Health Effects Institute

Protocol for a Systematic Review and Meta–Analysis of Selected Health Effects of Long–Term Exposure to Traffic–Related Air Pollution

JULY 31, 2019

Table of Contents

Citation: Health Effects Institute. 2019. Protocol for a Systematic Review and Meta–Analysis of Selected Health Effects of Long–Term Exposure to Traffic–Related Air Pollution. Boston, MA:Health Effects Institute.

1. INTRODUCTION

In 2018, the HEI Board of Directors appointed an expert HEI Panel to review the traffic–related air pollution (TRAP) and health literature. The findings will be published as an HEI Special Report in summer 2021. The Panel consists of 13 experts in epidemiology, exposure assessment and biostatistics, and is chaired by Francesco Forastiere (King's College London, UK) and Fred Lurmann (Sonoma Technology, Inc., Petaluma, California). See Appendix 1 for List A of Panel members. In addition, HEI hired a contractor team at Swiss Tropical and Public Health Institute, Switzerland, to execute certain parts of the review, particularly bibliographic searches, data extraction, and parts of the data synthesis, in close collaboration with HEI staff and Panel members. HEI may complement expertise on the Panel with a few consultants during the course of the review.

This document is a protocol for the conduct of a systematic review. It describes the rationale; objectives for the review; research questions; methods to search the literature, assess study quality, summarize results and reach conclusions; and requirements for reporting. The development of this protocol has been largely based on standards set by the Cochrane Collaboration (Higgins and Green, 2011), standards of the Preferred Reporting Items for Systematic Review and Meta–analysis Protocols (PRISMA–P) (Shamseer et al., 2015; Moher et al., 2015), the systematic reviews conducted as part of the WHO Air Quality Guidelines Update (e.g., Vilahur et al., 2017), and the NIEHS Office of Health Assessment and Translation handbook (OHAT, 2019). This protocol has been adapted for application to observational studies.

The review protocol was developed by the HEI Panel on the Health Effects of Long-Term Exposure to Traffic-Related Air Pollution, HEI scientific staff and the contractor team. The protocol will be published on the HEI website, and as an appendix to the HEI Special Report. In addition, the protocol will be published on PROSPERO.

2. RATIONALE AND OBJECTIVE

Traffic emissions are an important source of urban air pollution. The health burden associated with existing vehicle emissions levels remains significant in the United States and globally (Bhalla et al., 2014). Tailpipe emissions from motor vehicles and ambient concentrations of most monitored traffic–related pollutants have decreased steadily over the last several decades in most high–income countries. This trend is a result of air quality regulations and improvements in vehicular emission control technologies, and is likely to continue. However, decreases in emissions from individual motor vehicles do not fully compensate for the rapid growth and increased vehicular congestion of the motor vehicle fleet due to growth of the population and economic activity, as well as the presence of older or malfunctioning vehicles on the roads. In addition, interest in the contribution of non–tailpipe emissions to air quality and health is increasing in most–high income countries as vehicle miles traveled increase and regulations continue to be targeted almost exclusively to tailpipe emissions. Therefore, people continue to be exposed to TRAP, especially in urban settings and residences in proximity to busy roadways.

In 2010, HEI published Special Report 17, *Traffic–Related Air Pollution: A Critical Review of the Literature on Emissions, Exposure, and Health Effects*. This Report, developed by the HEI Panel on the Health Effects of Traffic–Related Air Pollution summarized and synthesized research on emissions, exposure, and health effects from TRAP and drew conclusions about whether the associations between exposure and health outcomes were causal. The Panel concluded that the evidence was 'sufficient' to support a causal relationship between exposure to TRAP and exacerbation of asthma. It also found 'suggestive' evidence of a causal relationship with onset of childhood asthma, non–asthma respiratory symptoms, impaired lung function, total and cardiovascular mortality, and cardiovascular morbidity, although the data were not sufficient to fully support causality. For a number of other health outcomes, there was limited evidence of associations, and the data were either 'inadequate' or 'insufficient' to draw firmer conclusions (HEI, 2010).

Since HEI published its review in 2010, many additional studies investigating the health effects of exposure to TRAP have been published and regulations and vehicular technology have advanced significantly. In addition, there is a better appreciation that, beyond air pollution, traffic can be a source of other exposures with potential relevance to health, most notably noise. These exposures may either confound or modify the health effect of TRAP. TRAP continues to be of public health interest and is of concern to policy makers and motor vehicle manufacturers alike. Therefore, HEI has decided to conduct a new literature review, as described in HEI's Strategic Plan 2015–2020 (HEI, 2015).

The overall objective is to systematically evaluate the epidemiological evidence regarding the associations between long–term exposure to TRAP and selected adverse health outcomes. The Panel will draw conclusions about the confidence in the strength of the evidence, discuss strengths and limitations of the existing studies, and make recommendations for future research. Results will be quantitatively combined to evaluate the strength of the evidence, where appropriate. In addition, the quantitative results of the review may be useful for future risk and health impact assessments of TRAP.

The current review differs from the earlier critical review in some important aspects: 1) It will follow a systematic approach using common methods and a published protocol; 2) it will evaluate the epidemiological literature only; 3) it will evaluate only studies of long–term exposure and health; 4) it will use a new exposure framework and will consider exposure contrasts beyond the near–road environment; 5) it will focus on a selected set of health outcomes chosen *a priori*, and 6) it will draw conclusions about the 'confidence' in the strength of the epidemiological evidence. The scope of the review was discussed extensively during several Panel meetings, and also took into account feasibility issues given the vast and rapidly growing literature on the potential adverse health effects of TRAP.

3. METHODS

3.1 Selection of health outcomes and prioritization

Five criteria were formulated by the Panel for selection of health outcomes in the review: 1) Previous strength of evidence; 2) Policy relevance; 3) Public health relevance; 4) Diversity of outcomes; and 5) Feasibility.

Table 1 lists the health outcomes that will be included in the systematic review (List A). Table 2 lists additional outcomes that may be evaluated if time and resources permit (List B). List A contains clinical outcomes, most of which have previously received 'causal' or 'likely to be causal' determinations for air pollution in general. The all–cause and cause–specific mortality outcomes are similar to the categories used in the 2017 Global Burden of Disease study of ambient air pollution (Cohen et al., 2017).

To assess the previous assessments of the strength of the evidence, HEI Special Report 17 was considered along with other authoritative reviews, such as the US EPA's Integrated Science Assessments of NO_2 (2016) and PM (2009), the International Agency for Research on Cancer (IARC) reviews on diesel (2014), and outdoor air pollution (2016), the Health Canada (HC) review of NO_2 (2016), the Global Burden of Disease project (Cohen et al., 2017), and the report from the UK–based Committee on the Medical Effects of Air Pollutants (COMEAP) on NO_2 and mortality (2018). Most outcomes on List B are outcomes that were previously not identified as having 'causal' or 'likely to be causal' determinations for air pollution in general, although the literature base is growing rapidly for some of these outcomes (e.g., neurological outcomes).

Related to policy and public health relevance, the Panel discussed the difference between clinical outcomes, subclinical outcomes useful for disease diagnosis, and other subclinical and physiologic outcomes primarily relevant to elucidating disease mechanisms, guided by the joint statement of the American Thoracic Society and the European Respiratory Society on this topic (Thurston et al., 2017). Preference was given to clinically relevant outcomes. As such, the selected health outcomes are either clinical outcomes (e.g., myocardial infarction or stroke in List A) or part of the definition of clinical outcomes (e.g., lung function or blood pressure in List B). Though not a clinical outcome or part of the definition, atherosclerosis was included in List B because it is the underlying pathology/disease process for some major clinical outcomes on List A (stroke and coronary events), remains subclinical for a long time, and adds to the diversity of outcomes.

Initial literature searches identified a large number of hits (~11,500), and it was considered unfeasible within the time and resources allotted to conduct a systematic review of all candidate outcomes. Hence, feasibility constraints led the Panel to prioritize the outcomes by categorizing them as high–priority (List A) lower priority (List B).

The level of detail in the HEI Special Report for outcomes on List B requires additional elaboration, but it is anticipated that high–level narrative summaries will be provided for some outcomes on List B, in particular for neurologic outcomes.

The ICD–10 codes listed in Tables 1 and 2 should be interpreted as guidance, and there will be some flexibility over exact diagnoses and codes as some studies use subsets or definitions based on their own assessments and other data. For all–cause mortality, preference is given to non–accidental (natural) mortality, and all–cause mortality will only be used if non–accidental mortality is not available.

For morbidity outcomes, both incidence and prevalence studies are included, where relevant. Note that some of the health outcomes represent a broad group of outcomes, and the selected health outcomes can be measured in various ways in different studies. Especially the definition of asthma is difficult and can be based on questionnaire data on asthma or asthma–like symptoms, hospitalization or emergency room visits for asthma, or medicine use. Further subcategories and specific outcomes will be identified in the course of the review and based on evidence and expert judgment of the Panel.

Table 1: Selected health effects of long–term exposure to traffic–related air pollution (List A: outcomes to be systematically reviewed)

Health Outcome Category	**Subcategory (ICD–10 codes from the WHO, version 2016, where applicable)**

All cause and cause–specific mortality	• Non–accidental mortality (A00–R99) or all–cause mortality (A00–Z99)
	• Respiratory mortality
	  ○ All respiratory mortality (J00–J99)
	  ○ Chronic Obstructive Pulmonary Disease (COPD) (J44)
	  ○ Acute Lower Respiratory Infections (ALRI) (J12–J18, J20–J22)
	• Circulatory mortality
	  ○ All circulatory mortality (I00–I99)
	  ○ Ischemic Heart Disease (IHD) (I20–I25)
	  ○ Stroke (I60–I69)
	• Lung cancer mortality (C33–C34)

Respiratory effects (both in children and adults)	• Asthma occurrence (J45–J46) • Asthma severity (asthma exacerbation) (J45–J46) • COPD occurrence (J44) • COPD severity (J44) • ALRI (J12–J18, J20–J22)
Cardiovascular effects including diabetes	• Coronary events such as fatal and non–fatal myocardial infarction (I21) and cardiac arrest (I46) • Stroke (I60–I69) • Type 2 diabetes (E11)
Birth outcomes	• Low birth weight (<2500 g) (P07.0–P07.1) and small for gestational age • Preterm birth (<37th week) (P07.2–P07.3)

Table 2: Selected health effects of long–term exposure to traffic–related air pollution (List B: lower–priority outcomes)

Health Outcome Category	**Subcategory (ICD–10 codes from the WHO, version 2016, where applicable)**
All cause and cause–specific mortality	• Diabetes mortality (E11)
Respiratory effects	• Lung function (FVC, FEV_1, FEF25–75)

Cardiovascular effects	• Heart failure (I50)
	• Atherosclerosis (I70)
	  ○ Intima–media thickness (IMT)
	  ○ Ankle–brachial index (ABI)
	  ○ Calcification: coronary artery calcium (CAC), or aorta calcification (AC), calcifications of left–sided heart valves
	• Blood pressure and hypertension (I10–I15)

Cancer incidence	• Childhood leukemia (C91–C95)
Pregnancy outcomes	• Gestational diabetes (O24) • Blood pressure and gestational hypertension (O13) • Preeclampsia / Eclampsia / Hemolysis Elevated Liver Enzyme Low Platelet Count (HELLP) syndrome (O14–O15)
Neurodevelopment outcomes (children)	• Cognitive function (e.g., general, verbal, and non–verbal IQ, language, memory, learning, visuospatial skills, attention, executive function) • Attention deficit hyperactivity disorder (ADHD) (F90) and related behaviors (e.g., inattention, impulse control, hyperactivity) • Autism spectrum disorders (ASD) (F84.0) and related behaviors (e.g., social cognition)
Neurocognitive outcomes (adults)	• Cognition or cognitive function in different cognitive domains • Cognitive decline in different cognitive domains • Dementia (includes Alzheimer's disease, vascular dementia, all–cause dementia) (F00–F03, G30), cognitive impairment and mild cognitive impairment (MCI) • Parkinson's disease (G20)

3.2 Development of PECOS questions and eligibility criteria

A PECOS framework (Population, Exposure, Comparator, Outcome and Study) was used to develop our review question (Higgins and Green, 2011). The formulation of an adequate PECOS question is a crucial step and will form the basis of the search for the evidence, and must therefore be framed in a way that enables systematic retrieval of the relevant literature that responds to the public health questions of interest.

