Project number: 003777 7205 UKAGEP
Report number: [trägt die UBA-Bibliothek ein]

Health Effects of Ultrafine Particles

Systematic literature search and the potential transferability of the results to the German setting

by

Simone Ohlwein MPH, Prof. Barbara Hoffmann MD et MPH
Institute for Occupational, Social and Environmental Medicine, University hospital Düsseldorf

Ron Kappeler MSc Med. Sci. Techn., Meltem Kutlar Joss MSc ETH Environmental Sc. et MPH, Prof. Nino Künzli MD et PhD
Swiss Tropical and Public Health Institute, Switzerland

Institute for Occupational, Social and Environmental Medicine.
Postal address: Universitätsklinikum Düsseldorf | AG Umweltepidemiologie |
PO box 101007; 40001 Düsseldorf
Visitor´s address: Gurlittstr. 55/II; 40223 Düsseldorf

On behalf of the German Environmental Agency and the Swiss Federal Office for the Environment

Financial support was provided by the Federal Ministry of Education and Research (Grant 01KX1460)

Completion date: February 2018

Abstract

Background/Aim: Due to their small size, the specific health effects of UFPs are related to their physical capacity to penetrate through the blood system, the nervous system, the brain and diverse organs. Five years ago scientific evidence pointed towards adverse effects of UFPs on health. Since then, numerous studies have been published. Therefore, the aims of this project were to review the literature on the effects of UFPs on health, to evaluate the selected studies and to assess the transferability of the results to the situation in Germany.

Methods: We systematically searched MEDLINE (Medical Literature Analysis and Retrieval System Online) for eligible studies published between 01.01.2011 until 11.5.2017 investigating health effects of ambient air pollution (AAP) related UFPs. In addition, we searched the LUDOK (Dokumentationsstelle Luftverschmutzung und Gesundheit)-database, provided by the Swiss Tropical and Public Health institute. We included epidemiologic studies with adequate study designs, containing an UFP measure, quantifiable measures of associations and a health outcome and extracted the relevant data based on previously elaborated evaluation criteria.

Results: Upon application of our search strategy, 85 references of original articles were identified for further evaluation. Most of included studies were conducted in North America (n=37) or Western Europe (n=27), investigating short-term effects (n=75). The short-term studies are dominated by panel studies (n=32), scripted exposure studies (n=16), and time-series studies (n=11). Ten studies investigated long-term associations using exposure estimates averaged over a period of months to years. Long-term studies most frequently applied cohort (n=4) and cross-sectional (n=4) study designs.

Conclusion: The evidence on health effects remains inconclusive or insufficient for most of the studied health outcomes. Specifically, while a number of studies have investigated mortality and emergency department/hospital admission outcomes, the relatively few studies with co-pollutant adjustment reveal mixed and, up to now, inconclusive evidence. In terms of number of studies, most evidence is available from studies investigating subclinical outcomes. Within this group of studies, cardiovascular outcomes and outcomes of pulmonary and systemic inflammation show the most consistent patterns with associations generally pointing into the direction of the adverse health outcome. A future challenge is the development of enhanced spatiotemporal models which can contribute to a more precise exposure assessment across larger areas as well as incorporating multi-pollutant models to become clear of independent effects.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

8-OHdG	8-hydroxy-2' –deoxyguanosine (biomarker of oxidative stress)
AccMP	Accumulation mode particles
AitMP	Aitken-mode particles
AAP	Ambient Air Pollution
BAFU	Schweizerisches Bundesamt für Umwelt
BC	Black Carbon
BDNF	brain-derieved neurotropic factor
BP	Blood Pressure
BREATHE	Brain Development and Air Pollution Ultrafine Particles in School Children
CAFEH	eine Kohorte, finde aber nichts im Internet (im Text auf Seite 53 oben)
CO	Carbon monoxide
COPD	Chronic obstructive pulmonary disease
CRP	C reactive protein
CTM	Chemistry-Transport-Model
CV	Cardiovascular
DBP	Diastolic Blood Pressure
DBP	diastolic blood pressure
DOI	Digital Object Identifier
EBC	Exhaled breath condensate
EC	Elemental Carbon
EHP	Environmental Health Perspectives
ETH	Eidgenössische Technische Hochschule Zürich
EURAD	EURopean Air Pollution Dispersion
FEF	Forced expiratory flow
FeNO	fractional exhaled nitric oxide
FEV_1	Forced Expiratory Volume in 1 second
FVC	Forced vital capacity
GUAN	German Ultrafine Aerosol Network
h	hour
HDL	High-Density-Lipoprotein
HEI	Health Effects Institute
HNR	Heinz Nixdorf Recall Study

