenthal study found positive associations between UFP and mortality in the CanCHEC cohort for Toronto and Montreal, with stronger associations after adjusting for average UFP size.

The current HEI studies add to the small evidence base, but a clear research gap remains, and additional long-term studies on UFP and health outcomes are needed. The lack of long-term UFP studies was also flagged in a 2022 HEI systematic review, which identified only 10 studies focused solely on traffic-related air pollution from road transport

using strict exposure criteria [23]. An area largely unexplored by epidemiology is whether the health effects of UFPs differ across emission sources (e.g., airports, ports, or road transport) and chemical composition [50], or size distribution [11]. Routine, long-term monitoring of UFPs and BC would be valuable to support such studies. There are promising nation-wide models available [32, 51] that might facilitate large-scale multicenter studies to investigate the long-term health effects of UFP exposure in the near future.

Compared to UFPs, there is more literature on the long-term health effects of BC, but similar questions remain as to the independent health effects of BC, particularly given the often-high correlation with UFPs, NO₂, and other combustion-related indicators [23, 52, 53, 54, 55]. BC is getting attention as a “superpollutant” alongside methane as it contributes significantly to global warming and has a short atmospheric lifetime. Several countries just launched the first coordinated effort to cut BC emissions with a new initiative called the “Super Pollutant Country Action Accelerator” at COP30 in Belem [56].

Focus on People's Mobility

Like many other outdoor air pollution and health studies, three (Hoek, Sheppard and Weichenthal) of the five studies in the HEI initiative only focused on estimating outdoor concentrations at residential locations. The justification for this approach is that a substantial proportion of time is spent at home, and information on time spent outside the home is typically not collected in epidemiological studies and is not available in administrative databases [57]. However, transport microenvironments (e.g., within different modes of transportation) tend to have higher air pollutant concentrations than other settings that most people encounter in their daily lives. Although commutes are typically brief, they might substantially contribute to total daily exposure and inhalation of air pollutants [22, 58, 59].

Steps toward assessing personal exposure to outdoor pollution include accounting for individual time–activity patterns, including commuting and time spent at nonresidential locations, rather than only residential addresses. Outdoor and indoor concentration measurements can be paired to estimate infiltration factors. These steps are challenging, costly, and demanding to perform for large cohorts, and thus have understandably been beyond the scope of most studies.

Lack of consideration of infiltration rates and time–activity data introduces exposure measurement error, which is often assumed to bias estimates of outdoor air pollution and health estimates toward the null, although the nature of the potential bias is not fully known [60, 61]). A similar “bias to the null” pattern was observed in the simulation studies conducted by Katsouyanni and colleagues, in which

alternative (e.g., outdoor residential) exposure estimates yielded smaller effect estimates than those that considered daily mobility patterns and the infiltration of outdoor pollution into indoor environments.

The de Hoogh study was specifically designed to evaluate whether long-term exposure estimates that account for people's mobility would improve exposure and health estimations. The results suggested that neglecting time away from home might not introduce substantial bias in epidemiological studies. Similar findings were reported in a recent review that documented health effects across studies using residential versus time-activity-integrated air pollution exposure estimates in a small body of literature [57] and corroborated also in a recent study using Medicare data in the United States [62].

Limitations of the de Hoogh study include that the evaluation of the agent-based modeling was limited because of the small convenience sample. Furthermore, workplace locations were estimated rather than known, and the study did not incorporate pollutant infiltration rates into homes or other buildings or account for exposure induced by different modes of transportation. These factors add uncertainty and measurement error to the mobility-enhanced exposure estimates, sometimes in unpredictable ways. For example, an HEI study in Hong Kong found that when pollutant infiltration indoors was taken into account, mobility-enhanced exposures were on average about 20% lower than residential exposures because infiltration of outdoor air pollution was higher in residences than in commercial buildings [63]. On the other hand, depending on the transportation mode used (e.g., car, bus, bicycle, or walking), commuting can increase exposures by as much as 400% compared to typical urban background concentrations [58]. Moreover, exposure can vary by transportation mode and by pollutant, as motorized modes (cars, motorcycles, buses, tram, rail, and subway) tend to have higher air pollutant concentrations of gaseous pollutants, such as NO₂, than active modes (cycling and walking). For particulate matter, the results varied by particle size and transportation mode [64, 65]. It is important to note that cumulative exposure and inhaled dose during commuting might also vary between transport modes, due to differences in route length and duration, and increased inhalation rates for active modes of transport [66, 67].

Although the finding that accounting for mobility might not be an important consideration is reassuring, a direct assessment from a personal exposure model would provide further insights and further research is needed — especially research about pollutants other than PM_{2.5} and NO₂ and in other locations and populations — to examine fully the added value of collecting mobility information on a large scale in air pollution cohort studies.

Accounting for Exposure Measurement Error in Health Analyses

Methodological complexity and lack of data prevented three of the five studies in the HEI initiative from propagating exposure measurement error into the health effects estimates. Only a very limited proportion of air pollution epidemiological studies apply bias-correction methods to account for measurement error [68, 69]. Each correction method relies on large assumptions that yield different advantages and disadvantages, and their complexity hinders widespread application.