The following PECOS question has been developed in relation to exposure to TRAP:

> "In the general population, including subgroups of adults and children (P), what is the increase in risk of health effect x (O) for a change (C) in long–term exposure to traffic–related air pollution (E), observed in studies relevant for the health outcome and exposure duration of interest (S)?

Table 3 presents inclusion and exclusion criteria for each PECOS domain in relation to the selected health effects of long–term exposure to TRAP.

The focus of the review is on the general population, except for the two severity outcomes (i.e., studies conducted in asthmatic and COPD patients). However, it was considered important to also be able to answer the question whether the health effects of TRAP are more pronounced for subgroups than in the general population. To this end, the Panel decided to include studies in selected patient populations, specifically patients with IHD, stroke, diabetes, heart failure, and hypertension, but only for all–cause and cause–specific mortality.

Long–term exposure was defined as a duration of months to years, similar to the definition of the World Health Organization (WHO) Air Quality Guidelines update (e.g., Vilahur et al., 2017).

Table 3: Inclusion and exclusion criteria for each PECOS domain in relation to the selected health effects of long–term exposure to traffic–related air pollution

PECOS	Inclusion	Exclusion
Population	General human population, of all ages, developed and developing areas, both urban and rural. No geographical restrictions. Asthma and COPD patients for outcomes on severity and recurrence of symptoms. Selected patient populations, specifically with IHD, stroke, diabetes, heart failure, and hypertension, but only for all–cause and cause–specific mortality.	Populations exposed in occupational settings or exclusively indoors.

Table 3: continued

PECOS	Inclusion	Exclusion
Exposure	Long–term exposure (months to years) to TRAP. Indirect measures of TRAP, such as distance to or length of roadways or traffic density at nearest road. Include studies regardless of whether they adjust for co–pollutant exposures. See section 3.3 Exposure framework for additional inclusion criteria.	Short–term exposure studies (minutes to months).
Comparator	Exposure to lower levels of TRAP in the same or in a referent population.	
Outcome	See Table 1 for the health outcomes selected (List A).	
Study	Human studies include cohort studies, case–cohort, case–control, cross–sectional studies, and intervention studies. Only human studies that are published (or accepted for publication i.e. in press) between January 1980 and July 2019, in peer–reviewed journal articles and written in English. Studies that report a quantitative measure of association and a measure of precision.	Qualitative studies, and studies reporting only unadjusted results. Studies without individual level data i.e. fully ecological outcome, exposure and covariates data. Studies where no original data were analysed, or methodological papers. Genome–wide association study (GWAS) and all other –omics studies. Non–human studies (*in vivo, in vitro,* other), and controlled exposure (chamber) studies. Grey literature, conference abstracts, conference papers, notes, editorials, letters and unpublished data.

3.3 Exposure framework

To guide selection and evaluation of epidemiological studies on TRAP, we developed a new framework for assessing the potential of different exposure assessment approaches used in epidemiological studies to be indicative of exposure to TRAP. That is, that the exposure signal or contrast driving the epidemiological associations reported is expected to be mainly due to TRAP. This framework will be used to identify studies in the systematic literature search that will be evaluated in the review and in the quantitative analysis. The framework builds on the 2010 HEI Report, and justification is provided for deviations. Chapter 3 (Exposure) and section 4.1 of Chapter 4 (Epidemiology) of the 2010 HEI Report are the key chapters in this respect. The exposure framework was tested on a selection of identified studies and adapted to clarify the selection process.

3.3.1 Exposure framework in the 2010 HEI Report

Traffic emissions affect air pollution concentrations and their variations at global, regional, urban and local scales (HEI, 2010: Figure 3.1). Regional scale refers to a large area of a country, urban scale refers to differences between urban and more rural areas and differences between neighborhoods of very large cities, and local scale reflects street–level differences between traffic density of major and minor roads or distance to major roads. In Chapter 3 (HEI, 2010: Table 3.2) a distinction is made among regional (100–1000 km), urban (4–50 km), neighborhood (50 m–4 km), and household (<50 m) scales. In the 2010 HEI Report, local is interpreted as less than 500m from a highway or a major road, and is often called the 'near–road' environment.

The 2010 HEI Report highlighted as a major issue that no pollutant is specific to traffic sources. Other sources contribute to commonly–used traffic–related pollutants such as EC, NO_2 and UFP. The 2010 HEI Panel decided to focus on primary (freshly emitted) pollutants on the local scale. Both long–term and short–term exposures were assessed. The choice of studies was based on exposure considerations; no *a priori* selection was made in terms of health outcomes at that time. The 2010 HEI Panel selected studies that used one of the following exposure assessment methods (HEI, 2010: Table 4.1):

1. Measures based on proximity to roadways or length of roads
2. Measures of traffic density
3. Modelling (dispersion models of traffic; other techniques such as LUR of traffic; traffic–specific source apportionment)
4. Participants in occupations characterized by exposure to traffic
5. Pollutant surrogates of traffic exposure such as NO_2, CO, and EC. This is interpreted as monitoring, with the requirement of road–side monitoring or participants living in short distances from the monitors.

Studies that used self–reported exposure to TRAP were specifically excluded.

Chapter 3 of the 2010 HEI Report provides a detailed discussion of surrogate metrics, modelling techniques and different pollutants used as indicators for exposure to TRAP, focusing extensively on exposure conditions in the United States. For the evaluated pollutants an important consideration was the fraction of the total emissions from motor vehicles. This was a somewhat controversial aspect of the 2010 HEI Report because it required assessment of how much of the spatial concentration contrast in a specific setting was due to different sources. Even $PM_{2.5}$ contrasts within a single city may be largely due to traffic sources. As an example, Figure 1 shows that most $PM_{2.5}$ in a study in Amsterdam, the Netherlands, reflected regional background air pollution, but the variation correlated well with UFP, which in Amsterdam is mainly traffic–related (Hoek et al., 2011).

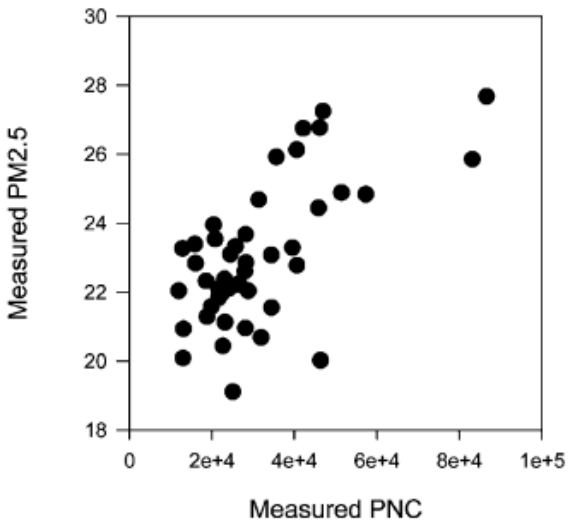

Figure 1: Relationship between measured PNC and $PM_{2.5}$ concentrations (Hoek et al., 2011).

The result of these strict exposure considerations was that most of the included studies used indirect measures of TRAP, such as proximity to roadways or traffic density (that are criticized because of validity and confounding issues). In addition, very few studies of short–term exposure were included in the review (e.g., 4 for all–cause mortality) and only 1 UFP study were included in the 2010 HEI Report. The numbers of time–series studies in general and short–term UFP studies specifically were limited by the road–side monitoring requirement. Most time–series studies select background monitors to represent daily variation of air pollution for larger areas or average all available stations including background and traffic sites.

Particularly for the studies based on monitoring, there is a dilemma regarding which studies to include. Limiting the review to those studies that unequivocally separate traffic from non–traffic sources even for traffic–related air pollutants would likely result in a small selection of studies and the exclusion of many studies that clearly provide evidence relevant to the health effects of TRAP.

3.3.2 Exposure framework for the current review

Traffic emissions may result in human exposure to TRAP at the local, urban and regional scales, and contribute to subsequent health effects. The highest exposure to TRAP is likely to occur at the local scale, that is, when in close proximity to traffic (walking, cycling, or being in a vehicle) or living or working close to major roads in urban environments. However, a more complete assessment of the health effects of TRAP requires consideration of exposure contrasts in TRAP more broadly within cities. In the European ESCAPE study the differences in NO_2 concentrations were similar comparing 1) urban background relative to nearby rural background and 2) street locations relative to urban background (Eeftens et al., 2012). In addition, as people spend time at locations other than their residence – which is often used to estimate individual spatially refined exposures – the broader neighborhood–scale exposure contrasts within an urban area are important in addition to the local contrasts. However, across broader spatial scales (urban and regional scale), traffic emissions are likely to be well–mixed with emissions from other sources, be lost to deposition, and undergo chemical and/or physical transformations to form secondary air pollutants which are difficult to link back to a traffic origin. Hence the focus of this review is on studies of primary pollutants derived from traffic, similar to the HEI 2010 Report.

We follow the 2010 HEI Report in recognizing that a major challenge for epidemiological research and for our objective of selecting and evaluating studies remains – i.e., that no pollutant is specific to traffic sources. Other sources contribute to commonly–used indicators of TRAP, such as EC, NO_2 and UFP. Therefore, use of commonly–accepted indicators of TRAP would ideally be evaluated in the context of the major drivers of exposure contrast in the geographic region under study and the specific design of the epidemiological study. However, given that a study–by–study evaluation is not feasible, we developed three guiding principles that when applied together give us reasonable confidence that exposure contrasts in a specific study were mainly related to traffic emissions:

1. Definition of traffic–related air pollutants (Table 4). We developed List A of air pollutants that are commonly considered to be related to traffic, though none of them has traffic as the only source. In general, studies of health effects of any of these pollutants are eligible for inclusion in this review, whether the exposure assessment is conducted by modeling or monitoring. However, $PM_{2.5}$, PM_{10} and PM_{coarse} have major sources other than traffic. PM studies are therefore only included if they used modeling approaches and PM monitoring studies are excluded, because it is unlikely that a major fraction of the measured PM exposure contrast is due to traffic sources. Also, studies that used self–reported exposure to TRAP are excluded, similar to the HEI 2010 Report.

2. Definition of scale of exposure contrasts related to specific study design (Table 5). The scale of the exploited exposure contrast affects how specifically an epidemiological study may reflect traffic impacts. Thus, scale refers to the area across which exposure levels are compared. We slightly modify the definition of the scales from the 2010 HEI Report and instead use: regional (>50 km), urban (5 km to 50 km), neighborhood (1 km to 5 km) and local scale (<1 km). Local is interpreted as less than 1 km from a major road, acknowledging that this is a large distance applicable only to major freeways, and the zone of elevated TRAP concentrations is typically within 100 m to 500 m from major roads. We will include studies that resolve TRAP variations within the local and neighborhood (within–city) scales, but not studies that were conducted at urban (i.e.,

primarily exploiting between–city contrasts in exposures) or regional scales. Thus, for example, studies that use city–wide average air pollution levels to all study participants are excluded because of a lack of specificity for TRAP. Within–city spatial variation in the exposure assessment is required.

3. Definition of specific exposure assessment methods and spatial resolution (Table 6).

Exposure assessment methods

A number of exposure assessment methods have been used in epidemiological studies of the health effects of TRAP and are considered in this review, including exposure assessments based on proximity to major roadways, traffic density, or length of nearby roads; dispersion or chemical transport models (CTMs); traffic–specific source apportionment; land use regression (LUR); surface monitoring; satellite monitoring; and personal exposure monitoring or modeling. Each exposure assessment approach has relative strengths and limitations, and these will be described in detail in the HEI Special Report.

To evaluate the traffic source contribution in studies using monitoring data is challenging. Most of those studies use regulatory sites (often measured at a location representative of 'background' levels) that monitor ambient air pollution concentrations and interpolate measured concentrations to estimate exposures to study participants. The sphere of influence of neighborhood background monitoring stations is typically about 3 – 5 km, dependent on the setting. As mentioned earlier, studies that use city averages (either monitored or modeled) to all study participants are excluded for all pollutants because of lack of traffic specificity, and within–city spatial variation in the exposure assessment is required. In addition, all PM monitoring studies will be excluded, because it is unlikely that a major fraction of the measured PM exposure contrast is due to traffic emissions. For the other pollutants in Table 4, a judgement will be made whether traffic emissions were a major contributor to the measured exposure contrast.