HR	Heart Rate
HRAPIE	Health risks of air pollution in Europe
HRV	heart rate variability
hs-CRP	high sensitivity c-reactive protein
ICAM-1	Intercellular Adhesion Molecule 1
ICD	International Classification of Diseases
IGT	Impaired glucose tolerance
ijerph	International Journal of Environmental Research and Public Health
IL	Interleukin
IQR	Interquartile range
ISA	Integrated Science Assessment
LBW	low birth weight
LDSA	Lung-deposited surface area
LUDOK	Dokumentationsstelle Luftverschmutzung und Gesundheit
ma	moving averages
MDA	malondialdehyde
MEDLINE	Medical Literature Analysis and Retrieval System Online
MI	Myocardial Infarction
NANOAPP	Nanomaterials & Applications
NIH	National Heart, Lung and Blood Institute of the National Institute of Health
nm	nucleation mode
NO_x	Nitrogen oxides
NO_2	Nitrogen dioxide
NOS	Newcastle-Ottawa Scale der Universitäten Newcastle, Australien und Ottawa, Cana-da
NO_x	nitrogen oxide
NSTEMI	Non-ST-Elevated Myocardial Infarction
NucMP	Nucleation mode particles
O_3	Ozone
OR	Odds Ratio
PAC	Particle Area Concentrations
PM	Particulate Matter
PNC	Particle Number Concentrations
PP	Pulse Pressure

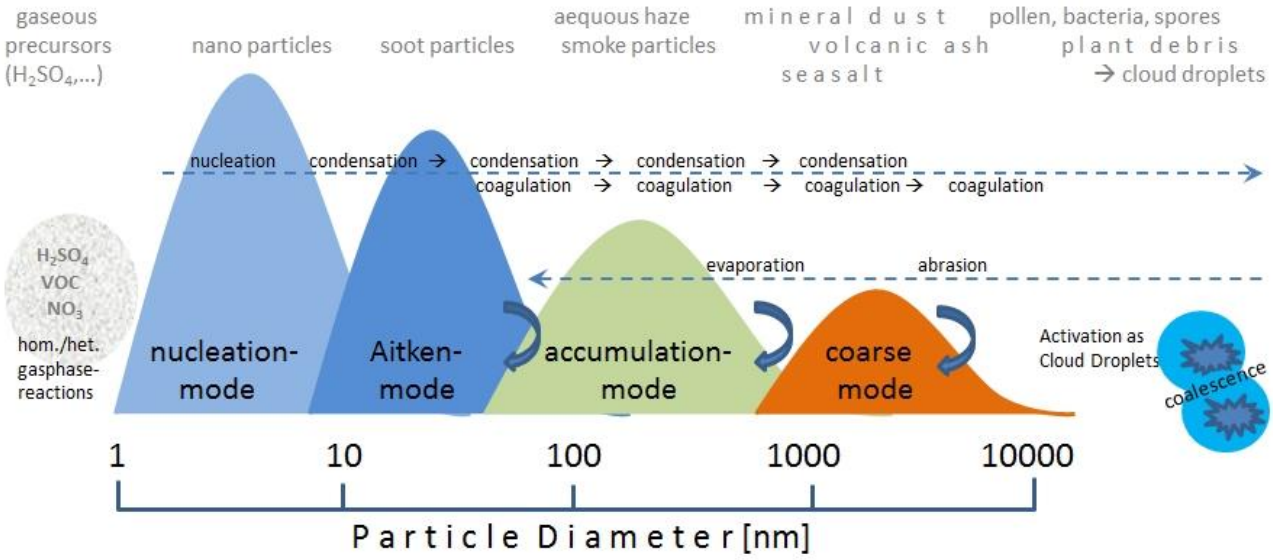
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PVC	Particle volume concentrations
REVIHAAP	Review of evidence on health aspects of air pollution
RHI	Reactive Hyperemia Index
RUPIOH	Relationship between Ultrafine and fine Particulate matter in Indoor and Outdoor air and respiratory Health
SAPALDIA	Swiss study on Air Pollution And Lung Disease in Adults
SBP	Systolic Blood Pressure
SO_2	Sulphur Dioxide
SOA	Secondary Organic Aerosols
STEMI	ST-elevation myocardial infarction
Swiss TPH	Schweizerische Tropen- und Public Health-Institut
T2DM	Type 2 Diabetes Mellitus
TNFRII	Tumour Necrosis Factor Type II
TNR	Gene that encodes the protein Tenascin-R
TU	Technische Universität
UFIPOLNET	Ultrafine particle size distributions in air pollution monitoring networks
UFIREG	Ultrafine Particles - an evidence based contribution to the development of regional and European environmental and health policy
UFPs	Ultrafine Particles
UKD	Universitätsklinikum Düsseldorf
US EPA	United States Environmental Protection Agency
VCAM	Vascular Cell Adhesion Molecule
VOCs	Volatile Organic Compounds
VT	Ventricular Tachycardia
WHO	World Health Organization
WP	Work package

Summary

Background

Ultrafine particles (UFPs) represent the smallest size fractions of air pollutants measured from a nanometer to few micrometers. By convention, UFPs are defined as particles not exceeding an aerodynamic diameter of 100 nm. Measurement procedures mostly assess particle number per ml since UFPs contribute only little to the particle mass of ambient air. Further size fractions used in epidemiological research are nucleation mode particles (precursor substances sized up to 20 nm), Aitken-mode particles (condensation particles sized 10 - 80 nm), and accumulation mode particles (condensation and coagulation particles sized 50 - 1,000 nm) covering different particle fractions (Figure 1).