Two HEI studies (Katsouyanni and Sheppard) adjusted for exposure measurement error in the health analyses. Sheppard used a nonparametric bootstrap approach that is complex and computationally demanding as the approach also requires the simulation of the exposure surface, not just correction in the epidemiological analysis stage. Hence, bootstrap approaches cannot be applied in very large administrative cohorts. Also, bootstrap approaches rely on the assumption of spatial “compatibility” of monitoring sites and cohort locations (which was part of the design of the Sheppard monitoring campaign), whereas in practice, spatially misaligned exposure data is more common. Katsouyanni used simulation extrapolation and regression calibration, both of which entail their own strong assumptions and simplifications. Importantly, with the exception of bootstrap approaches, measurement error correction methods typically rely on the existence of a “gold-standard”, which is true long-term personal exposure to outdoor air pollution. Such a gold standard is nearly impossible to obtain.

The Katsouyanni study demonstrated some of the difficulties in obtaining a “true” measure. Even with exceptionally long and detailed personal monitoring data for the London study area, major assumptions still had to be made to estimate long-term personal exposures to outdoor air pollution, which adds uncertainty. In practice, however, true personal exposure to outdoor air pollution cannot be observed directly. Measurement errors derived between so-called “true” and other estimated exposures to air pollution might be related to many different issues, including the reliability and validity of measurement instruments, methods for estimating infiltration efficiency, and incomplete information about where people spend their time. Personal exposure studies are generally limited to small numbers of participants, raising concerns about representativeness.

Accounting for the inherent (spatially varying) uncertainty and biases in modeled estimates of air pollution exposure remains largely an unresolved problem in air pollution epidemiology [61, 68], and advances in this area are much needed. Additional future work in this field might also consider applying such correction methods to epidemiological

models that simultaneously consider multiple pollutants. More work may also be needed to consider spatially correlated exposures that might either confound or modify the health effects of air pollution, such as noise and lack of green space [70].

For the multi-pollutant analyses in the current initiative, the investigators added up to four pollutants as potential confounders in the health model typically as linear terms. These models help understand how pollutants affect the risk when additional adjustments for other pollutants are being made, but they might not adequately capture the complex relationships among the different pollutants. The development of statistical approaches to address multi-pollutant mixtures and multiple exposures remains an active area of research, and many advanced approaches have been developed, particularly for omics analyses and in studies of the exposome [71]. Most multi-pollutant statistical approaches to date, however, cannot accommodate very large datasets such as DUELS.

As shown by the Sheppard study, accounting for exposure measurement error does not negate the importance of careful exposure study design. The collection of long-term personal measurements for large populations can be challenging but carefully designed validation subsamples might provide insight for exposure measurement error corrections. For future cohort studies, collecting more detailed information on workplace locations, and information on time spent outdoors could reduce uncertainty in activity-based exposure assessments.

Low-cost Sensors

Regulatory monitoring networks are sparse due to the high cost, complexity of the equipment, and need for trained operators [72]. In recent years, low-cost air quality sensors gained popularity for supplementing traditional regulatory or research-grade air quality monitoring, enhancing spatial and temporal coverage across the globe. Low-cost sensors have made it possible to gather real-time air pollution data at a high spatial resolution and increase public awareness and public information if the data are made publicly available [73]. Additionally, low-cost sensors are relatively easy to operate, allowing non-technical users, such as community groups and individuals, to participate in air quality monitoring. Hence, the application of low-cost sensors is only set to grow.

Low-cost sensors only played a modest role in the HEI initiative. Their application was primarily limited to testing the added value of low-cost sensor data in hybrid exposure models that combined data from regulatory monitors and other research-grade monitors or other data sources (e.g., dispersion model estimates). Unsurprisingly, the added

value of low-cost sensor data is diminished when a large amount of spatial information is already available in the model from many additional locations and data sources. Sheppard and colleagues found that low-cost sensor data deployed at ACT-participants' residences improved exposure estimates for $PM_{2.5}$ but had little effect on estimates for NO_2 . Hoek and colleagues saw little improvement in exposure estimates for $PM_{2.5}$ or NO_2 when low-cost sensor data were used together with dispersion modeling or data from the regulatory monitoring network. They noted that although equipment costs were low, a large personnel effort was needed to calibrate and maintain the sensors.

Despite the appeal and advances, uncertainties and limitations of low-cost sensors — most importantly, their lower accuracy compared to reference instruments — remain. Known limitations include interferences from environmental conditions such as temperature, humidity, cross-sensitivities with other gases, and a low signal-to-noise ratio. Low-cost sensors must therefore be calibrated for their data to be incorporated into exposure models, and this might take substantial time and effort. The U.S. EPA has established performance assessment protocols for low-cost sensors measuring $PM_{2.5}$, ozone, and other pollutants [74, 75, 76]. Those protocols provide information on collocation, calibration, performance evaluation, and deployment. Common protocols are needed given the variety of techniques currently deployed [77]. Additional ethical, privacy, logistical, and data-science challenges remain as areas for further research [73].

Harnessing Machine Learning for Environmental Epidemiology

So far, linear regression methods have been the most popular techniques for developing land-use regression models to estimate outdoor air pollutant concentrations in epidemiological studies [9]. In recent years, the use of machine learning methods for air pollution modelling has rapidly increased to address outliers, complex nonlinear relationships, and highly correlated predictors. Potential limitations of machine learning methods include the need for large numbers of variables for prediction, which might not be available in all study areas or time periods, and high computational resource needs. The parameters used to predict exposures are less interpretable than standard regression-based approaches [78].