In addition to spatial resolution, the various methods used for exposure assessment have intrinsic differences to consider. For example, studies using measures based on proximity to roadways, traffic density, and length of roads are thought to be very specific for traffic and will be included in this review, provided they meet the resolution requirements listed in the table. All dispersion and chemical transport models applied specifically to resolve spatial patterns of TRAP will be included, provided they meet the resolution requirements (equivalent of 5x5 km grids or higher). Studies that use dispersion and chemical transport models that predict air pollutant concentrations of all traffic and non–traffic sources combined (e.g., total $PM_{2.5}$ or NO_2 from mobile, area and point sources) will require judgment regarding the study area (extent, present sources), modelled sources and other considerations.

LUR modeling methods generally predict air pollutant concentrations from all sources, similar to all–source dispersion and CTM models. Thus, judgement regarding their applicability will need to be made, taking into account the extent of the study area, presence of other sources and which sources are represented by predictive variables in the models. Models that contain at least one traffic predictor (e.g., traffic intensity or road density) or broader surrogate of traffic (e.g., address

density, household density, population density, impervious surface) will be included. We will thus exclude LUR models that include only non–traffic sources.

Spatial resolution

The spatial resolution of the exposure assessment method is important to judge how well a specific method characterizes TRAP for the population being studied. Studies will be considered potentially eligible for conclusion if the spatial resolution of the air pollution surface was finer than 5 km, referring to surfaces of 5x5 km grids or smaller; the upper limit of what is considered the neighborhood scale. Information will be added during the course of the review on the typical size of zip codes tabulation areas, zip codes, census tracts and census blocks in cities and suburban and rural areas in the US, Canada and other countries. For measures based on roadway proximity, length of road, or traffic density, the largest acceptable distance was set at 1 km, the upper limit of comparisons at local scales. We will use the 1 km criterion to select studies and may consider more stringent criteria distinguishing the type of major road (traffic intensity) in the evaluation stage and data synthesis.

The second key aspect of spatial resolution of the exposure assessment is the resolution of the location assignment for participants in the epidemiological study. For studies based on roadway proximity, length of road, or traffic density, fine spatial resolution (<100 m) of participant address geolocation is required for inclusion in the review. For other exposure assessment methods, participant address geolocation will be considered sufficient if the exposure assignment was at 5 km resolution or finer and the spatial resolution of the exposure model and address used in the epidemiology analysis to be matched (using appropriate spatial techniques). While exact participant locations would be ideal, exact addresses of study participants are not always used either due to lack of availability or in order to protect privacy. For example, only 4 or 5 position zip code can typically be obtained for participants in the Netherlands because of privacy reasons (i.e., by having the spatial element contain a sufficiently large number of people, in this case ~10,000, identification of the participants is much less likely.), a spatially resolved pollution map would need to be aggregated to the 4 or 5 position zip code level. We include studies based on residential location, school location or work location, as all these environments represent a considerable amount of potential exposure to TRAP.

For monitoring studies, the equivalent spatial 'resolution' was defined based on distances between monitors and study participants. For inclusion in the review, we require that the majority of the population analyzed lives within 5 km of a monitor; this may mean that in some instances we will use only a subgroup analysis, if provided. The choice to allow 5 km distance between the residence and the monitor is consistent with the definition of neighborhood scale and accepted spatial resolutions for modeling surfaces. If no information about the distance to monitors is available, we require that the average distance between sites be less than 10 km or that the density of sites be more than 1 site per 50 km^2.

Note that all spatial resolutions are provided as general guidance and need to be interpreted in conjunction with actual land use (i.e., the factors actually impacting the exposure contrasts) and spatial extent of the study area. We prefer to use absolute spatial criteria rather than using terms as 'address' or zip code, because the resolution varies across and within countries.

Combination of the guiding principles. For selection of the studies into the review, we will take an inclusive approach to the combination of these three guiding principles to select studies that are informative about health effects of TRAP (See Table 7 below). Therefore the contractor team will conduct initial screening of studies, and any studies where the inclusion decision is not clear will be brought to the exposure subgroup of the Panel to decide using the full text of the epidemiological study and the accompanying exposure paper, if needed.

Application of the exposure framework will probably be most challenging for nationwide or large regional or statewide (e.g., California) studies. In general, the larger the study area extent, the less likely a measured or modelled contrast in pollution for a given exposure assessment is mainly due to traffic emissions. For example if an LUR model with traffic and other predictors is applied in a nationwide epidemiological study, much of the exposure contrast is likely due to non–traffic air pollutant sources and will therefore not be included in the review unless the analysis approach taken was able to address this issue, for example by an adjustment for city or area in the health analysis. If the same LUR model is applied in a single metropolitan area, the model may be accepted. A nationwide epidemiological study based on a dispersion model of road traffic sources at high resolution may be included.

Because of the requirement for city or area adjustment in the health analysis for studies with large spatial extent, nationwide studies of several cities (e.g., MESA, a study of participants living across 6 US cities) are more likely to be included than nationwide studies with a very large number of communities precluding adjustment for city (e.g., Nurses' Health study). Statewide epidemiological studies (e.g., California) may require the same treatment as some states in the US are bigger than some countries in Europe. Authors often present area– or region–specific results (sometimes in a supplement) in addition to the overall nationwide results, which may be included in the review even if the main analyses do not meet the inclusion criteria.

Additional considerations related to how well exposure contrast in the included studies represents traffic emissions may be incorporated in the evaluation stage and data synthesis, and not in the study identification. For example, the Panel may want to weigh the three guiding principles together with other potential factors to categorize studies according to their 'traffic specificity' (e.g., 'High', 'Medium', and 'Low'). This 'traffic specificity' variable may be used for sensitivity analyses. We note that the 'traffic specificity' of a large number of PM LUR models will be likely 'Low'.

3.3.3 Application of the framework

The exposure framework will be used to support a transparent selection and evaluation of studies included in the review. It requires evidence that the spatial contrast in exposure is mainly related to traffic emissions, which depends on the design of the epidemiological study, the exposure assessment method and model resolution. Studies are included that resolve TRAP variations within the local and neighborhood (within–city) scales, but not studies that were conducted at urban (between–city) and regional scales. Additional considerations related to how well exposure contrast in the included studies represents participants' exposure to TRAP may be incorporated in the evaluation stage and data synthesis.

The ability to apply the exposure framework and obtain positive confirmation that exposures are mainly traffic–related is required to make the determination. Thus, if there is insufficient information in either the paper reporting the health analysis or an accompanying exposure paper, the study will be excluded. Sometimes the authors express in the paper that their exposure assessment approach only represents urban or regional scale variation and if so the paper will also be excluded.

In this review, results from single pollutant models will be used in the main analyses; results from multipollutant models will be explored in sensitivity analyses, where available. Hence there is a focus on single pollutant models in the exposure framework. The framework will be applied to all pollutants in a study. It is possible that one pollutant fulfills the criteria for inclusion whereas another pollutant does not in the same study. An example would be if one pollutant is assessed at a finer spatial resolution (e.g. NO_2) and another pollutant (e.g. $PM_{2.5}$) at a coarser spatial resolution (e.g., larger air pollutant surface grid size). Likewise, the applicability of metrics such as proximity to roadways will be assessed separately from the air pollutants because it is possible that the available address information is sufficient for the modelled pollutant but not sufficient for the proximity measure.

It is anticipated that the Special Report will have a separate section to lay out these exposure considerations in detail.

Table 4: Traffic–related air pollutants and exposure indicators included in review

Pollutant / exposure indicators	**Consideration**
NO_2, NO_x, NO	Frequently used in TRAP studies; NAAQS / limit values.
CO	Frequently used particularly in earlier TRAP studies; NAAQS / limit values
$PM_{2.5}$, PM_{10}, and PM_{coarse}	Frequently used in TRAP studies; NAAQS / limit values. All PM monitoring studies are excluded, because it is unlikely that a major fraction of the measured PM exposure contrast is due to traffic emissions.
Non–tailpipe PM trace metals from wearing of brakes and tires or from the resuspension of road dust, such as Ba, Cu, Fe and Zn	Increased interest because of reduction of tailpipe emissions.
UFP, PNC, quasi–ultrafine, different particle modes (nucleation, Aitken, accumulation), particle size distribution	Component of fine particles with distinct exposure patterns that often reflect traffic emissions.
EC, BC, BS, PM absorption ('soot')	Frequently used in epidemiological studies.
PAH	Added for completeness; Some PAHs are increased by traffic emissions, though PAHs are generally not specific to traffic.
Benzene	Added for completeness; Some VOCs are increased by traffic emissions, though VOCs are generally not specific to traffic emissions. Benzene chosen as a marker for mobile source air toxics.
Measures based on distance, length of roads, or traffic density	Added for completeness. Very specific, but concerns about validity and no pollutant estimates available.

Table 5: Traffic-related air pollution impacts at different scales

Scale (area of impact)[a]	**Within scope of review**	**Rationale**
Increase in regional scale (> 50 km), average background concentration of secondary pollutants including O_3, nitrates (part of $PM_{2.5}$).	No.	Other sources than traffic contribute to contrast that cannot be reliably separated.
Increase in regional scale (>50 km) average background concentration of traffic–related pollutants as listed in Table 4.	No, when this is the only source of spatial contrast. An example is a study using county–level pollution as the exposure metric or a study evaluating only rural communities, where the contrast is derived completely from difference in regional background. Will be a rare study design given that we exclude geographical (ecological/correlation) studies.	Other sources than traffic contribute to contrast that cannot be reliably separated.
Increase in neighborhood (1–5 km) and urban scale (5–50 km) average background concentration of traffic–related pollutants as listed in Table 4. This category also includes nationwide epidemiological studies that evaluate contrast as the sum of regional and urban or neighborhood background.	Yes. For the assessment we will classify studies based on contrast: a. Exclusively between city b. Within city and between city c. Within city only Studies that exclusively use between–city contrast (e.g., original Six Cities Study that estimate one value to all participants in a city) will not be included because of lack of traffic specificity. Thus studies that estimate city averages to all participants are excluded. An exception may be a study in a small non–industrial region (e.g. <100 km) where the dominant contrast is between cities and smaller towns. Studies will be included if the epidemiological study or accompanying exposure paper satisfactorily documents that the contrast between study locations/cities has an important traffic source contribution and is not dominated by other sources (e.g., industry, wood smoke). This is particularly important for studies that include a between city component.	All three types of contrast studies may contain a traffic signal, though the certainty of attributing contrast to traffic differs. Category '*a*' has the most uncertainty in whether the contrast is related to traffic sources, and is therefore not considered. For studies in category '*b*', if traffic is a documented important source and an adjustment is made for city or area in the epidemiological analysis, the study is more likely to show a traffic signal than if no adjustment is made.

Table 5: continued

Scale (area of impact)[a]	**Within scope of review**	**Rationale**
Increase in local scale (<1 km) average background concentration of traffic–related pollutants as listed in Table 4.	Yes. Studies will be included, if the contrast between study locations has an important traffic source contribution.	
Increase in commuting exposures for all traffic–related pollutants as listed in Table 4.	Yes, but likely few studies on long–term exposure.	
Increase in occupational exposure (taxi drivers, postal delivery workers).	No.	Not considered useful in 2010 HEI Report. Difficult to combine with general environmental exposures.

[a] Scale refers to the area across which exposure levels are compared. Studies typically have multiple scales, we use the smallest scale to categorize a study.

Table 6: Exposure assessment methods eligible for inclusion in the review

Exposure metric	**Considerations**	**Spatial resolution "pollution surface"[a]**	**Spatial resolution "address" (health) data[a]**
Measures based on distance	Specific marker for local scale. Limited validity. As continuous distance (preferably non–linear) or distance categories.	<1000 m away from a highway or a major road	<100 m
Measures of traffic density or length of roads	Specific marker for local scale. Limited validity. Typically buffers or traffic intensity of nearest street.	Buffers with radius of <1000 m around address	<100 m
Dispersion models or chemical transport models (CTM)	May cover local, neighborhood and urban scale. If a dispersion model or CTM of traffic is used, the study will be included, provided it meets the scale and resolution requirements. Studies that use dispersion or CTM models that predict air pollutant concentrations of all sources combined (e.g., total $PM_{2.5}$ or NO_2 from mobile, area and point sources) require judgment regarding the study area and modelled sources.	<5 km	<5 km
Traffic–specific source apportionment	Specific to the degree that source apportionment is successful. May cover local, neighborhood and urban scale. Difficult to apply spatially resolved.	< 5 km	<5 km
Land use regression (LUR). Includes hybrid models with CTM and/or satellites; universal kriging; Bayesian methods; models by machine learning techniques.	Studies that use LUR require judgment including study area and predictors in the model. For inclusion we require at least one traffic predictor (traffic intensity or road density) or broader surrogate of traffic (e.g., address density, household density, population density, impervious surface).	< 5 km	<5 km

Table 6: continued

Exposure metric	**Considerations**	**Spatial resolution "pollution surface"[1]**	**Spatial resolution "address" (health) data[1]**
Surface monitoring. Involves exposure assignment by interpolation, including nearest neighbor, Thiessen polygon, inverse distance weighing and kriging without covariates (e.g., ordinary kriging)	Not fully specific for traffic. Main issue is exploited spatial scale of exposure contrast of study. PM studies are excluded.	< 5 km	<5 km
Satellite monitoring	Less specific for traffic than surface monitoring if used directly. If satellite monitoring is combined with other approaches (e.g., hybrid model), the overall specificity may be sufficient. PM studies are excluded.	< 5 km	<5 km
Personal exposure monitoring or modeling (time weighted average of micro–environment exposures)	Unlikely to be applied in long–term studies. Separation of indoor and outdoor sources needed. PM studies are excluded.	Not applicable	Not applicable

[a] All spatial resolutions are provided as general guidance and need to be interpreted in conjunction with actual land use (i.e., the factors actually impacting the exposure contrasts) and spatial extent of the study area. We prefer to use absolute spatial criteria rather than terms such as 'address' or zip code, because the resolution varies across and within countries.