Figure 1: Size-fractions of airborne particles (Deutscher Wetterdienst, 2018)



UFPs vary with regard to their chemical composition and physical reactivity. They are emitted directly or are formed from precursors in atmospherical processes. In urban areas, a great proportion of UFPs originate from combustion processes of motorized vehicles (Health Effects Insitute, 2013; Kelly et al., 2012).

The specific health effects of UFPs are related to their physical capacity to penetrate through diverse organ systems (i.e., blood system, nervous system, brain, organs) due to their small size. Hypothesized health effects of UFP include cardiovascular and respiratory morbidity and mortality, the elicitation of local pulmonary and systemic inflammation and oxidative stress, and adverse actions on the brain and the metabolism (Health Effects Insitute, 2013; Rückerl et al., 2011).

In contrast to other air pollutants, there are no regulations on UFP exposure concentrations. The expert commission of the Health Effects Institute (HEI) and the World Health Organization (WHO concluded five years ago that scientific studies point towards adverse effects of UFPs on health. However, the evidence base of epidemiologic studies was not sufficient to recommend regulations on UFP exposure concentrations.

At first, the HEI provides the most thorough and complete information on a possible relationship between UFPs and various health effects. The body of research was rated as suggestive but not definitive on the adverse health effects of UFPs on respiratory and cardiovascular outcomes. Reasons for the lack of clarity were (1) inconsistencies of outcomes and methodological aspects of the study designs, (2) inconsistent and possibly biased exposure assessments and (3) a lack of studies adjusting for co-pollutants. On top of those issues, HEI couldn't find any studies on long-term exposure effects of UFPs. Therefore the evidence base in 2013 on epidemiologic studies was not sufficient to recommend regulations on UFP exposure concentrations.

In February 2015 the United States Environmental Protection Agency invited experts from around the world to discuss and present evidence of health effects associated with UFP exposure, which has been summarized in 2016 . According to that workshop, short-term epidemiological studies provided evidence that exposure to traffic pollution (rich in UFPs) was associated with adverse cardiovascular outcomes, however, the effects still couldn't be reliably disentangled from other PM fractions or other gaseous pollutants. Similar to HEI's conclusion, epidemiological studies did not provide enough evidence that UFPs are more potent than other PM size fractions. Nevertheless, toxicological concerns about health effects of UFPs suggested that particle size may need to be considered in assessing potential adverse effects of exposures to PM.

Chen et al. (2016) thoroughly reviewed articles on composition of UFPs, their sources, typical characters, oxidative effects and potential exposure routes with a main focus on toxicology. Furthermore they also considered evidences emerging from nanotoxicology, as this research field contributes to the understanding of toxicity mechanisms of airborne UFPs in AAP. They concluded that UFPs play a major role in adverse impacts on human health.

An American working group (Li et al. 2016) reevaluated the conclusions made by the HEI report by assessing experimental, epidemiological and clinical trial studies published in 2014 and 2015. The authors mentioned a critical knowledge gap in clearly identifying the impact of exposure to the nano-scale pollutants on human health. However, due to new evidence, especially from experimental and toxicological studies, they questioned the validity of HEI's conclusion that there is no evidence that the adverse health effects of UFP were dramatically different from those of $PM_{2.5}$. Nevertheless, the issues of epidemiological studies assessing health effects of UFPs reported by the HEI Panel still remain.

Heinzerling et al. (2016), examining respiratory health effects of UFPs in children, identified 12 relevant articles from which 4 are not included in HEI. In single pollutant models, exposure to UFPs were associated with incident wheezing, current asthma, lung function and emergency department visits due to exacerbation of asthma. Only one study that reported significant association between asthma emergency department visits and UFPs, also adjusted for co-pollutants (Halonen et al., 2008). In this study, the association was no longer significant after adjusting for NO_2 exposure. Even though the evidence between UFPs and children's respiratory health is accumulating, the authors concluded for the same reasons stated by the HEI Panel that the evidence remains inconclusive.

In addition, Clark et al. published in 2016 a study focusing on biological mechanisms of cardiovascular effects beyond the alveolar barrier within the body or in vitro tissues exposed to UFPs and quasi-UFPs of up to 500 nm size. They concluded that there is some (e.g. altered autonomic modulation with increases of heart rate in animal models) up to strong evidence (e.g. vasoconstriction induced by endothelium-dependent and independent pathways mediated through UFPs) for various cardiovascular outcomes (heart rate, vasoactivity, atherosclerotic advancement, oxidative stress, coagulability, inflammatory changes).

Recently published epidemiologic studies now make it necessary to reevaluate the evidence base on the health effects of UFPs.

Aims of the project

The aims of this project were to systematically review the literature on the effects of UFPs on health, to evaluate the selected studies and to assess the transferability of the results to the situation in Germany. For this purpose, we focus on the following objectives:

1. Conducting a systematic literature review

   - Focus on health effects associated with ultrafine particles
   - Emphasis on epidemiologic studies and quantitative effect measures (e.g., relative risks, dose-response relationships)
   - Documentation of the literature search results and storage of all considered articles using a literature management database (EndNote).