To date, whether machine learning methods outperform linear regression methods is unclear. Differences in performance between machine learning and more established land-use regression models were modest in the HEI initiative, suggesting that both approaches can be used interchangeably. The Weichenthal study found generally similar

exposure estimates using a convolutional neural network operating on aerial images in comparison to estimates from a traditional land-use regression framework but suggested the neural network approach would be valuable for locations without extensive traffic or land use data. Hoek and colleagues found little difference between supervised linear regression, random forest and Least Absolute Shrinkage and Selection Operator algorithms across a variety of exposure estimation approaches using different input datasets. The de Hoogh study found that a random forest algorithm modestly outperformed supervised linear regression, while the Shepard study found a universal kriging approach performed slightly better than the machine learning models they tested.

These findings contrast with a recent review that reported that machine learning models outperformed the traditional linear models by 0.12 (R^2) [79]. The superior performance of machine learning models was observed with varied study designs and methods, different spatial extents (from urban to global), temporal resolutions (from hourly to annual), monitoring types (mobile and fixed), and algorithms [79]. Only few algorithms were tested in the HEI initiative, however, and the focus was largely on predicting spatial contrasts, which is most relevant for estimating long-term exposure and health outcomes. Indeed, spatiotemporal models (e.g., often used to predict hourly, daily, or monthly concentrations) appeared to benefit more from machine learning methods than spatial models (e.g., used to predict annual averages). A reason might be that longer averaging times can simplify the relationships between predictors and the pollutants, making it easier to model [79]. Also, findings might differ in different settings and might depend on whether internal or external validation data are used to compare model performance. Very few studies [32, 80, 81] in the Vachon review used external validation datasets partly because they are rare for pollutants not routinely measured such as UFPs and BC [79]. Two of the comparisons had conflicting performance differences between cross-validation and external validation [32, 81]. Chen et al. evaluated 16 modeling techniques to predict $PM_{2.5}$ and NO_2 concentrations in Europe and found no performance difference between statistical and machine learning methods using either cross-validation or external validation [80].

Artificial intelligence technology, including machine learning methods is being increasingly adopted for environmental health research. Artificial intelligence provides the opportunity to augment and accelerate research, helping scientists to generate hypotheses, design studies, collect and interpret large datasets, and gain insights that might not have been possible using traditional scientific methods alone. However, there are fears concerning loss of privacy, lack of transparency and accountability, increase in inequality, the carbon footprint and other issues affecting society [78, 82,

83]. The use of artificial intelligence beyond machine learning models for exposure assessment was outside the scope of the current HEI initiative. HEI has taken various steps incorporating artificial intelligence technologies, including a recent research initiative focused on advancing satellite remote sensing for use in health studies [84].

Conclusion

HEI's initiative to improve long-term exposure assessment of outdoor air pollution for health studies in North America and Europe provided valuable lessons to inform decisions about estimating air pollution in epidemiological studies, as investigators balance concerns about bias, efficiency, and resource needs.

Exposures predicted from the various models generally resulted in moderately to highly correlated exposure estimates. Health effect estimates associated with the alternative exposures were similar, albeit with some notable differences in the magnitude – but not the direction — of the associations. Notably, neglecting time away from home might not introduce substantial bias in epidemiological studies that are typically focused on estimating outdoor concentrations at residential locations. Accounting for exposure measurement error in health analyses typically led to increases in the magnitude of the associations and wider confidence intervals in simulation studies. Importantly, accounting for exposure measurement error does not negate the importance of careful exposure study design, such as including samples beyond business hours, extended times of the day and weekends in mobile monitoring campaigns. Further, differences in performance between machine learning and more established land-use regression models were modest, suggesting that both approaches can be used interchangeably. The added value of low-cost sensor data is diminished when a large amount of spatial information is already available in the exposure model from many additional locations and data sources, and when the personnel effort required to calibrate and maintain the sensors is substantial.

Across the studies, common limitations included the short duration and limited number of samples available for validation, and the temporal misalignment between health outcome data and recently acquired air pollution data for exposure modeling. Reconciling modeled exposures with measured long-term personal exposure remains a challenging research topic. Future work is needed to replicate findings in other studies and regions. Other areas of future work include generating representative validation data for long-term exposure estimates, propagating exposure measurement error more fully into health estimates, and improving understanding of the influence of indoor environments.