Table 7: Combination of criteria for all accepted combinations

Pollutant	**Exposure assessment methods**	**Spatial resolution "pollution surface"**	**Spatial resolution "address"**	**Spatial resolution of "address" for study identification**	**Traffic contribution major? Other considerations?**
All pollutants from Table 4	Dispersion or CTM models of traffic or traffic–specific source–apportionment	<5 km	<5 km	Residential address as exact address, census block, census block group, neighborhood, census tract, zip code acceptable (city or county not)	Assumed by method
All pollutants from Table 4	Dispersion or CTM models of all sources	<5 km	<5 km	Residential address as exact address, census block, census block group, neighborhood, census tract, zip code acceptable (city or county not)	Judgement needed (e.g., spatial extent of the study area, other sources such as coal burning absent)
All pollutants from Table 4	LUR. Models that contain at least one traffic predictor (e.g., traffic intensity or road density) or broader surrogates of traffic (e.g., address density, household density, population density, impervious surface)	<5 km	<5 km	Residential address as exact address, census block, census block group, neighborhood, census tract, zip code acceptable (city or county not)	Judgement needed (e.g., spatial extent of the study area, other sources such as coal burning absent)
$PM_{2.5}$, PM_{10}, and PM_{coarse}	Surface, satellite and personal monitoring	excluded	excluded	excluded	excluded
	Measures based on distance, length of roads, or traffic density	<1000 m from a highway or a major road	<100 m	Residential address as exact address or detailed zip code (street segment)	Assumed by method

3.4 Information sources and search strategy

Studies matching the PECOS questions will be searched comprehensively in the PubMed electronic database. Initial literature searches revealed that the addition of a second electronic database, Web of Science, added very few relevant papers to the PubMed search but added a large number of hits to screen. Thus the Panel decided to restrict the search to PubMed. In addition to PubMed, studies included in the LUDOK database will be checked as well The LUDOK database is developed and maintained by the contractor team and provides a rich compilation of air pollution and health studies. In addition, references included in other systematic reviews will be considered for inclusion. To further ensure that relevant published studies not captured through the search are included in the review, we will check with the individual bibliographic databases curated by HEI and Panel members. We will also check against the selected studies included in the HEI 2010 Special Report.

The search strategy has been developed by the contractor team, borrowing from other reviews as much as possible, in particular the OHAT systematic review on TRAP and selected health outcomes (OHAT, 2016). See Appendix 2 for the search strategy (dated February 12, 2019). To ensure that the appropriate studies are identified, the search strategy has been designed to capture studies considering all the inclusion criteria derived from the PECOS questions. The search strategy has been rigorously tested against a large set of reviews relevant to the topic.

Literature searches will include studies from January 1980 through the end of July 2019. The HEI 2010 review included studies published between January 1980 and October 2008; the current review will include studies published during the same period and the following 10 years.

3.5 Data management and selection process

Two reviewers from the contractor team will independently screen titles and abstracts of the search results to determine whether a reference meets the inclusion criteria. Disagreements between screeners will be resolved through discussion, involving one or more additional members of the contractor team or HEI staff and Panel members, as necessary.

After completion of the title/abstract screen, full–text articles will be retrieved for those studies that either clearly meet the inclusion criteria or for which it is not possible to make a clear assessment from the bibliographic information and abstract alone. For those studies, full–text review will be independently conducted by two reviewers from the contractor team and HEI staff. Disagreements will be resolved by discussion involving one or more additional members of the contractor team, HEI staff and Panel members. The reason for exclusion at the full–text–review stage will be annotated and reported in a flow chart in the HEI Special Report. In particular, the exposure will be screened in detail at the full text stage using a structured form.

DistillerSR, a web–based, systematic review software program with structured forms and procedures will be used to screen articles for relevance and eligibility to ensure standardization of process.

3.6. Data extraction

For each included study, relevant data will be extracted using a structured form, which will be built in DistillerSR. The extracted data will be used to help summarize study designs and findings, facilitate assessment of risk of bias and to conduct statistical analyses during data synthesis. Data extraction will be performed by one reviewer from the contractor team and checked by a second reviewer (a member of the contractor team, HEI staff or Panel) for completeness and accuracy. To formally assess accuracy, the contractor team will select a random sample of included studies to re–enter the effect estimates.

To obtain information about the exposure assessment, we will make use of information reported in the cited exposure papers, if needed.

No attempts will be made to contact authors of included studies to obtain missing data considered important for evaluating key study findings (e.g., data required to conduct a meta–analysis).

3.7 Risk of bias in individual studies

The optimal methods for assessing risk of bias (RoB) in the context of observational studies in environmental health is currently an active area of research with no clear consensus. There are different RoB assessment tools available, for example ROBINS–I, designed for non–randomized studies of interventions (Sterne et al., 2016), the NIEHS Office of Health Assessment and Translation Risk of Bias assessment (OHAT, 2015), the Newcastle – Ottawa Quality Assessment scale for assessing the quality of non–randomized studies in meta–analysis (Wells et al., 2019), and the RoB assessment that is used for the updated WHO Air Quality Guidelines (AQGs) (Vilahur et al., 2017). All assessments have their own strengths and limitations, and there is no validated instrument yet for observational studies (Bero et al., 2018).

The Panel decided to use the RoB Tool and Guidance from the WHO AQG review because it is designed for assessment of risk of bias in observational air pollution epidemiology studies (Vilahur et al., 2017). The WHO RoB tool was finalized in February 2019.

In brief, the WHO RoB tool guides judgement of each study across 6 domains (in no particular order): 1) confounding; 2) selection bias; 3) exposure assessment; 4) outcome measurement; 5) missing data and 6) selective reporting. Most domains have subdomains. Each subdomain and an overall rating per domain is derived using three categories ('Low'/'Moderate'/'High'). For each domain and subdomain the WHO provided guidance for making a judgment about whether the study presents 'Low', 'Moderate', or 'High' RoB. The RoB instrument mandates that a rationale is provided for each judgement. No summary classification will be derived across the domains at the study level.

To come to an overall judgment for a domain the WHO formulated the following rules: if any of the subdomains has a rating of high risk of bias, the whole domain will be rated as high risk of bias; if all the subdomains have a rating of low risk of bias, the whole domain will be rated as low risk of bias;

when at least one subdomain has a rating of moderate risk of bias and none of the other subdomains is at high risk of bias, the whole domain will be rated as moderate risk of bias.

The tool will be slightly adapted based upon Panel members' experiences in applying the tool in the systematic reviews of the WHO AQG. For example, the distinction between critical and potential confounders will be removed and we will make use of a small list of critical confounders only. The methods and changes made to the WHO RoB tool will be documented and discussed in the HEI Special Report.

The RoB assessment will be conducted for each exposure–outcome pair. In other words, if a study reports on two relevant exposure–outcome associations, the RoB will be applied separately for each exposure–outcome pair. One Panel member will assess the risk of bias. The assessments will be checked by HEI staff and other Panel members for completeness, accuracy and consistency. Disagreements will be resolved through discussions, and involving one or more additional members of the Panel and HEI staff. Panel members will test and calibrate the RoB tool on a limited number of studies to ensure comparability across assessors.

Sensitivity analyses will be performed per RoB domain across studies, grouping studies at higher RoB for that domain and studies at lower RoB for that domain, provided there is a sufficient number of studies. Similar to WHO, the RoB assessment will be conducted only for exposure–outcome associations to be included in subsequent meta–analyses. In addition, the RoB assessments will be informative in the overall evaluation of the epidemiological evidence.

3.8. Data synthesis

Data synthesis will consist of three parts: (1) description of evidence base; (2) quantitative analyses of study results; and (3) meta–analysis where appropriate.

(1) Description of evidence base

This will be undertaken at the publication and cohort or study population level. We will provide an overview of the numbers of publications available for the outcomes and exposures included in the review. We will stratify publications by cut–off date of the 2010 HEI Report to indicate growth in publications since the last review. We will also stratify publications by region of the world to indicate the geographical spread of the evidence base.

We will also provide an overview of the numbers of independent studies available for the exposure–outcome associations to be included in meta–analyses.

(2) Quantitative analyses of study results

We will provide forest plots of the evidence available for the evaluation of the strength of evidence. We will produce individual plots by exposure–outcome pairs.

The plots will include the point estimate, 95% confidence interval, and study descriptors (e.g., location, year of publication, study population) but no meta–analysis. Results selected for subsequent meta–analysis will be indicated in the plots. Where exposure has been categorized (e.g.,

proximity to roadway studies) the exposure categories will be reported in the plots. Otherwise risk estimates for continuous exposures will be standardized to a common pollutant increment. A narrative description of the evidence base will accompany the plots.

(3) Quantitative analyses of study results using meta–analysis techniques

Results will be quantitatively combined to evaluate the strength of the evidence, where appropriate. In addition, the quantitative results of the review may be useful for future risk and health impact assessments of TRAP. Please see Table 8 for exclusion, inclusion and selection criteria for meta–analysis. Additional criteria may be developed during the course of the review.

Relative risks (RRs) will be used as the common effect measure of association across studies, and hazard ratios (HRs) may be considered equivalent to RRs. If odds ratios (ORs) are reported in the study and the outcome prevalence is lower than 10%, they will be considered equivalent. Estimates for continuous outcomes will be standardized to a common exposure increment.

In case three or more studies are identified for the same pollutant and health outcome, a meta–analysis will be performed.

Results will be quantitatively combined using random effects (RE) models (DerSimonian et al., 1986; Veroniki et al., 2016). We will report summary estimate, 95% confidence intervals, Chi^2, tau^2, I^2 statistics and 95% prediction intervals. We will use the statistical program R for the analyses and plots.

Table 8 Exclusion, inclusion and selection criteria for meta–analysis

Exclusion criteria
(i) no measure of precision of the point estimate (standard error or confidence interval) (ii) insufficient information available to standardize estimates & precision (e.g., not reported, pollutant increment not clear) (iii) clear evidence of an analytical error
Inclusion criteria
(i) general population studies and studies in selected population sub–groups e.g. truck drivers (not occupational exposure); Californian teachers study, Adventist Health Study of Smog, Nurses' Health study)) will be considered as population based studies. In addition, selected patient populations, specifically with IHD, stroke, diabetes, heart failure, and hypertension will be included, but only for all–cause and cause–specific mortality. (ii) single pollutant model result. In addition to single–pollutant results, we will extract effect estimates corrected for co–pollutants (general PM2.5 and ozone) as well as estimates corrected for traffic noise, where available for sensitivity analyses. (iii) ability to standardize the results (see text below).

Table 8: continued

Selection criteria
(i) when the same study population is used in several publications on the same exposure-outcome, select on basis of the following order: • largest population sample size, number of events or number of cases • most appropriate adjustment for confounders • most recent publication date (ii) exclude cohorts/studies that are already included in other multi–cohort analyses. Cohorts included in summary analysis such as ESCAPE have more standardized methods and focus that is optimal in order to detect non–design related heterogeneity. However, there may be some exceptions, for example if the exposure assessment is more specific to traffic. (iii) for (i), exclude sub–group analyses of larger cohorts or populations, even if the subgroup analyses are reported in the most recent publications.