2. Evaluation of the identified literature

   - Evaluation of individual study quality based on defined criteria
   - Evaluation of the transferability of the identified findings to the present conditions in Germany

3. Evaluation of the health relevance of ultrafine particles, specifically:

   - Within the context of other ambient air pollution (AAP) exposures (e.g., PM_{10}, $PM_{2.5}$, ozone, nitrogen dioxide)
   - With regard to the current German situation
   - When considering the projected trajectory of ultrafine particle exposure in Germany.

Methods

We systematically searched MEDLINE (Medical Literature Analysis and Retrieval System Online) for eligible studies investigating health effects of AAP related UFPs. The period included in the search was 01.01.2011 until 11.05.2017. In addition, we searched the LUDOK (Dokumentationsstelle Luftverschmutzung und Gesundheit)-database, which is provided by the Swiss Tropical and Public Health institute (Swiss TPH). This database contains scientific literature on the effects of AAP on human health.

The focus of the systematic search was on epidemiologic studies that explore health effects of UFPs including quantitative effect measures.

Another selection critera was the use of one UFP-measure (particle numbers (PNC) for particles with a diameter of less than 100nm, $PM_{0.1}$, nucleation mode particles, Aitken-mode particles as well as quasi-UFPs-measures: PNC for particles with a maximum diameter of > 100 nm, $PM_{0.25}$,

surface area concentrations and accumulation mode particles. Health outcomes were required to include mortality, morbidity, emergency/hospital admissions or subclinical[1] outcomes.

Toxicological studies were assessed only with regard to supporting evidence of the evaluation of UFP-related health relevance as stated in work package 3. Studies which investigate population related exposure to UFPs were assessed in order to evaluate the transferability of the reviewed results to the situation in Germany (work package 2b) and to evaluate the health related relevance of UFPs with regard to the situation in Germany (work package 3b) and in consideration of the potential trends of UFP exposure in Germany (work package 3c).

Search Strategy

The last comprehensive review was performed by the HEI including a systematic literature research in MEDLINE and Web of science up to May 2011. Within our project, we replicated their search strategy and discussed specific issues on the search strategy. We set the starting time of our search half a year earlier than the end poind of the search period of HEI in order to assess publications which may not have been indexed yet during the search period of the HEI.

The search strategy of the LUDOK database includes epidemiological and experimental original works studying the effects of „classical"/traditional ambient air particles on humans, as well as effects of further air pollutants . The search is conducted monthly using a constant, very broad search strategy in PubMed. The LUDOK search is complemented by hand search in more than 20 relevant journals, reference lists of publications and other sources. The search strategy within this project consisted of a modified HEI search strategy, completed by a search in LUDOK and hand searches. The keywords were extended in comparison to the HEI search keywords, following the very general search strategy of the LUDOK database. An alternative search strategy was applied using specific disease related keywords instead of the general keywords "health" and "epidemiology/ic/ical".

Further hand searches considered reviews of the last six years as well as reviews identified by our search. Finally, published abstract bands from relevant conferences and symposia were searched as well es publications by authors identified by our search.

Study Selection

Two reviewers screened title, abstracts and – if needed – full texts of the studies with regard to the inclusion and exclusion criteria (see below). 10 % of the studies were screened by both reviewers. In case of uncertainties concerning the selection of a study the case was discussed by the whole team. If necessary, inclusion and exclusion criteria were clarified and extended. The process of the study selection is illustrated in a Flowchart and documented in a chart adapted to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) method (Figure 4).

All references were organized within a library of a reference management program "Endnote" providing access for all project members (Figure 3).

Inclusion criteria

[1] Subclinical endpoints indicate biological measurements, e.g., of lung function, of heart rate variability, of atherosclerosis and arrhythmia as well as the examination of body fluids to determine systemic or pulmonary inflammation markers.

- Epidemiologic studies with an adequate study design, i.e.: cohort, case-control, cross-sectional, case-crossover, panel-studies, scripted exposures, time-series studies.
- Quantifiable measures of association containing at least one UFP measure/metric: Number (PNC) or size-fractioned PNC for particles < 100 nm, $PM_{0.1,}$ nucleation-mode particles (NucMP) and Aitken-mode particles (AitMP) or containing at least one quasi-UFP effect measures: PNC < 3000, $PM_{0.25}$, $PM_{0.1}$, surface-area concentration or accumulation mode particles (AccMP).
- Quantifiable measures of association including at least one measure: Odds ratio, relative risk, hazard ratio, β-estimates of percent change or exposure-response functions.
- Health outcomes including mortality or ICD-coded diseases, symptoms, emergency/hospital admissions/visits, preclinical outcomes.
- Languages: English, German.
- Year: Studies published from 2011 onward until 11.05.2017 which were not included in the HEI review; studies published after the deadline are listed in the appendix

Exclusion criteria

- Toxicological studies, controlled exposure studies, animal experiments, in-vitro studies,
- Exposure to industrially engineered nanoparticles,
- Exposure to nanoparticles/ UFPs in occupational settings,
- Exposure to source-related indoor nanoparticles/ UFPs,
- Exposure to diesel particles, BC or EC only,
- Distance measures in substitution of exposure measurements
- Health outcomes of unclear health relevance, e.g. epigenetics, metabolomics, methylation.