Key References

- Hoek G, Vienneau D, de Hoogh K. Does residential address-based exposure assessment for outdoor air pollution lead to bias in epidemiological studies? *Environ Health*. 2024;23:75. <https://doi.org/10.1186/s12940-024-01111-0>.
 - A review that documented health effects across studies using residential versus time-activity-integrated air pollution exposure estimates in a small body of literature.
- Kerckhoffs J, Hofman J, Khan J, Adams MD, Blanco MN, deSouza P, et al. Mobile monitoring of air pollution - a position paper on use cases, good practices, challenges, and opportunities. *Environ Int*. 2025;202:109582. <https://doi.org/10.1016/j.envint.2025.109582>.
 - A position paper that discussed good practices in terms of mobile monitoring design, data processing and modeling of air pollution.
- Kim SY, Blanco MN, Bi J, Larson TV, Sheppard L. Exposure assessment for air pollution epidemiology: A scoping review of emerging monitoring platforms and designs. *Environ Res*. 2023;223:115451. <https://doi.org/10.1016/j.envres.2023.115451>.
 - A review that described novel monitoring platforms and designs for air pollution cohort studies.
- Vachon J, Kerckhoffs J, Buteau S, Smargiassi A. Do machine learning methods improve prediction of ambient air pollutants with high spatial contrast? A systematic review. *Environ Res*. 2024;262:119751. <https://doi.org/10.1016/j.envres.2024.119751>.
 - A review that compared machine learning and linear methods to predict air pollution concentrations.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s40572-026-00560-8>.

Acknowledgements HEI is indebted to an independent expert panel for project oversight. Specifically, we would like to thank Jeff Brook, Francesca Dominici, Markey Johnson, Heather Holmes, Marianthi-Anna Kioumourtoglou, Neil Pearce, Allen Robinson, Evi Samoli, David Savitz, and Greg Wellenius for providing advice and feedback on the study design, analytical plans, and progress. In addition, we would like to thank Pallavi Pant, Elise Elliott, and Elena Craft for valuable guidance. We also would like to thank the original investigators Gerard Hoek, Kees de Hoogh, Klea Katsouyanni, Lianne Sheppard, Scott Weichenthal, and their teams for their efforts in designing

and conducting these comprehensive studies. Research described in this article was conducted under contract to the HEI, an organization jointly funded by the U.S. Environmental Protection Agency (Assistance Award No. CR- 83998101) and certain motor vehicle and engine manufacturers. The views in this article are those of the authors and do not necessarily reflect the views of the HEI or its sponsors.

Author contributions Conceptualization: HB, DLC, APP, EM; Investigation: HB, DLC; Supervision: HB, EM, JBM; Writing original draft: HB; Writing review and editing: DLC, APP, EM, SBB, JEM, AO, SY, JDD, JV, JBM.

Funding HEI receives balanced funding from the United States Environmental Protection Agency and the worldwide motor vehicle industry.

Data Availability No datasets were generated or analysed during the current study.

Declarations

Human and Animal Rights and Informed Consent This article does not contain any studies with human or animal subjects performed by any of the authors.

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. SOGA (State of Global Air). A report on air pollution and its role in the world's leading causes of death. Boston, MA: Health Effects Institute; 2025.
2. US EPA (United States Environmental Protection Agency). Integrated science assessment for oxides of nitrogen—health Criteria. EPA/600/R-15/068. Washington, DC: US EPA; 2016.
3. US EPA (United States Environmental Protection Agency). Integrated science assessment for particulate matter. EPA/600/R-19/188. Washington, DC: US EPA; 2019.
4. US EPA (United States Environmental Protection Agency). Supplement to the 2019 integrated science assessment for particulate matter. EPA/600/R-22/028. Washington, DC: US EPA; 2022.
5. WHO (World Health Organization). WHO Global Air Quality Guidelines: Particulate matter (PM_{2.5} and PM₁₀), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. Geneva: WHO; 2021.
6. Dockery DW, Pope CA 3rd, Xu X, Spengler JD, Ware JH, Fay ME, et al. An association between air pollution and mortality in