3.8.1 Standardization of estimates

We will conduct separate meta–analyses for the individual pollutants included in the review (standardization of pollutant). We might explore ways to convert different combustion–related PM indicators such as EC, BC, BS, Soot, into a common index, for example consider all as EC –equivalent estimates (Cyrys et al., 2003; Janssen et al., 2011). However, we are not planning to convert NO_x to NO_2 or PM_{10} to $PM_{2.5}$.

Effect estimates for pollutants expressed by ppb or ppm will be converted to $\mu g/m^3$, using standard WHO scaling factors (standardization of units). For example, 1 ppb NO_2 = 1.88 $\mu g/m^3$, assuming an ambient pressure of 1 atmosphere and a temperature of 25 degrees Celsius.

In addition, effects will need to be expressed using a standardized increment in exposure, assuming a linear exposure–outcome relationship (standardization of pollutant–specific increments. The Panel will explore the possibility of including studies that use transformed pollutant concentrations in the meta–analyses. The analyses will focus on studies that have analyzed the exposure as a continuous variable. Studies using categories of exposure, such as quartiles of exposures, high versus low, or categories of distance from major road, road length, or traffic intensity are excluded from the meta–analyses.

3.8.2 Heterogeneity

The primary aim of the heterogeneity assessment is to inform the evaluation of consistency of a given exposure–outcome association across subgroups of studies or populations, which is one of the factors listed in the overall evaluation of the epidemiological evidence (see section 3.10). An exhaustive exploration of all sources of heterogeneity is beyond the scope of the review. We have identified subgroups of interest for potential sensitivity analyses, provided there is a sufficient number of studies:

- general population versus selected patient subgroups (only for mortality outcomes)
- time period (e.g., recent exposures versus past exposures)
- geographical areas (e.g., North America vs. Europe vs. Asia)
- higher RoB versus lower RoB per domain of the RoB tool
- confounder adjustment for individual–level behavioral factors (e.g., smoking, alcohol use, physical activity, diet)
- adjustment for traffic noise (selected outcomes only)
- adjustment for co–pollutants (general $PM_{2.5}$ and O_3)
- mean age of the study participants
- 'traffic specificity' (high versus lower)

Some of the analyses listed above require additional considerations:

(1) Traffic–related air pollution is a complex mixture

Because TRAP is a complex mixture, a key question that remains largely unresolved is whether a given traffic pollutant is a 'causal' agent or only an indicator of TRAP given that correlation in space and time between individual traffic pollutants is often high. For each exposure–outcome pair in the review, results from single pollutant models will be used in the main analyses; in sensitivity analyses results from multipollutant models will be explored. Specifically, effect estimates corrected for general $PM_{2.5}$ mass and O_3 will be explored because typically correlation with a traffic–related pollutant is low to moderate for those two pollutants, allowing a meaningful interpretation.

The difficulty in interpreting regression coefficients for correlated variables in multipollutant (regression) models is well documented (Dominici et al., 2010). These difficulties include: (1) correlation between pollutants (arising from common sources and meteorological conditions) can lead to unstable effect estimation; (2) different degrees of measurement error across pollutants can lead to the 'transfer' of an association from the less well measured (but true) pollutant to the better measured (but incorrect) pollutant; and (3) interactions between pollutants are often not evaluated, but such evaluations are required to correctly interpret the pollutants' estimated effects.

Those issues will be discussed in detail in the HEI Special Report.

(2) Traffic noise

Since the 2010 HEI Report, there is now a better appreciation that motor vehicles also contribute to other potentially harmful exposures in addition to TRAP – most notably traffic noise. Traffic noise may either confound or modify the health effects of TRAP. Yet, relatively few studies have sought to

quantitatively disentangle the possible effects of TRAP and traffic noise (e.g., Münzel et al., 2017; Stansfeld 2015; Tetrault et al., 2013; Tzivian et al., 2015). In the review, the influence of correction for traffic noise will be explored in sensitivity analyses.

(3) Changes in composition of traffic–related air pollution over time or country

Emissions from motor vehicles and ambient concentrations of most monitored traffic–related pollutants have decreased steadily over the last several decades in most high–income countries as a result of air quality regulations and related improvements in vehicular emission control technologies. Therefore, in the review we will explore how study results vary depending on the time period (recent exposure versus past exposures) in sensitivity analyses.

Those changes might differ across countries because both fleet composition and regulatory emission limits are different across countries. For example, diesel–fueled vehicles make up a larger portion of the vehicle fleet in Europe than in the United States. To address these possible influences on the composition of TRAP we will conduct sensitivity analyses of studies grouped by geographical area, as possible.

3.9 Publication bias

We will use funnel plots with Egger test on asymmetry to assess publication bias, if there is a sufficient number of studies (Egger et al., 1997). Those methods are recommended when at least 10 studies are included in the meta–analysis. However, even 10 studies may be low, because the results of the Egger test also depend on study size and magnitude of effects (Macaskill et al., 2001). Hence those methods will be applied with caution, since differences in study results may also reflect true heterogeneity (Lau et al., 2006). We will additionally apply the trim and fill method to determine whether potential publication bias produces a meaningful change in effect estimates (Duval and Tweedie, 2000).

3.10 Overall evaluation of the epidemiological evidence

The Panel will evaluate the strength of the body of evidence that TRAP is associated with specific health effects, and decided to do so using the strategy proposed by the Office of Health Assessment and Translation (OHAT) (OHAT, 2019; Rooney et al., 2014). The OHAT method is based on the Grading of Recommendations Assessment, Development and Evaluation (GRADE), which has been adopted by the Cochrane Collaboration, and many other organizations. OHAT refined the GRADE approach to include observational human studies in addition to randomized clinical trials. Moreover, OHAT applies the framework separately for animal and human data, which is relevant for the focus of the review on epidemiological studies. The Panel is aware that adaptations to the environmental health field are also underway within the GRADE Working Group – also for use in the update of the WHO AQG. This is critical because experience with GRADE in the environmental health context is as yet limited (Morgan et al., 2016).

The OHAT method uses four descriptors to indicate **the level of confidence in a body of evidence** (Table 9, Figure 2). More detailed guidance on reaching confidence ratings in the body of evidence as "high", "moderate", "low" or "very low" is provided in the OHAT Handbook for Conducting a Literature–Based Health Assessment (OHAT, 2019). In brief, for each outcome or group of related outcomes, studies are given an initial confidence rating that reflects the presence or absence of key study design features. The initial rating of each outcome group is downgraded for factors that decrease confidence and upgraded for factors that increase confidence in the results. Confidence across all studies with the same outcome is then assessed by considering the ratings for all groups of studies with that outcome, and conclusions are based on the highest rating for that outcome.

Factors that may decrease confidence in evidence:

- Risk of Bias
- Unexplained inconsistency
- Indirectness
- Imprecision
- Publication bias

Factors that may increase confidence in evidence:

- Larger magnitude of association
- Exposure–response
- Consideration of residual confounding
- Consistency
- Other factors

In OHAT, the four key study design features used to delineate the studies for initial confidence ratings are: (1) the exposure is experimentally controlled; (2) the exposure assessment demonstrates that exposures occurred prior to the development of the outcome (or concurrent with aggravation or amplification of an existing condition); (3) the outcome is assessed on the individual level (i.e., not through population aggregate data), and (4) an appropriate comparison group is included in the study. The first key feature, "controlled exposure", reflects the ability of experimental studies to largely eliminate confounding by randomizing allocation of exposure. Therefore, these studies usually have all four features and receive an initial rating of "high

confidence." By definition, observational studies do not have controlled exposure and are differentiated by the presence or absence of the three remaining study design features. For example, cohort studies usually have all three remaining features and receive an initial rating of "moderate confidence".

To translate confidence ratings into **level of evidence for health effects** in OHAT (Table 10), the nature of the association ("health effect" or "no health effect") is considered. Three descriptors ("high," "moderate," and "low" level of evidence) directly translate from the ratings of confidence in that the exposure is associated with a heath effect. If the confidence rating conclusion is "very low" or no evidence is identified, then the level–of–evidence conclusion is characterized as "inadequate evidence." The descriptor "evidence of no health effect" is used to indicate confidence that the exposure is not associated with a health effect. Because of the inherent difficulty in proving a negative, the conclusion "evidence of no health effect" is only reached when there is high confidence in the body of evidence.

The Panel recognizes that the scientific judgments involved in developing these ratings are inherently subjective. A key advantage of the evaluation is that it provides a framework to document and explain the decisions made, and thereby provide transparency into the scientific basis of judgments made in reaching conclusions. The Panel will use the OHAT method as a guide, and will not apply the methods in a mechanistic way; the Panel will not 'automatically' downgrade or upgrade based on the listed factors. Some factors are quite controversial, such as unexplained inconsistency because of the nature of observational studies in different populations, contexts, and exposure conditions. In addition, some factors may be considered more important than others (Risk of Bias versus imprecision).

The determinations will be based on consensus among members of the Panel, and considering the overall strengths and limitations of the available evidence. The Panel will evaluate consistency of the effects across pollutants and exposure indicators. For example, recent suggestions of integration ('triangulation') have been made within the epidemiology community (Lawlor et al., 2016). The underlying premise is that if different epidemiological approaches all point to the same conclusion, the confidence is strengthened. This seems particularly compelling when the key sources of bias of some of the approaches are predicted to influence estimates in opposite directions (Pearce et al., 2019). The method adopted by the Panel, including deviations from the OHAT method will be documented and discussed in the HEI Special Report.

Table 9: Confidence ratings in the body of evidence (OHAT, 2019; Rooney et al., 2014).

Confidence rating	Definition
High confidence (++++)	High confidence in the association between exposure to the substance and the outcome. The true effect is highly likely to be reflected in the apparent relationship.
Moderate confidence (+++)	Moderate confidence in the association between exposure to the substance and the outcome. The true effect may be reflected in the apparent relationship.
Low confidence (++)	Low confidence in the association between exposure to the substance and the outcome. The true effect may be different from the apparent relationship.
Very low confidence (+)	Very low confidence in the association between exposure to the substance and the outcome. The true effect is highly likely to be different from the apparent relationship.

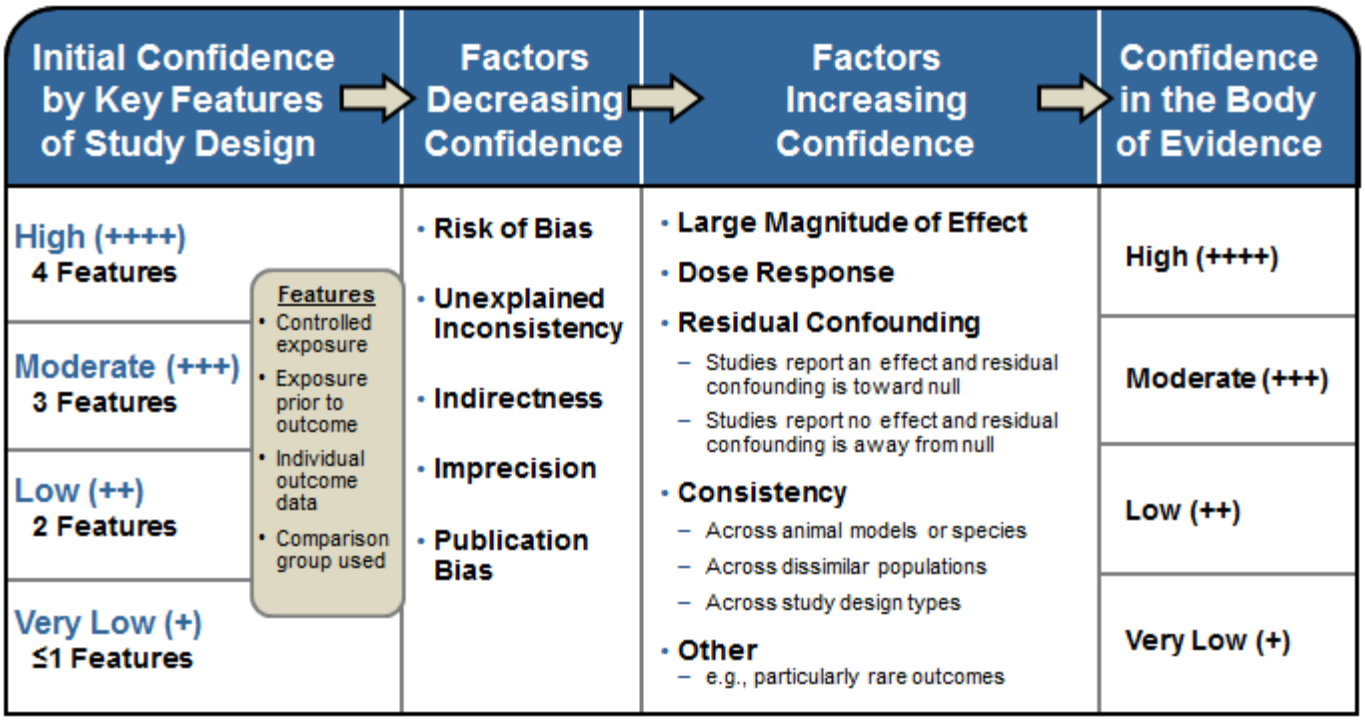

Figure 2: Assessing confidence in the body of evidence (OHAT, 2019)

Table 10: Level of evidence ratings for health effects (OHAT, 2019; Rooney et al., 2014).