Evaluation of the identified literature

The identified articles were evaluated concerning their quality of report, significance and contents as well as their transferability to the German context. The established quality criteria are adapted from the Quality Assessment Tools of the National Heart, Lung and Blood Institute of the National Institute of Health (2014). When developing the different criteria, specific attention was paid to the exposure assessment. In particular, criteria to evaluate the applied measurement devices, the representativeness of the measurement sites for the exposure of the target population, the validity of potentially used exposure models and for the assessment/modeling of several air pollutants.

Results

Literature search

The application of the main search strategy in MEDLINE yielded 1,114 references, the application of the alternative outcome-specific MEDLINE search strategy yielded 992 references, of which 332 were not included in the main search strategy (Figure 4). Together, the MEDLINE search yielded 1,446 references. The search in the LUDOK database yielded 106 references, of which 30 were additional to the MEDLINE search. Another 8 additional references were identified through the hand search in other sources, yielding an overall total of 1,484 unique references that were examined for in- and exclusion criteria.

The final number of 85 original references included in this systematic review was achieved from the following sources: Of the 1,114 unique references identified by the main MEDLINE search strategy, 70 references were included in the analysis. Of the 332 unique references identified by the alternative outcome-specific MEDLINE search strategy, 3 additional references were identified for the review.

Of the 106 LUDOK references, 8 relevant studies were identified additionally. Of the 8 studies identified through hand search, 4 studies were added to the final analysis database.

In a repeated search on 23.02.2018, limited to articles published or accepted after the closing date of the full search, we identified another 13 articles, which are listed in the appendix.

Study characteristics

Most of included studies (n=85) were conducted in North America (n=37) or Western Europe (n=27). Further 12 studies took place in the Western-Pacific region. Only very few studies were conducted in Middle/ South America (n=1), Eastern Europe (n=2) and South-East-Asia (n=1). Three out of five multi-center studies included studies conducted in several Western Europe countries (Karakatsani et al., 2012; Manney et al., 2012; Samoli, Andersen, et al., 2016), two multi-center studies included study sites located both in Western and Eastern Europe countries (Lanzinger et al., 2016a, 2016b).

The majority of the studies were related to the investigation of short-term effects (n=75) measuring outcomes during hours to weeks after exposure. Ten studies investigated long-term associations using exposure estimates averaged over a period of months to years. The studies with a long-term study design consisted of cohort studies (n=4), cross-sectional studies (n=4), one case-cohort and one case-control study, respectively (Table 1). Short-term studies are dominated by panel studies - 31 as repeated measures and one in a cross-sectional design, scripted exposure studies (n=16), and time-series studies (n=11). Further studies investigating short-term associations were case-crossover (n=8), cohort (n=4) and cross-sectional studies (n=4).

Table 1: Study design by long-term/ short-term studies

Design	Number of studies	%
Long-term	all=10	
Case-cohort study	1	1.2%
Case-control study	1	1.2%
Cohort study	4	4.7%
Cross-sectional study	4	4.7%
Short-Term	all=75	
Cohort study	4	4.7%
Cross-Sectional study	4	4.7%
Panel (cross-sectional)	1	1.2%
Panel (repeated measure)	31	36.5%
Case-crossover	8	9.4%
Scripted exposure	16	18.8%

Time-series	11	12.9%
Total	**85**	**100.0%**

Overall, most studies used measurement-based exposure assessments (87.1%). Model based exposures were used in 10.6% of the studies. In long-term studies, mostly model-based exposure were used (9 out of 10), whereas the majority of short-term studies used measurement-based exposures (71 out of 75). This pattern is attributable to the fact, that model-based exposures are necessary to capture the spatial variation in exposure, which is the required exposure contrast for the assessment of long-term effects in different study design.

The majority of the studies applied central-site measurements (n=45), followed by mobile measurement techniques (n=17) and combination of different modeling/ measurements (n=10), e.g., central-site measurements in combination with spatio-temporal LUR models, residential measurements or microscale personal exposure models (Table 2).

Table 2: Type of exposure models/ measurements used in the studies

Exposure model/measurement	Number of studies	(%)
Chemical-transport model	3	3.5%
Land-use regression model	1	1.2%
Dispersion model	1	1.2%
Measurement: Central site	45	52.,9%
Measurement: Residential	2	2.4%
Measurement: Mobile	17	20.0%
Microscale personal exposure model	2	2.4%
Other	4	4.7%
Combination of different types	10	11.8%
Total	**85**	**100.0%**

In most studies, UFPs were assessed as particle number concentrations (PNCs) per volume. In about one third of the studies, PNCs sized up to 100 nm were used (29 out of 95[2]). In 66 studies, quasi-UFPs sized PNC fractions up to 3,000 nm were used. In relation to different size modes, only few studies used nucleation mode particles (n=1), representing particles with a diameter of less than 10 nm or Aitken-mode particles (n=1), representing particles with a diameter of 10-100 nm. In 14 studies, Accumulation mode particles were used, representing particles with a diameter of 100-1,000 nm[3] (see Figure 1, p.13). Particles measured as mass per m^3 are used in 11 studies: In six studies, submicron $PM_{0.1}$ particles were assessed, in seven studies, quasi-UFP $PM_{0.25}$ or $PM_{0.1}$ particles were assessed. LDSA was only used in two studies.