- six US cities. *N Engl J Med*. 1993;329:1753–9. <https://doi.org/10.1056/nejm199312093292401>.
7. Pope CA 3rd, Burnett RT, Thun MJ, Calle EE, Krewski D, Ito K, et al. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *JAMA*. 2002;287:1132–41. <https://doi.org/10.1001/jama.287.9.1132>.
 8. Roque NA, Andrews H, Santos-Lozada AR. Identifying air quality monitoring deserts in the United States. *Proc Natl Acad Sci U S A*. 2025;122:e2425310122. <https://doi.org/10.1073/pnas.2425310122>.
 9. Hoek G. Methods for assessing long-term exposures to outdoor air pollutants. *Curr Environ Health Rep*. 2017;4:450–62. <https://doi.org/10.1007/s40572-017-0169-5>.
 10. HEI (Health Effects Institute). RFA 19–1: Applying novel approaches to improve long-term exposure assessment of outdoor air pollution for health studies. Boston, MA: Health Effects Institute; 2019.
 11. Weichenthal S, Lloyd M, Ganji A, Simon L, Xu J, Venuta A, et al. Long-term exposure to outdoor ultrafine particles and black carbon and effects on mortality in Montreal and Toronto, Canada. *Res Rep Health Eff Inst*. 2024;217:1–63.
 12. Sheppard L, Blanco MN, Doubleday A, Szpiro AA, Bi J, Zuidema C, et al. Optimizing air pollution exposure assessment with application to cognitive function. *Res Rep Health Eff Inst*. 2025;228:1–117.
 13. de Hoogh K, Flückiger B, Probst-Hensch N, Vienneau D, Jeong A, Imboden M, et al. Accounting for mobility in air pollution exposure estimates in studies on long-term health effects. *Res Rep Health Eff Inst*. 2025;229:1–60.
 14. Hoek G, Bouma F, Janssen N, Wesseling J, van Ratingen S, Kerckhoffs J, et al. Comparison of long-term air pollution exposure from mobile and routine monitoring, low-cost sensors, and dispersion models. *Res Rep Health Eff Inst*. 2025;226:1–101.
 15. Katsouyanni K, Evangelopoulos D, Wood D, Barratt B, Zhang H, Walton H, et al. Investigating the consequences of measurement error of gradually more sophisticated long-term personal exposure models in assessing health effects: The London study (MELONS). *Res Rep Health Eff Inst*. 2025;227:1–78.
 16. Bergen S, Sheppard L, Kaufman JD, Szpiro AA. 2016. Multi-pollutant measurement error in air pollution epidemiology studies arising from predicting exposures with penalized regression splines. *J R Stat Soc Ser C-Appl Stat*. 2016;65:731–753. <https://doi.org/10.1111/rssc.12144>
 17. Szpiro AA, Paciorek CJ, Sheppard L. Does more accurate exposure prediction necessarily improve health effect estimates? *Epidemiology*. 2011;22:680–5. <https://doi.org/10.1097/EDE.0b013e3182254cc6>.
 18. Szpiro AA, Sheppard L, Lumley T. Efficient measurement error correction with spatially misaligned data. *Biostatistics*. 2011;12:610–23. <https://doi.org/10.1093/biostatistics/kxq083>.
 19. Crooks AT, Heppenstall AJ. Introduction to agent-based modeling. Agent-Based Models of Geographical Systems. In: Heppenstall AJ, Crooks AT, See LM, Batty M, editors. Springer; 2012. p. 85–105.
 20. de Hoogh K, Héritier H, Stafoggia M, Künzli N, Kloog I. Modelling daily PM_{2.5} concentrations at high spatio-temporal resolution across Switzerland. *Environ Pollut*. 2018;233:1147–54. <https://doi.org/10.1016/j.envpol.2017.10.025>.
 21. de Hoogh K, Saucy A, Shtein A, Schwartz J, West EA, Strassmann A, et al. Predicting fine-scale daily NO₂ for 2005–2016: incorporating OMI satellite data across Switzerland. *Environ Sci Technol*. 2019;53:10279–87. <https://doi.org/10.1021/acs.est.9b03107>.
 22. de Nazelle A, Seto E, Donaire-Gonzalez D, Mendez M, Matamala J, Nieuwenhuijsen MJ, et al. Improving estimates of air pollution exposure through ubiquitous sensing technologies. *Environ Pollut*. 2013;176:92–9. <https://doi.org/10.1016/j.envpol.2012.12.032>.