Evidence descriptors	**Definition**
High level of evidence	There is high confidence in the body of evidence for an association between exposure to the substance and the health outcome(s).
Moderate level of evidence	There is moderate confidence in the body of evidence for an association between exposure to the substance and the health outcome(s).
Low level of evidence	There is low confidence in the body of evidence for an association between exposure to the substance and the health outcome(s), or no data are available.
Evidence of no health effect	There is high confidence in the body of evidence that exposure to the substance is not associated with the health outcome(s).
Inadequate evidence	There is insufficient evidence available to assess if the exposure to the substance is associated with the health outcome(s).

4. REPORTING

The reporting of the systematic review will comply with the Preferred Reporting Items for Systematic Reviews and Meta–Analyses (PRISMA) standards (Moher et al., 2015) with slight adaptations, because these standards were originally intended for health care intervention evaluation.

As such, the completed systematic review will include a clear formulation of the rationale and the objective of the review according to the protocol. The search strategy developed will be presented in an Appendix. In addition, a flow chart on studies included and excluded in every stage (from identification to screening, eligibility and inclusion) will be provided. Characteristics of included studies will be summarized in detail. The final review will describe the risk of bias assessment conducted for individual studies. Results from the main analyses, as well as results from sensitivity analyses will be presented in tables and figures. For the main analyses, there will be an assessment of the confidence in the body of evidence. Changes to the protocol will be specifically listed.

In addition to reporting the systematic review of the epidemiologic evidence, the HEI Special Report will include a separate exposure section to lay out the exposure criteria considerations in detail. Moreover, the HEI Special Report will also include a background section that contains: text addressing some other important issues related to emissions, exposure to TRAP and health effects; a high-level succinct review on the toxicological evidence of long-term exposure to TRAP; and summaries of some recent key short-term studies.

The level of detail in the HEI Special Report for outcomes on List B requires additional elaboration, but it is anticipated that high–level narrative summaries will be provided for some outcomes on List B, in particular for neurologic outcomes.

The HEI Special Report is expected to undergo peer review in 2020 and publication is aimed for summer 2021.

5. REFERENCES

Bero L, Chartres N, Diong J, Fabbri A, Ghersi D, Lam J, et al. The risk of bias in observational studies of exposures (ROBINS–E) tool: concerns arising from application to observational studies of exposures. Systematic Reviews 2018;7(1):242.

Bhalla K, Shotten M, Cohen A, Brauer M, Shahraz S, Burnett R, et al. 2014. Transport for health: the global burden of disease from motorized road transport (English). Washington, DC: World Bank Group.

Cohen AJ, Brauer M, Burnett R, Anderson HR, Frostad J, Estep K, et al. Estimates and 25–year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. Lancet 2017;389(10082):1907–18.

Committee on the Medical Effects of Air Pollutants (COMEAP). 2018. Associations of long–term average concentrations of nitrogen dioxide with mortality. Public Health England.

Cyrys J, Heinrich J, Hoek G, Meliefste K, Lewne M, Gehring U, et al. Comparison between different traffic related particle indicators: elemental carbon (EC), PM2.5 mass, and absorbance. J Expo Anal Environ Epidemiol 2003;13(2):134–43.

DerSimonian R, Laird N. Meta–analysis in clinical trials. Control Clin Trials 1986;7(3):177–88.

Dominici F, Peng RD, Barr CD, Bell ML. Protecting human health from air pollution: shifting from a single–pollutant to a multipollutant approach. Epidemiology 2010; 21(2): 187–94.

Duval S, Tweedie R. Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics 2000; 56(2): 455-63.

Eeftens M, Tsai M–Y, Ampe C, Anwander B, Beelen R, Bellander T, et al. Spatial variation of PM2.5, PM10, PM2.5 absorbance and PM–coarse concentrations between and within 20 European study areas and the relationship with NO2 – results of the ESCAPE project. Atmos Environ 2012;62:303–317.

Egger M, Davey Smith G, Schneider M, et al. Bias in meta–analysis detected by a simple, graphical test. BMJ 1997;315: 629–34.

Health Canada. 2016. Human health risk assessment for ambient nitrogen dioxide. Water and Air Quality Bureau. Health Canada, Ottawa, ON.

HEI Panel on the Health Effects of Traffic–Related Air Pollution. 2010. Traffic–Related Air Pollution: A Critical Review of the Literature on Emissions, Exposure, and Health Effects. HEI Special Report 17. Boston, MA:Health Effects Institute.

Health Effects Institute. 2015. HEI Strategic Plan for Understanding the Health Effects of Air Pollution 2015–2020. Boston, MA:Health Effects Institute.

Higgins JPT, Green S (editors). Cochrane handbook for systematic reviews of interventions version 5.1.0 [updated March 2011]. The Cochrane Collaboration; 2011. Available from www.handbook.cochrane.org.

Hoek G, Beelen R, Kos G, Dijkema M, van der Zee SC, Fischer PH, Brunekreef B. Land Use Regression Model for Ultrafine Particles in Amsterdam. Environ Sci Technol 2011;45:622–28.

International Agency for Research and Cancer (IARC). Monograph on the evaluation of carcinogenic risks to humans. Diesel and Gasoline Engine Exhausts and Some Nitroarenes. 2014. Volume 105.

International Agency for Research and Cancer (IARC). Monograph on the evaluation of carcinogenic risks to humans. Outdoor air pollution. 2016. Volume 109.

Janssen NAH, Hoek G, Simic–Lawson M, Fischer P, van Bree L, 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 Persp 2011;119:1691–9.

Lau J, Ionnidis JP, Olkin I. The case of the misleading funnel plot. BMJ 2006;333:597–600.

Lawlor DA, Tilling K, Davey Smith G. Triangulation in aetiological epidemiology. Int J Epidemiol 2016;45:1866–86.

Macaskill P,Walter SD, Irwig L. A comparison of methods to detect publication bias in meta–analysis. Stat Med 2001;20:641–54.

Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta–analysis protocols (PRISMA–P) 2015 statement. Systematic Reviews 2015;4:1.

Morgan RL, Thayer KA, Bero L, Bruce N, Falck–Ytter Y, Ghersi D, et al. GRADE: Assessing the quality of evidence in environmental and occupational health. Environ Int 2016:92–93:611–6.

Münzel T, Sorensen M, Gori T, Schmidt FP, Rao X, Brook J, et al. Environmental stressors and cardio–metabolic disease: Part I–epidemiologic evidence supporting a role for noise and air pollution and effects of mitigation strategies. Eur Heart J 2017;38:550–6.

Office of Health Assessment and Translation (OHAT). 2019. Handbook for Conducting a Literature–Based Health Assessment Using OHAT Approach for Systematic Review and Evidence Integration. Division of the National Toxicology Program. National Institute of Environmental Health Sciences.

Office of Health Assessment and Translation (OHAT). 2015. OHAT Risk of Bias Tool for human and Animal Studies. Division of the National Toxicology Program. National Institute of Environmental Health Sciences.

Office of Health Assessment and Translation (OHAT). 2016. Protocol for a systematic review of traffic–related air pollution and selected health outcomes. Division of the National Toxicology Program. National Institute of Environmental Health Sciences.

Pearce N, Vandenbroucke JP, Lawlor DA. Causal Inference in Environmental Epidemiology: Old and New Approaches. Epidemiology 2019;30:311–6.

Rooney AA, Boyles AL, Wolfe MS, Bucher JR, Thayer KA. Systematic review and evidence integration for literature–based environmental health science assessments. Environ Health Perspect 2014; 122:711–8.

Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta–analysis protocols (PRISMA–P) 2015: elaboration and explanation. BMJ 2015; 349:g7647.

Stansfeld SA. Noise effects on health in the context of air pollution exposure. Int J Environ Res Public Health 2015; 12:12735–60.

Sterne JA, Hernan MA, Reeves BC, Savovic J, Berkman ND, Viswanathan M, et al. Robins–I: A tool for assessing risk of bias in non–randomised studies of interventions. BMJ 2016; 355:i4919.

Tetreault LF, Perron S, Smargiassi A. Cardiovascular health, traffic–related air pollution and noise: Are associations mutually confounded? A systematic review. Int J Public Health 2013:58:649–66.

Thurston GD, Kipen H, Annesi–Maesano I, Balmes J, Brook RD, Cromar K, et al. A joint ERS/ATS policy statement: What constitutes an adverse health effect of air pollution? An analytical framework. Eur Resp J 2017:49:1600419.

Tzivian L, Winkler A, Dlugaj M, Schikowski T, Vossoughi M, Fuks K, et al. Effect of long–term outdoor air pollution and noise on cognitive and psychological functions in adults. Int J Hyg Environ Health 2015:218:1–11.

U.S. EPA. 2016. Integrated Science Assessment (ISA) for Oxides of Nitrogen – Health Criteria (Final Report, 2016). U.S. Environmental Protection Agency, Washington, DC, EPA/600/R–15/068, 2016.

U.S. EPA. 2009. Integrated Science Assessment (ISA) for Particulate Matter (Final Report, Dec 2009). U.S. Environmental Protection Agency, Washington, DC, EPA/600/R–08/139F, 2009.

Veroniki AA, Jackson D, Viechtbauer W, Bender R, Bowden J, Knapp G, et al. Methods to estimate the between–study variance and its uncertainty in meta–analysis. Res Synth Methods 2016;7:55–79.

Vilahur N, Héroux M, Román P, Verbeek J, Chen J, Hoek G. Long–term exposure to PM2.5 and PM10 and all–cause and cause–specific mortality: a systematic review and meta–analysis protocol – Update of WHO Global AQGs. August 2017.
https://www.crd.york.ac.uk/PROSPEROFILES/82577_PROTOCOL_20190211.pdf (accessed June 7, 2019)

Wells GA, Shea B, Connell DO, et al. The Newcastle–Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in metaanalyses.
http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp (accessed June 7, 2019).

6. APPENDICES

Appendix 1. HEI Panel on the Health Effects of Long–term Exposure to Traffic–Related Air Pollution

Francesco Forastiere, **Panel Co–Chair**
Visiting Professor, Science Policy & Epidemiology, Environmental Research Group, King's College London, United Kingdom

Frederick Lurmann, **Panel Co–Chair**
Manager of Exposure Assessment Studies. Sonoma Technology, Inc., Petaluma, California

Richard Atkinson
Reader in Epidemiology, Population Health Research Institute and MRC–PHE Centre for Environment and Health, St George's, University of London, United Kingdom

Jeffrey Brook
Assistant Professor, Occupational and Environmental Health Division, Dalla Lana School of Public Health, University of Toronto, Canada; HEI Research Committee Member

Howard Chang
Associate Professor, Department of Biostatistics and Bioinformatics Information, Rollins School of Public Health, Emory University, Atlanta, Georgia

Gerard Hoek
Associate Professor, Institute for Risk Assessment Sciences, Environmental Epidemiology, Utrecht University, Netherlands

Barbara Hoffmann
Professor of Environmental Epidemiology, Institute of Occupational, Social and Environmental Medicine, Heinrich Heine University of Düsseldorf, Germany; HEI Research Committee Member

Sharon Sagiv
Assistant Adjunct Professor of Epidemiology, Center for Environmental Research and Children's Health, Division of Epidemiology, University of California Berkeley School of Public Health

Audrey Smargiassi
Associate Professor, Department of Environmental and Occupational Health, School of Public Health, University of Montreal, Montreal, Quebec, Canada

Adam Szpiro
Associate Professor of Biostatistics, Department of Biostatics, University of Washington, Seattle, Washington

Danielle Vienneau
Senior Scientific Collaborator, Physical Hazards and Health, Swiss Tropical and Public Health Institute, University of Basel, Switzerland

Gregory Wellenius
Associate Professor, Department of Epidemiology, Brown University, Providence, Rhode Island