[2] As many studies used various size-fractioned PNCs, the number of analyses using PNCs with a size up to 100 nm (n=29) and/or up to 3,000 nm (n=66) exceed the number of 75 included studies that assessed PNCs.

[3] In literature, different cutpoints are used to divide particles in the different modi.

Table 3: Health outcome types of long-term and short-term-studies

	Number of studies	%
Long-term	all=10	
Mortality	1	1.1%
Morbidity	4	4.5%
Emergency/hospital call/admission	0	0.0%
Subclinical	5	5.7%
Short-term	all=78	
Mortality	7	8.0%
Morbidity	5	5.7%
Emergency/hospital call/admission	11	12.5%
Subclinical	55	62.5%
Total	88	100.0%

Eight studies assessing mortality analyzed the effects of UFPs on total, cardiovascular or respiratory mortality. Nine studies analyzed the effects on cardiovascular, respiratory, or other morbidity outcomes. Eleven studies investigated UFP effects on cardiovascular or respiratory disease-related emergency calls/ hospital admissions. The vast majority of studies used various subclinical measures as health outcomes. Three studies investigated several different types of main outcome types. Most studies measured cardiovascular organ system-related outcomes, followed by inflammatory and respiratory/atopy health outcomes. Few studies investigated total mortality, oxidative stress and other outcomes.

Quality indicators

In more than half of the studies (n=49), convenience samples were used. Six studies used random samples. Further seven studies used a combination of random and convenience samples. In 13 studies, study participants represented the general population in terms of sociodemographic aspects. In most of the included studies (n=62, 72.9%), the study population was a selected group, not representative for the general population. A sample size justification was rarely provided (n=3). Most of the study participants were recruited from the same populations and the same time period (n=71 and n=82).

The majority of the studies (n=66, 77.6%) reported the size-ranges of the measured UFPs. Almost all studies (n=79, 92.9%) reported the technical device used to measure the particles. Less than half (n=34) of the studies that assessed other air pollutants (n=78) adjusted for co-pollutants within multi-pollutant-models. Studies without adjustment for co-pollutant were considered as “high risk of bias”. 66 studies adjusted for meteorology, from which the majority (n=64) were short-term studies.

In all but one study (n=84) assigned exposure values were measured or modeled for time periods prior or parallel to the assessment of the outcome or for the time period of follow-up. In five of the included long-term studies, this was achieved by the use of chemical transport modeling, which allows the estimation of daily air pollutant concentrations for specific time periods. Fur-

thermore, all but one study (n=84) defined and described the outcome measures clearly. In 68 of the studies, a blinding of the outcome assessors could be presumed. In 15 studies, no blinding was ensured.

Short-term health effects

In comparison to the prior evidence, seven additional studies have been conducted with overall mixed results (Table 4). For all-cause **mortality**, only two out of four studies found positive estimates in analyses not adjusted for co-pollutants. Of these, only one study showed positive associations for quasi-UFPs after adjustment for other pollutants, while in the other study, elevated point estimates decreased towards the null upon adjustment.

The evidence of respiratory mortality is also scarce and inconsistent. Out of the five studies on respiratory mortality, four studies found positive, though mostly non-significant associations for UFPs or quasi-UFPs. Three studies adjusted for co-pollutants, with opposite effects after NO_2 adjustment, leading either to an enhancement or to an attenuation of effect estimates after adjustment for NO_2. The studies presented two-pollutant associations only for those models/ lags/ size fractions showing the strongest associations. Thus, the specific effect estimates are difficult to compare and consistency of the results cannot be fully assessed.

Similar to the overall results for respiratory mortality, associations of UFP/quasi-UFP with cardiovascular (CV) mortality are inconsistent. The six single exposure studies observe positive (three studies) as well as inverse associations (three studies) with CV mortality. In the two multi-pollutant studies, adjustment for NO_2 led to a decrease in effect estimates, causing the loss of significance in one study and a decrease to a significantly inverse relationship in the other study. Adjustment for $PM_{2.5}$ only caused small or no changes in the UFP estimate.

Evidence from this as well as from prior reviews suggests that effects may be larger in the warm season; therefore possible effect modification by season is an important factor to consider in future short-term effect studies. Moreover, the observed effects at least partially overlap with other air pollutant effects, most clearly seen for NO_2. Due to differences in investigated size fractions, no conclusions can be made about the most important fractions.