 23. HEI (Health Effects Institute). Systematic review and meta-analysis of selected health effects of long-term exposure to traffic-related air pollution. Special Report 23. Boston, MA: Health Effects Institute; 2022.
 24. Kim SY, Blanco MN, Bi J, Larson TV, Sheppard L. Exposure assessment for air pollution epidemiology: A scoping review of emerging monitoring platforms and designs. *Environ Res*. 2023;223:115451. <https://doi.org/10.1016/j.envres.2023.115451>.
 25. Lloyd M, Olaniyan T, Ganji A, Xu J, Venuta A, Simon L, et al. Airborne Nanoparticle Concentrations Are Associated with Increased Mortality Risk in Canada's Two Largest Cities. *Am J Respir Crit Care Med*. 2024;210:1338–47. <https://doi.org/10.1164/rccm.202311-2013OC>.
 26. Ndiaye A, Vienneau D, Flückiger B, Probst-Hensch N, Jeong A, Imboden M, et al. Associations between long-term air pollution exposure and mortality and cardiovascular morbidity: A comparison of mobility-integrated and residential-only exposure assessment. *Environ Int*. 2025;198:109387. <https://doi.org/10.1016/j.envint.2025.109387>.
 27. Apte JS, Chambliss SE, Messier KP, Gani S, Upadhyaya AR, Kushwaha M, et al. Scalable multipollutant exposure assessment using routine mobile monitoring platforms. *Res Rep Health Eff Inst*. 2024;216:1–54.
 28. Messier KP, Chambliss SE, Gani S, Alvarez R, Brauer M, Choi JJ, et al. Mapping air pollution with Google street view cars: Efficient approaches with mobile monitoring and land use regression. *Environ Sci Technol*. 2018;52:12563–72. <https://doi.org/10.1021/acs.est.8b03395>.
 29. Hatzopoulou M, Valois MF, Levy I, Mihele C, Lu G, Bagg S, et al. Robustness of land-use regression models developed from mobile air pollutant measurements. *Environ Sci Technol*. 2017;51:3938–47. <https://doi.org/10.1021/acs.est.7b00366>.
 30. Kerckhoffs J, Hofman J, Khan J, Adams MD, Blanco MN, deSouza P, et al. Mobile monitoring of air pollution - a position paper on use cases, good practices, challenges, and opportunities. *Environ Int*. 2025;202:109582. <https://doi.org/10.1016/j.envint.2025.109582>.
 31. Blanco MN, Gasset A, Gould T, Doubleday A, Slager DL, Austin E, et al. Characterization of annual average traffic-related air pollution concentrations in the Greater Seattle Area from a year-long mobile monitoring campaign. *Environ Sci Technol*. 2022;56:11460–72. <https://doi.org/10.1021/acs.est.2c01077>.
 32. Shen Y, Kerckhoffs J, Stafoggia M, Nobile F, Cattani G, Peters A, et al. Europe-Wide Spatial Models for Ultrafine Particles Based on Mobile Monitoring. *Environ Sci Technol*. 2025;59:27033–43. <https://doi.org/10.1021/acs.est.5c03380>.
 33. Boogaard H, Crouse DL, Tanner E, Mantus E, van Erp AM, Vedal S, et al. Assessing adverse health effects of long-term exposure to low levels of ambient air pollution: The HEI experience and what's next? *Environ Sci Technol*. 2024;58:12767–83. <https://doi.org/10.1021/acs.est.3c09745>.
 34. Alexeeff SE, Roy A, Shan J, Liu X, Messier K, Apte JS, et al. High-resolution mapping of traffic-related air pollution with Google street view cars and incidence of cardiovascular events within neighborhoods in Oakland, CA. *Environ Health*. 2018;17:38. <https://doi.org/10.1186/s12940-018-0382-1>.
 35. Bai L, Weichenthal S, Kwong JC, Burnett RT, Hatzopoulou M, Jerrett M, et al. Associations of long-term exposure to ultrafine particles and nitrogen dioxide with increased incidence of congestive heart failure and acute myocardial infarction. *Am J Epidemiol*. 2019;188:151–9. <https://doi.org/10.1093/aje/kwy194>.
 36. Downward GS, van Nunen EJHM, Kerckhoffs J, Vineis P, Brunekreef B, Boer JMA, et al. Long-term exposure to ultrafine particles and incidence of cardiovascular and cerebrovascular disease in a prospective study of a Dutch cohort. *Environ Health Perspect*. 2018;126:127007. <https://doi.org/10.1289/EHP3047>.