Jennifer Weuve
Associate Professor, Department of Epidemiology, Boston University School of Public Health, Boston, Massachusetts

Consultants to the Panel
Frank Kelly
Professor of Environmental Health and Director of the Environmental Research Group, Kings College London, London, United Kingdom; Member, HEI Review Committee Member

Evangelia Samoli
Assistant Professor of Epidemiology and Medical Statistics, Dept. of Hygiene, Epidemiology and Medical Statistics, School of Medicine, National and Kapodistrian University of Athens, Athens, Greece

Contractor team
Meltem Kutlar Joss, Ron Kappeler and students supervised by Professor Nino Künzli
Swiss Literature Database and Services on Health Effects of Ambient Air Pollution (LUDOK), Swiss Tropical and Public Health Institute, Basel, Switzerland

Appendix 2. Search strategy (February 12, 2019)

PECOS		PubMed Search Terms
Population		adult[tiab] OR adults[tiab] OR child[tiab] OR children[tiab] OR pupils[tiab] OR preschooler[tiab] OR preschoolers[tiab] OR student[tiab] OR students[tiab] OR adolescent[tiab] OR adolescents[tiab] OR infant[tiab] OR infants[tiab] OR toddler[tiab] OR toddlers[tiab] OR newborn[tiab] OR baby[tiab] OR babies[tiab] OR person[tiab] OR persons[tiab] OR human[tiab] OR humans[tiab] OR people[tiab] OR man[tiab] OR men[tiab] OR woman[tiab] OR women[tiab] OR elderly[tiab] OR boy[tiab] OR boys[tiab] OR girl[tiab] OR girls[tiab] OR patients[tiab] OR population[tiab] OR populations[tiab] OR survivor[tiab] OR survivors[tiab] OR spouse[tiab] OR spouses[tiab] OR wife[tiab] OR husband[tiab] OR smoker[tiab] OR smokers[tiab] OR resident[tiab] OR residents[tiab] OR veteran[tiab] OR mother[tiab] OR mothers[tiab] OR father[tiab] OR fathers[tiab] OR “population based”[tiab] OR “cohort”[tiab] OR (("persons"[Mesh] OR "humans"[Mesh]) NOT (animals[Mesh] NOT humans[Mesh]))
Exposure	General Terms to be combined with pollutants Different Pollutants to be combined with OR	("Environmental Exposure"[Mesh] OR "Environmental Pollution"[Mesh] OR "Air Pollutants"[Mesh] OR "Air Pollution"[Mesh] OR "air pollution"[tiab] OR "air pollutants"[tiab] OR "polluted atmosphere"[tiab] OR "atmospheric pollution"[tiab] OR "polluted air"[tiab] OR "ambient air"[tiab] OR "Inhalation Exposure/adverse effects"[Mesh] OR "Motor Vehicles"[Mesh] OR "Vehicle Emissions"[Mesh] OR "traffic–related"[tiab]) OR ((traffic OR transport) AND air)
	NO_x	((("Nitrogen Oxides"[Mesh] OR "Nitrogen dioxide"[tiab] OR "NO2"[tiab] OR "NO(2)"[tiab] OR "NOx"[tiab] OR "NO(x)"[tiab] OR "Nitrogen oxide"[tiab] OR "nitrogen oxides"[tiab])) OR "oxides of nitrogen"[tiab]
	CO	"Carbon Monoxide"[Mesh] OR "carbon monoxide"[tiab]
	Traffic PM	"Particulate Matter"[Mesh:NoExp] OR "Smog"[Mesh] OR “smog”[tiab] OR "Particle Size"[Mesh] OR "PM10"[tiab] OR PM2.5[tiab] OR PM10–2.5[tiab] OR PM2.5–10[tiab] OR PM1[tiab] OR “fine particulate”[tiab] OR "PM10"[tiab] OR "PM2.5"[tiab] OR "PM10–2.5"[tiab] OR "PM2.5–10"[tiab] OR "PM1"[tiab] OR "PM(10)"[tiab] OR "PM(2.5)"[tiab] OR "PM(10–2.5)"[tiab] OR "PM(2.5–10)"[tiab] OR "PM(1)"[tiab] OR

		"particulate matter"[tiab] OR "PMcoarse"[tiab] OR "PMcoarse"[tiab]
	Non–tailpipe emissions and metals	resuspended dust[tiab] OR re–suspended dust[tiab] OR road dust[tiab] OR brake dust[tiab] OR tire dust[tiab] OR tyre dust[Text Word] OR brake wear[tiab] OR tire wear[tiab] OR tyre wear[tiab] OR road wear[tiab] OR debris dust[tiab] OR fugitive dust[tiab] OR diffuse dust[tiab] OR wear dust[tiab] OR non–exhaust[tiab] OR source apportionment[tiab] OR windblown dust[tiab] OR non–tailpipe[tiab] OR mineral dust[tiab] (nickel[tiab] OR Ni[tiab] OR Copper[tiab] OR Cu[tiab] OR aluminium[tiab] OR aluminum[tiab] OR Al[tiab] OR zinc[tiab] OR Zn[tiab] OR barium[tiab] OR Ba[tiab] OR iron[tiab] OR Fe[tiab] OR copper[tiab] OR Cu[tiab] OR Antimon[tiab] OR Sb[tiab] OR Tinn[tiab] OR Sn[tiab] OR Zirconium[tiab] OR Zr[tiab] OR "trace metals"[tiab] AND ("Particulate Matter"[Mesh:NoExp] OR "Smog"[Mesh] OR “smog”[tiab] OR "Particle Size"[Mesh] OR "PM10"[tiab] OR PM2.5[tiab] OR PM10–2.5[tiab] OR PM2.5–10[tiab] OR PM1[tiab] OR “fine particulate”[tiab] OR "PM10"[tiab] OR "PM2.5"[tiab] OR "PM10–2.5"[tiab] OR "PM2.5–10"[tiab] OR "PM1"[tiab] OR "PM(10)"[tiab] OR "PM(2.5)"[tiab] OR "PM(10–2.5)"[tiab] OR "PM(2.5–10)"[tiab] OR "PM(1)"[tiab] OR "particulate matter"[tiab] OR "PMcoarse"[tiab] OR "PMcoarse"[tiab]))
	UFPs	“submicron“[tiab] OR “surface area“[tiab] OR “ultrafine“[tiab] OR “ultrafine particles“[tiab] OR “ultrafine particle“[tiab] OR “nano particle“[tiab] OR “nano particles“[tiab] OR “nanoparticles“[tiab] OR “nanoparticle“[tiab] OR PM0.1[tiab] OR “PM0.1“[tiab] OR “PM(0.1)“[tiab] OR PM0.25[tiab] OR “PM(0.25)“[tiab] OR “PM0.25“[tiab] OR “quasi–ultrafine“[tiab] OR “quasi ultrafine“[tiab] OR “PNC“[tiab] OR “accumulation mode“[tiab] OR “particle number“[tiab] OR "number of particles"[tiab] OR “aitken mode“[tiab]
	Soot/BC	"Soot"[Mesh] OR soot[tiab] OR "PM2.5 absorbance"[tiab] OR "PM2.5absorbance"[tiab] OR “PM2.5abs”[tiab] OR "black carbon"[tiab] OR "carbon black"[tiab] OR “organic carbon”[tiab] OR “elemental carbon”[tiab] OR “black smoke”[tiab]
	PAHs	"Polycyclic Aromatic Hydrocarbons"[Mesh:NoExp] OR “polycyclic aromatic hydrocarbons”[tiab] OR PAH[tiab] OR

		"PAH's"[tiab] OR PAHs[tiab] OR "benzo(a)pyrene"[tiab] OR benzopyrene[tiab]
	Benzene	"benzene"[Mesh] OR benzene[tiab] OR BTEX[tiab]
	Proxy measures for traffic including OHAT 2016 traffic terms	((((traffic[tiab]) NOT ("Accidents, Traffic"[Mesh] OR safety[tiab] OR accident[tiab] OR accidents[tiab] OR injur*[tiab] OR collision*[tiab] OR crash*[tiab])) OR "traffic intensity"[tiab] OR "traffic density"[tiab] OR "traffic load"[tiab] OR "traffic count"[tiab] OR "road length"[tiab] OR ((proximity[tiab] OR near[tiab] OR distance[tiab] OR nearest[tiab] OR next[tiab] OR close[tiab] OR closest[tiab]) AND (road*[tiab] OR highway*[tiab] OR freeway*[tiab] OR motorway*[tiab] OR interstate[tiab] OR expressway[tiab]))))) OR ((vehicle[tiab] OR vehicles[tiab] OR vehicular[tiab] OR auto[tiab] OR automobile[tiab] OR bus[tiab] OR buses[tiab] OR car[tiab] OR truck[tiab] OR trucker[tiab] OR trucks[tiab] OR engine[tiab] OR transport[tiab] OR traffic[tiab]) AND (emissions[tiab] OR exhaust[tiab] OR fume*[tiab]))
Comparator	Measures of effect	“risk”[Mesh] OR “risk”[tiab] OR “risks”[tiab] OR “incidence”[Mesh] OR “incidence”[tiab] OR “incident”[tiab] OR "Prevalence"[Mesh] OR “prevalence”[tiab] OR “prevalent”[tiab] OR "Risk Factors"[Mesh] OR "risk factor"[tiab] OR "Odds Ratio"[Mesh] OR "odds"[tiab] OR “onset”[tiab] OR “associated”[tiab] OR “association”[tiab] OR “cause”[tiab] OR “causes”[tiab] OR “caused”[tiab] OR “develop”[tiab] OR “developed”[tiab] OR “prevent”[tiab] OR “prevents”[tiab] OR “prevented”[tiab] OR “increase”[tiab] OR “increased”[tiab] OR “increases”[tiab] OR “effect”[tiab] OR “effects”[tiab] OR “affect”[tiab] OR “affects”[tiab] OR “affected”[tiab] OR “protective”[tiab] OR “protect”[tiab] OR “protected”[tiab] OR “harm”[tiab] OR “harms”[tiab] OR “harmed”[tiab] OR “harmful”[tiab] OR “hazard”[tiab] OR “hazardous”[tiab] OR "Proportional Hazards Models"[Mesh] OR "proportional hazard"[tiab]
Outcome	Mortality	("Mortality"[Mesh] OR "mortality"[MeSH Subheading] OR "Cardiovascular Diseases/mortality"[Mesh] OR "Myocardial Ischemia/mortality"[Mesh] OR "Respiratory Tract Diseases/mortality"[Mesh] OR "Respiratory Tract Infections/mortality"[Mesh] OR "Respiration Disorders/mortality"[Mesh] OR "Lung Neoplasms/mortality"[Mesh] OR "Pulmonary Disease, Chronic Obstructive/mortality"[Mesh]) OR (("cause–specific"[tiab] OR "all–cause"[tiab] OR "non–accidental"[tiab] OR "natural"[tiab] OR "natural–cause"[tiab] OR "cardiovascular"[tiab] OR "respiratory"[tiab] OR "cardiorespiratory"[tiab] OR "cardio

		respiratory"[tiab] OR "lung cancer"[tiab] OR "COPD"[tiab]) AND (mortality[tiab] OR death[tiab] OR "deadly"[tiab] OR died[tiab] OR fatal*[tiab] OR surviv*[tiab])) OR ("mortality"[tiab] OR "death"[tiab])
Outcome	Respiratory effects	"Pulmonary Ventilation"[Mesh] OR "Respiratory Function Tests"[Mesh] OR “spirometry”[tiab] OR "plethysmography"[tiab] OR “forced expiratory”[tiab] OR “FEV”[tiab] OR “FVC”[tiab] OR “FEF25–75”[tiab] OR “MEF”[tiab] OR “expiratory flow”[tiab] OR “expiration flow”[tiab] OR “small airway”[tiab] OR “impulse oscillometry”[tiab] OR “FOT”[tiab] OR “peripheral airway”[tiab] OR (("pulmonary"[tiab] OR "respiratory"[tiab] OR "lung"[tiab]) AND ("volume"[tiab] OR "function"[tiab] OR "ventilation"[tiab] OR "capacity"[tiab])) OR "Asthma"[Mesh] OR asthma[tiab] OR asthmatic[tiab] OR wheezing[tiab] OR wheeze[tiab] OR whistle[tiab] OR whistling[tiab] OR "bronchial hyperreactivity"[tiab] OR "Bronchial Hyperreactivity"[Mesh] OR "bronchial hyperresponsiveness"[tiab] OR "airway hyperresponsiveness"[tiab] OR ISAAC[tiab] OR "Respiratory Hypersensitivity/chemically induced"[Mesh] OR bronchiodilat*[tiab] OR "bronchial dilation"[tiab] OR "bronchial dilatation"[tiab] OR bronchioconstrict*[tiab] OR salbutamol*[tiab] OR "methacholine"[tiab] OR "mannitol"[tiab] OR "Breath Tests"[Mesh] OR “exhaled nitric oxide”[tiab] OR “FeNO”[tiab] OR “fractional exhaled NO”[tiab] OR "Acute lower respiratory infection"[tiab] OR "Acute lower respiratory tract infection"[tiab] OR "ALRI"[tiab] OR ("respiration tract"[tiab] AND "infection"[tiab]) OR "Pneumonia"[Mesh] OR "pneumonia"[tiab] OR "Bronchiolitis"[tiab] OR "Bronchitis"[Mesh] OR "Bronchitis"[tiab] OR "Pulmonary Disease, Chronic Obstructive"[Mesh] OR COPD[tiab] OR ((“chronic obstructive”[tiab]) AND (bronchitis[tiab] OR “bronchopulmonary disease”[tiab] OR “lung disorder”[tiab] OR “pulmonary disease”[tiab] OR “pulmonary disorder”[tiab] OR “respiratory disease”[tiab] OR disease[tiab])) OR "emphysema"[tiab] OR "chronic airway obstruction"[tiab] OR "chronic airflow obstruction"[tiab]