Table 4: Summary table of conducted analyses in the seven mortality studies

Study	All-Cause	Single pollutant associations	Multi-pollutant associations	Respiratory	Single pollutant associations	Multi-pollutant associations	Cardiovascular	Single pollutant associations	Multi-pollutant associations
Lanzinger et al. 2016a	✓	0	0	✓	(+)	+	✓	(-)	-
Leitte et al. 2012				✓	UFP: (+), quasi-UFP: +	UFP: 0 quasi-UFP: (+)			
Meng et al. 2013, (only quasi-UFP)	✓	+	+	✓	(+)	nc	✓	+	nc
Samoli et al. 2016	✓	0	0	✓	-	-	✓	(-)	nc
Stafoggia et al.,2017	✓	(+)	(-)	✓	+	nc	✓	(-)/(+)*	nc
Su et al. 2015							✓	+	(+)

Wolf et al. 2015							✓	(+)	nc

0 indicates no association. (+) and (-) indicates primarily non-significant associations, + and - indicate significant associations. Nc: not conducted. *varying across lags

Of the few studies investigating short-term effects of UFPs/quasi-UFPs on **morbidity** outcomes, only two studies observed significantly elevated estimates with a marker of perceived stress and with various symptoms. Since none of the above mentioned studies adjusted for co-pollutants or were by design able to disentangle the independent effects of different constituents of the air pollution mixture, we cannot conclude an independent effect of UFPs on morbidity outcomes. The evidence base for CV morbidity outcomes is scarce with only two studies available on different outcomes. This evidence suggests that participants with preexisting cardiovascular disease might be more susceptible to adverse associations with elevated UFP/quasi-UFP concentrations.

However, while both studies show generally positive associations, no inference on the independence on the reported UFPs effect can be made. The evidence for associations with short-term changes in mental health symptoms is insufficient.

The evidence base for UFP-related effects on **utilization of the healthcare system** due to respiratory symptoms is scarce (Table 5). Possible associations seem to be most probable for children as a susceptible subgroup. While single-pollutant associations were observed in few studies, multi-pollutant models of the studies could not verify independent associations of UFPs/quasi-UFPs with respiratory hospital admissions/emergency department visits. Specifically adjustment for NO_2 led to a decrease in estimates, which mostly reached the null in co-pollutant models.

Most studies investigating cardiovascular disease-related use of the healthcare system indicate weak associations being stronger for shorter time lags of up to 24 hours. These associations decreased upon adjustment for co-pollutants with no clear evidence for independent associations of UFPs/quasi-UFPs with cardiovascular emergency department visits/hospital admission.

Table 5: Summary table of conducted analyses in the 11 studies on emergency department visits/hospital admissions

Study	Respiratory	Single pollutant associations	Multi-pollutant associations	Cardiovascular	Single pollutant associations	Multi-pollutant associations
Evans et al., 2014	✓	(+)	(+) (no NO_2 adjustment)			
Gardner et al., 2015				✓	(+)/0	nc
Iskandar et al., 2012	✓	(+)	0			
Rosenthal et al., 2013				✓	(+)/+	0
Wichmann et al., 2013				✓	(+)/0	nc
Delfino et al., 2014	✓	nr	nc			
Diaz-Robles et al., 2014	✓	+				
Lanzinger et al., 2016	✓	(+)	0	✓	(+)/0	0
Samoli UK, 2016	✓	(+)/(-)	(+)	✓	(+)	(-)/(+)
Samoli EU, 2016	✓	(+)/(-)	(-)/-			
Liu et al., 2013				✓	+/(+)	nc

0 indicates no association. (+) and (-) indicates primarily non-significant associations, + and - indicate significant associations. Nc: not conducted, nr: not reported

An overview over **subclinical outcomes** is provided in table 6.

Most of the studies on subclinical respiratory endpoints have only limited sample sizes (15-84 participants). Moreover, study samples were frequently selective, either representing healthy young adults or persons suffering from atopy and/or asthma. The investigated lags and averaging periods differ across studies, but generally, most associations were found in a time range of 0-48 hours after increased exposure. Finally, results of the studies are mostly inconsistent in relation to the specific respiratory endpoints. With regard to peak-flow endpoints, measurement error could be an issue in this self-monitored endpoint, especially in the study by Cole-Hunter et al. (2013) which could not be blinded. Due to the lack of adjustment for co-pollutants, little can be concluded regarding the independence of effects. The scarce evidence on studies with co-pollutant adjustment suggests an at least partial overlap of UFP, respectively PNC effects, with NO_2-effects.

The majority of studies found adverse associations between exposure to UFP/quasi-UFP and **blood pressure indices**, indicating increases in blood pressure (BP). These results differed across different endpoints (systolic blood pressure (SBP), diastolic blood pressure (DBP), pulse pressure (PP)), different size fractions and lag periods. Apart from one study with more than 1,000 participants, the studies consisted of smaller study populations. In addition, all study samples represented selected group, impeding a transfer to the general population. Apart from these limitations, the evidence from two-pollutant studies is too scarce to draw conclusions on independent UFP effects on blood pressure indices.