37. Pond ZA, Saha PK, Coleman CJ, Presto AA, Robinson AL, Arden Pope C. Mortality risk and long-term exposure to ultrafine particles and primary fine particle components in a national U.S. Cohort. *Environ Int*. 2022;167:107439. <https://doi.org/10.1016/j.envint.2022.107439>.
38. Weichenthal S, Bai L, Hatzopoulou M, Van Ryswyk K, Kwong JC, Jerrett M, et al. Long-term exposure to ambient ultrafine particles and respiratory disease incidence in Toronto, Canada: A cohort study. *Environ Health*. 2017;16:64. <https://doi.org/10.1186/s12940-017-0276-7>.
39. Montagne DR, Hoek G, Klompmaker JO, Wang M, Meliefste K, Brunekreef B. Land use regression models for ultrafine particles and black carbon based on short-term monitoring predict past spatial variation. *Environ Sci Technol*. 2015;49:8712–20. <https://doi.org/10.1021/es505791g>.
40. Bakola M, Hernandez Carballo I, Jelastopulu E, Stuckler D. 2022. The impact of COVID-19 lockdown on air pollution in Europe and North America: a systematic review. *Eur J Public Health*. 2022;32:962–968. <https://doi.org/10.1093/eurpub/ckac118>
41. Venter ZS, Aunan K, Chowdhury S, Lelieveld J. COVID-19 lockdowns cause global air pollution declines. *Proc Natl Acad Sci U S A*. 2020;117:18984–90. <https://doi.org/10.1073/pnas.2006853117>.
42. Samet J, Shairsingh K, Ye W, Gumy S, Mudu P, Andersen Z, et al. Transforming air pollution and health research into action in low- and middle-income countries. *Environ Epidemiol*. 2025;9:e434. <https://doi.org/10.1097/EE9.0000000000000434>.
43. EU (European Union). 2024. Directive (EU) 2024/2881 of the European Parliament and of the Council of 23 October 2024 on Ambient Air Quality and Cleaner Air for Europe. Available: <https://eurlex.europa.eu/eli/dir/2024/2881/oj/eng>
44. HEI Review Panel on Ultrafine Particles. 2013. Understanding the health effects of ambient ultrafine particles. HEI Perspective 3. Health Effects Institute, Boston, MA.
45. Ohlwein S, Kappeler R, Kutlar Joss M, Künzli N, Hoffmann B. Health effects of ultrafine particles: A systematic literature review update of epidemiological evidence. *Int J Public Health*. 2019;64:547–59. <https://doi.org/10.1007/s00038-019-01202-7>.
46. Ostro B, Hu J, Goldberg D, Reynolds P, Hertz A, Bernstein L, et al. Associations of mortality with long-term exposures to fine and ultrafine particles, species and sources: Results from the California Teachers Study Cohort. *Environ Health Perspect*. 2015;123:549–56. <https://doi.org/10.1289/ehp.1408565>.
47. Bergmann M, Haddad-Thoelke P, Jeong H, Kappeler R, Altug H, Oberwinster L, et al. Systematic review and meta-analysis on the health effects of long-term exposure to ultrafine particles. *Eur Respir Rev*. 2026;35:250156. <https://doi.org/10.1183/16000617.0156-2025>.
48. Bouma F, Janssen NA, Wesseling J, van Ratingen S, Strak M, Kerckhoffs J, et al. Long-term exposure to ultrafine particles and natural and cause-specific mortality. *Environ Int*. 2023;175:107960. <https://doi.org/10.1016/j.envint.2023.107960>.
49. Qi Q, Yu F, Nair AA, Lau SSS, Luo G, Mithu I, et al. Hidden danger: The long-term effect of ultrafine particles on mortality and its sociodemographic disparities in New York State. *J Hazard Mater*. 2024;471:134317. <https://doi.org/10.1016/j.jhazmat.2024.134317>.
50. Vallabani NVS, Gruziova O, Elihn K, Juárez-Facio AT, Steimer SS, Kuhn J, et al. Toxicity and health effects of ultrafine particles: Towards an understanding of the relative impacts of different transport modes. *Environ Res*. 2023;231:116186. <https://doi.org/10.1016/j.envres.2023.116186>.
51. Saha PK, Hankey S, Marshall JD, Robinson AL, Presto AA. High-spatial-resolution estimates of ultrafine particle concentrations across the continental United States. *Environ Sci Technol*. 2021;55:10320–31. <https://doi.org/10.1021/acs.est.1c03237>.
52. Janssen NA, Hoek G, Simic-Lawson M, Fischer P, van Bree L, ten Brink H, et al. Black carbon as an additional indicator of the adverse health effects of airborne particles compared with PM10 and PM2.5. *Environ Health Perspect*. 2011;119:1691–9. <https://doi.org/10.1289/ehp.1003369>.
53. Luben TJ, Nichols JL, Dutton SJ, Kirrane E, Owens EO, Datko-Williams L, et al. A systematic review of cardiovascular emergency department visits, hospital admissions and mortality associated with ambient black carbon. *Environ Int*. 2017;107:154–62. <https://doi.org/10.1016/j.envint.2017.07.005>.
54. Song X, Hu Y, Ma Y, Jiang L, Wang X, Shi A, et al. Is short-term and long-term exposure to black carbon associated with cardiovascular and respiratory diseases? A systematic review and meta-analysis based on evidence reliability. *BMJ Open*. 2022;12:e049516. <https://doi.org/10.1136/bmjopen-2021-049516>.
55. Zhu X, Liu B, Guo C, Li Z, Cheng M, Zhu X, Wei Y. Short and long-term association of exposure to ambient black carbon with all-cause and cause-specific mortality: A systematic review and meta-analysis. *Environ Pollut*. 2023;324:121086. <https://doi.org/10.1016/j.envpol.2023.121086>.
56. CCAC Super Pollutant Country Action Accelerator. Embedding super pollutant coordination teams in governments. Available here: <https://www.ccacoalition.org/content/ccac-super-pollutant-country-action-accelerator>
57. Hoek G, Vienneau D, de Hoogh K. Does residential address-based exposure assessment for outdoor air pollution lead to bias in epidemiological studies? *Environ Health*. 2024;23:75. <https://doi.org/10.1186/s12940-024-01111-0>.
58. de Nazelle A, Bode O, Orjuela JP. Comparison of air pollution exposures in active vs. passive travel modes in European cities: a quantitative review. *Environ Int*. 2017;99:151–60. <https://doi.org/10.1016/j.envint.2016.12.023>.
59. Koehler K, Good N, Wilson A, Mölter A, Moore BF, Carpenter T, et al. The Fort Collins commuter study: Variability in personal exposure to air pollutants by microenvironment. *Indoor Air*. 2019;29:231–41. <https://doi.org/10.1111/ina.12533>.
60. Richmond-Bryant J, Long TC. Influence of exposure measurement errors on results from epidemiologic studies of different designs. *J Expo Sci Environ Epidemiol*. 2020;30:420–9. <https://doi.org/10.1038/s41370-019-0164-z>.
61. Sheppard L, Burnett RT, Szpiro AA, Kim SY, Jerrett M, Pope CA 3rd, et al. Confounding and exposure measurement error in air pollution epidemiology. *Air Qual Atmos Health*. 2012;5:203–16. <https://doi.org/10.1007/s11869-011-0140-9>.
62. Antonelli J, Shin H, Kang S, Franks A, Audirac M, Braun D. Robust statistical approaches to understanding the causal effects of air pollution mixtures. *Res Rep Health Eff Inst*. 2025;234:1–63.
63. Barratt B, Lee M, Wong P, Tang R, Tsui TH, Cheng W, et al. A dynamic three-dimensional air pollution exposure model for Hong Kong. *Res Rep Health Eff Inst*. 2018;194:1–65.
64. Ramel-Delobel M, Heydari S, de Nazelle A, Praud D, Salizzoni P, Fervers B, et al. Air pollution exposure in active versus passive travel modes across five continents: a Bayesian random-effects meta-analysis. *Environ Res*. 2024;261:119666. <https://doi.org/10.1016/j.envres.2024.119666>.
65. Ridolfo S, Amato F, Querol X. Particle number size distributions and concentrations in transportation environments: a review. *Environ Int*. 2024;187:108696. <https://doi.org/10.1016/j.envint.2024>.
66. Good N, Carpenter T, Anderson GB, Wilson A, Peel JL, Browning RC, et al. Development and validation of models to predict personal ventilation rate for air pollution research. *J Expo Sci Environ Epidemiol*. 2019;29:568–77. <https://doi.org/10.1038/s41370-018-0067-4>.
67. Good N, Mölter A, Ackerson C, Bachand A, Carpenter T, Clark ML, et al. The Fort Collins Commuter Study: Impact of route