Outcome	Cardiovascular effects	(“cardiovascular”[Title/Abstract] OR “cardiorespiratory”[Title/Abstract] OR “cardio–respiratory”[Title/Abstract]) OR ("Myocardial Ischemia"[Mesh] OR ((myocardial[tiab] OR myocard[tiab] OR heart[tiab] OR cardiac[tiab] OR cardial[tiab] OR myocardium[tiab]) AND (infarct[tiab] OR infarction[tiab] OR attack[tiab] OR failure[tiab] OR disease[tiab])) OR "Heart Failure"[Mesh] OR “fatal MI”[tiab] OR “coronary event”[tiab] OR “coronary syndrome”[tiab] OR “coronary syndrom”[tiab] OR “cardiac death”[tiab] OR “revascularization”[tiab] OR “revascularisation”[tiab]) OR ("Stroke"[Mesh] OR "Stroke"[tiab] OR "acute cerebrovascular lesion"[tiab] OR "cerebral vasculopathy"[tiab] OR "brain attack"[tiab] OR "cerebral apoplexy"[tiab] OR "brain ischemic attack"[tiab] OR (("cerebrovascular"[tiab] OR "cerebro vascular"[tiab] OR cerebral[tiab]) AND (insufficiency[tiab] OR "accident"[tiab] OR arrest[tiab] OR "failure"[tiab] OR "injury"[tiab] OR "attack"[tiab]))) OR ("Arteriosclerosis"[Mesh] OR “atherosclerosis”[tiab] OR “arteriosclerosis”[tiab] OR “vascular sclerosis”[tiab] OR "Carotid Intima–Media Thickness"[Mesh] OR “CIMT”[tiab] OR "aorta wall thickness"[tiab] OR "aortic thickness"[tiab] OR "aortic wall thickness"[tiab] OR "arterial thickness"[tiab] OR "artery thickness"[tiab] OR "artery wall thickness"[tiab] OR "carotid intima media thickness"[tiab] OR "carotid intima–media thickness"[tiab] OR "carotid intimamedia thickness"[tiab] OR "intima–media thickness"[tiab] OR "intimal medial thickness"[tiab] OR "intimamedia thickness"[tiab]) OR "Ankle Brachial Index"[Mesh] OR “ankle–brachial index”[tiab] OR "ankle brachial pressure index"[tiab] OR "ankle brachial ratio"[tiab] OR "Pulse Wave Analysis"[Mesh] OR "pulse wave velocity"[tiab] OR "pulse wave analysis"[tiab] OR "augmentation pressure"[tiab] OR "augmentation index"[tiab] OR "vascular reactivity"[tiab] OR "vascular function"[tiab] OR "Vascular Stiffness"[Mesh] OR ((aorta[tiab] OR arterial[tiab] OR aortic[tiab] OR artery[tiab] OR vascular[tiab]) AND (stiffness[tiab] OR stiffening[tiab])) OR "Calcinosis"[Mesh] OR "artery calcification"[tiab] OR "aortic calcification"[tiab] OR ("Blood Pressure"[Mesh] OR “blood pressure”[tiab] OR “systolic pressure”[tiab] OR “diastolic pressure”[tiab] OR "Hypertension"[Mesh] OR “hypertension”[tiab] OR “intravascular pressure”[tiab] OR “vascular pressure”[tiab] OR

		“blood tension”[tiab] OR “normotension”[tiab] OR “hypertensive”[tiab]) OR ("Plaque, Atherosclerotic"[Mesh] OR "plaque area"[tiab] OR “atherosclerotic plaque”[tiab] OR “arteriosclerotic plaque”[tiab] OR "atheromatous plaque”[tiab] OR "intima plaque”[tiab]) "Diabetes Mellitus, Type 2"[Mesh] OR "diabetes"[tiab] OR "diabetic"[tiab] OR T2DM[tiab] OR "type 2 DM"[tiab] OR "fasting blood glucose"[tiab] OR "fasting glucose"[tiab] OR "glucose metabolism"[tiab] OR "glucose homeostasis"[tiab] OR Hba1c[tiab] OR IDDM[tiab] OR NIDDM[tiab] OR HOMA–IR[tiab] OR hyperglycemia[tiab]
Outcome	Childhood leukemia	(("Leukemia"[Mesh] OR "Leukemia"[tiab] OR "Leukaemia"[tiab] OR leucemia[tiab] OR leucaemia[tiab] OR "childhood cancer"[tiab] OR hemoblastoma[tiab]) AND ("Child"[Mesh] OR "Adolescent"[Mesh] OR "Young Adult"[Mesh] OR "Infant"[Mesh] OR "children"[tiab] OR "childhood"[tiab] OR child[tiab] OR preschooler[tiab] OR preschoolers[tiab] OR pupil[tiab] OR pupils[tiab] OR student[tiab] OR students[tiab] OR adolescent[tiab] OR adolescents[tiab] OR infant[tiab] OR infants[tiab] OR toddler[tiab] OR toddlers[tiab] OR newborn[tiab] OR newborns[tiab] OR baby[tiab] OR babies[tiab] OR boy[tiab] OR boys[tiab] OR girl[tiab] OR girls[tiab]))
Outcome	Birth outcomes	"Fetal Growth Retardation"[Mesh] OR "Birth Weight"[Mesh] OR "Infant, Low Birth Weight"[Mesh] OR "Premature Birth"[Mesh] OR “intrauterine growth restriction”[tiab] OR "Fetal Development"[Mesh] OR “fetal development”[tiab] OR “foetal development”[tiab] OR “intrauterine growth retardation”[tiab] OR "birth weight"[tiab] OR “small for gestational age”[tiab] OR “preterm birth”[tiab] OR “premature birth”[tiab] OR "birth outcome"[tiab] OR "pregnancy outcome"[tiab] OR “neonatal weight”[tiab] OR “newborn weight”[tiab] OR “fetal growth”[tiab] OR “foetal growth”[tiab] OR “foetus growth”[tiab] OR “fetus growth”[tiab] OR “foetal growth restriction”[tiab] OR “foetal growth retardation”[tiab] OR “in utero growth retardation”[tiab] OR “in utero growth restriction”[tiab] OR “congenital hypotrophy”[tiab] OR “prenatal growth retardation”[tiab] OR “prenatal growth restriction”[tiab] OR “retarded intrauterine growth”[tiab] OR “premature childbirth”[tiab] OR “premature birth”[tiab] OR “small for date”[tiab] OR “low birth weight”[tiab] OR (LBW[tiab] AND (infant[tiab] OR baby[tiab] OR newborn[tiab] OR child[tiab])) OR (premature[tiab] AND (infant[tiab] OR baby[tiab] OR

		newborn[tiab] OR child[tiab])) OR (“preterm”[tiab] AND (infant[tiab] OR baby[tiab] OR newborn[tiab] OR child[tiab]))
Outcome	Pregnancy outcomes	"Diabetes, Gestational"[Mesh] OR "Hypertension, Pregnancy–Induced"[Mesh] OR "Gestational Hypertension"[tiab] OR "pregnancy–induced hypertension"[tiab] OR (pregnan*[tiab] AND hypertens*[tiab]) OR pre–eclampsia[tiab] OR preeclampsia[tiab] OR (pregnan*[tiab] AND toxemia*[tiab])
Outcome	Neurodevelopment outcomes (children) and neurocognitive outcomes (adults)	"Cognition Disorders"[Mesh] OR cognition[tiab] OR cognitive[tiab] OR neurobehavio*[tiab] OR neuropsych*[tiab] OR "Mental Processes"[Mesh] OR memory[tiab] OR "mental recall"[tiab] OR (verbal[tiab] OR language[tiab] OR reading[tiab] AND (comprehension[tiab])) OR “language”[tiab] OR learning[tiab] OR perception[tiab] OR perceptual[tiab] OR neurodevelop*[tiab] OR intelligen*[tiab] OR intellect*[tiab] OR “IQ”[tiab] OR behavior[Mesh:NoExp] OR Child behavior[Mesh] OR Adolescent behavior[Mesh] OR Behavioral symptoms[Mesh] OR Spatial behavior[Mesh] OR executive function[tiab] OR “academic achievement”[tiab] OR “academic performance”[tiab] OR "Neurodevelopmental Disorders"[Mesh] OR attention[tiab] OR inattenti*[tiab] OR hyperactiv*[tiab] OR "impulsive behavior"[Mesh] OR impulsive[tiab] OR impulse–control[tiab] OR impulsivity[tiab] OR “response inhibition”[tiab] OR “inhibitory control”[tiab] OR “vigilance”[tiab] OR “social–behavior”[tiab] OR “social–behaviour”[tiab] OR “social skills”[tiab] OR aggression[tiab] OR aggressive[tiab] OR “ADDH”[tiab] OR “ADHS”[tiab] OR “ADHD”[tiab] OR “ADH”[tiab] OR "Autism Spectrum Disorder"[Mesh] OR autistic[tiab] OR autism[tiab] OR “Tic–disorder”[tiab] OR Asperger*[tiab] OR “communication–disorder*”[tiab] OR language[tiab] OR agraphia[tiab] OR dyslexi*[tiab] OR dyscalculia[tiab] OR speech[tiab] OR aphasia[tiab] OR echolalia[tiab] OR “stereotyp*”[tiab] OR “Pervasive Developmental Disorder”[tiab] OR “social cognition”[tiab] OR “social communication”[tiab] OR “social reciprocity”[tiab] OR “repetitive behavior*”[tiab] OR “repetitive behaviour”[tiab] OR “restricted interests”[tiab] OR “maladaptive behavior”[tiab] OR “maladaptive behaviour”[tiab] OR “adaptive behavior”[tiab] OR “behavioral regulation”[tiab] OR

		"Aging"[Mesh] OR "Cognitive Dysfunction"[Mesh] OR “dementia”[Mesh] OR dementia[tiab] OR alzheime*[tiab] OR neurotox*[tiab] OR “Neurodegenerative Diseases”[Mesh] OR neurodegenerat*[tiab] OR neurodisease*[tiab] OR Parkinson*[tiab] OR neuropsycholog*[tiab]
Study	Filters	NOT ((((((“shortterm"[ti] OR "short–term"[ti] OR “time series”[ti] OR time–series[ti]) AND (("shortterm"[ti] OR "short–term"[ti] OR “time series”[ti] OR time–series[ti]) NOT ("longterm"[tiab] OR "long term"[tiab] OR "medium term"[tiab] OR "intermediate term"[tiab] OR “chronic”[tiab])))))) OR ("Clinical Trial"[Publication Type] OR "Treatment Outcome"[MeSH] OR "Cross–Over Studies"[Mesh] OR "case cross over"[tiab])) OR ("Air Pollutants, Occupational"[Mesh] OR "Accidents, Traffic"[Mesh] OR "Protective Devices"[Mesh])) OR (mouse[Title/Abstract] OR mice[Title/Abstract] OR rat[Title/Abstract] OR rats[Title/Abstract]) AND English[Language] AND ("1980/01/01"[Date – Publication] : "3000"[Date – Publication])