- type and transport mode on personal exposure to multiple air pollutants. *J Expo Sci Environ Epidemiol*. 2016;26:397–404. <https://doi.org/10.1038/jes.2015.68>.
68. Samoli E, Butland BK. Incorporating measurement error from modeled air pollution exposures into epidemiological analyses. *Curr Environ Health Rep*. 2017;4:472–80. <https://doi.org/10.1007/s40572-017-0160-1>.
69. Shaw PA, Deffner V, Keogh RH, Tooze JA, Dodd KW, Küchenhoff H, et al. Epidemiologic analyses with error-prone exposures: review of current practice and recommendations. *Ann Epidemiol*. 2018;28:821–8. <https://doi.org/10.1016/j.annepidem.2018.09.001>.
70. Williams H, Andersen ZJ, Boogaard H, Brage S, Browning MHEM, Cai S, et al. Expert perspectives on exposure-response functions for urban health policy: Lessons from a UBDPolicy workshop. *Environ Res*. 2026;288:123150. <https://doi.org/10.1016/j.envres.2025.123150>.
71. Stafoggia M, Breitner S, Hampel R, Basagaña X. Statistical approaches to address multi-pollutant mixtures and multiple exposures: The state of the science. *Curr Environ Health Rep*. 2017;4:481–90. <https://doi.org/10.1007/s40572-017-0162-z>.
72. Snyder EG, Watkins TH, Solomon PA, Thoma ED, Williams RW, Hagler GS, et al. The changing paradigm of air pollution monitoring. *Environ Sci Technol*. 2013;47:11369–77. <https://doi.org/10.1021/es4022602>.
73. Bagkis E, Hassani A, Schneider P, DeSouza P, Shetty S, Kassandros T, et al. Evolving trends in application of low-cost air quality sensor networks: challenges and future directions. *npj Clim Atmos Sci*. 2025;8:335. <https://doi.org/10.1038/s41612-025-01216-4>.
74. Duvall R, Clements A, Hagler G, Kamal A, Kilaru V, Goodman L et al. 2021. Performance testing protocols, metrics, and target values for fine particulate matter air sensors: Use in ambient, outdoor, fixed sites, non-regulatory supplemental and informational monitoring applications, US EPA Office of Research and Development, EPA/600/R-20/280.
75. Duvall R, Clements A, Hagler G, Kamal A, Kilaru V, Goodman L, US EPA Office of Research and Development, Washington. Performance testing protocols, metrics, and target values for ozone air sensors: Use in ambient, outdoor, fixed sites, non-regulatory supplemental and informational monitoring applications. Volume 600/R–20/279. EPA; 2021.
76. Duvall R, Clements A, Barkjohn K, Kumar M, Greene D, Dye T, U.S. Environmental Protection Agency. NO₂, CO, and SO₂ supplement to the 2021 report on performance testing protocols, metrics, and target values for ozone air sensors. Washington, DC: EPA/600/R-23/146; 2024.
77. Hayward I, Martin NA, Ferracci V, Kazemimanesh M, Kumar P. Low-cost air quality sensors: biases, corrections and challenges in their comparability. *Atmosphere*. 2024;15:1523. <https://doi.org/10.3390/atmos15121523>.
78. Iyer HS, Karasaki S, Yi L, Hswen Y, James P, VoPham T. Harnessing geospatial artificial intelligence (GeoAI) for environmental epidemiology: A narrative review. *Curr Environ Health Rep*. 2025;12:34. <https://doi.org/10.1007/s40572-025-00497-4>.
79. Vachon J, Kerckhoffs J, Buteau S, Smargiassi A. Do machine learning methods improve prediction of ambient air pollutants with high spatial contrast? A systematic review. *Environ Res*. 2024;262:119751. <https://doi.org/10.1016/j.envres.2024.119751>.
80. Chen J, de Hoogh K, Gulliver J, Hoffmann B, Hertel O, Ketzel M, et al. A comparison of linear regression, regularization, and machine learning algorithms to develop Europe-wide spatial models of fine particles and nitrogen dioxide. *Environ Int*. 2019;130:104934. <https://doi.org/10.1016/j.envint.2019>.
81. Kerckhoffs J, Hoek G, Portengen L, Brunekreef B, Vermeulen RCH. Performance of prediction algorithms for modeling outdoor air pollution spatial surfaces. *Environ Sci Technol*. 2019;53:1413–21. <https://doi.org/10.1021/acs.est.8b06038>.
82. Huang L, Duan Q, Liu Y, Wu Y, Li Z, Guo Z, et al. Artificial intelligence: A key fulcrum for addressing complex environmental health issues. *Environ Int*. 2025;198:109389. <https://doi.org/10.1016/j.envint.2025.109389>.
83. Weichenthal S, Hatzopoulou M, Brauer M. A picture tells a thousand... exposures: Opportunities and challenges of deep learning image analyses in exposure science and environmental epidemiology. *Environ Int*. 2019;122:3–10. <https://doi.org/10.1016/j.envint.2018.11.042>.
84. HEI (Health Effects Institute). RFA 25–1: Advancing satellite-derived air quality data and approaches for use in health studies. Boston, MA: Health Effects Institute; 2025.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.