A relatively large body of evidence (16 studies) is available for **heart-rate variability** (HRV) **indices**, of which 12 showed UFP-related associations on at least on one HRV outcome. Upon adjustment for co-pollutant, associations changed in both directions. Across studies, different time-windows and different co-pollutants were examined, so that no clear pattern can be observed.

Considering the limited number of studies on **arrhythmia** outcomeswith only one study, the evidence base is still insufficient.

The majority of the seven studies examining associations between UFP/quasi-UFP and **vascular function** indicate a possible association. However, a lack of consistency regarding the study design, specifically the outcome parameters, as well as missing co-pollutant models do not allow overall conclusions.

All 12 studies which have been investigated UFP-effects on pulmonary inflammations suggest positive associations between UFP and adverse changes in the pulmonary inflammation marker, in particular immediately after exposure. Nevertheless, the evidence base for pulmonary inflammation in response to UFP is still limited as the studies used different subgroups, exposure metrics, outcome measures and time frames. The two studies that conducted two-pollutant models observed overall robust effect estimates.

The majority of the 18 studies investigating UFP effects on systemic inflammation markers indicate inconsistent associations. Effects of UFP on indices for high sensitivity C-reactive protein (hs-CRP), fibrinogen, blood cell counts, myeloperoxidase varied, which may originate from different compositions of participants, assessed PNC fractions and exposure assessment types. In most studies, effects seem to be most pronounced for shorter lag periods. Only few multi-

pollutant models do not allow statements on independent effects of UFPs/ quasi-UFPs, as only two of the five conducted studies with multi-pollutant models showed robust results.

Table 6: Summary table of conducted analyses in the 55 studies on subclinical outcomes

Outcome	Number of studies	Number of studies with single-pollutant--associations in expected direction	Number of studies with multi-pollutant associations in expected direction	Comments (i.e. studies with significant results in the non-expected direction)
Respiratory indices	11	4/11	3/3	Li et al. (2016) found significantly positive associations between UFP and FEV_1 & FVC
Blood pressure	13	9/13	2/4[4]	Two of the nine studies with associ. showed inconsistent results across lags
HRV	16	12/16	3/5	In Zhang et al. (2013), effect estimates decreased upon adj. for NO_2 and increased upon adj. for O_3
Arrhythmia	1	1/1	nc	Strong associations with $PM_{0.25}$, nearly protective associations between PN and hourly nighttime measured tachycardia
Vascular function	7	4/7	1/2	
Pulmonary inflammation	12	12/12	2/2	Most studies investigated effects on FeNO
Systemic inflammation (incl. fibrinogen)	18	7/18[5]	2/5	Significant inverse associations between fibrinogen & PNC upon adjustment for NO_2 (Strak et al., 2013)
Neurocognitive outcomes	2	1	nc	-

HRV: Heart rate variability.

Long-term health effects

A limited number of studies, varying outcomes and exposure assessment methods as well as lacking co-pollutant adjustment do not allow to draw final conclusions. The summarized results are presented in table 7.

Table 7: Summary table of the ten long-term studies in single and multi-pollutant associations.

Outcome type/ study	Outcome	Single pollutant associations	Multi-pollutant associations
Mortality/ Ostro et al. 2015	- all-cause - cardiovascular/ IHD - pulmonary	0 (+)/0 0	nc nc nc

[4] One of the four studies did not show assoc. in single-pollutant models, either. A further study (Rich et al., 2012) did not show all results, therefore rated as non-associated here

[5] Most positive associations relate to fibrinogen

Morbidity / Li et al. 2017	- cardiometabolic	(+)	nc
Laurent et al. 2014/2016b	- low birth weight	+/(+)	nc
Laurent 2016a	- preterm birth	-/+	nc
Subclinical/			
Aguilera et al. 2016	- carotid-intima-media thickness (PNC/LDSA)	+/+	-/(+)
Viehmann et al. 2015	- hs-CRP/ fibrinogen/ WBC	(+)/+/(+)	nc
Lane et al. 2015	- hs-CRP/ IL-6	(+)/(+)	nc
Lane et al. 2016	- hs-CRP/ IL-6/ TNRFIII/ fibrinogen	(+)/(+)/(+)/(-)	nc
Sunyer et al. 2016	- working memory,	(+)	nc
	- superior working memory	+	
	- inattentiveness	+	

IHD: Ischemic heart disease, 0 indicates no association. (+) and (-) indicates primarily non-significant associations, + and - indicate significant associations. nc: not conducted.

Summary of short-term and long-term health effects

An overview on all included short-term and long-term studies reflects the inconsistency of the results (Table 8). More than half (n=49) of the studies on short-term effects (n=79) reported at least one significant effect in the single pollutant model, especially those studying mortality or subclinical outcomes. For less than half of the single-pollutant associations (21 of 49), the general pattern of the association was consistent regardless of the significance level. 18 out of 32 studies found at least one significant association in multi-pollutant models. The associations in multi-pollutant studies remained consistent in about half of the studies (n=7).

Associations between health outcomes and long-term exposure with ultrafines were more consistent in the single pollutant models even though there were considerably fewer studies. Nevertheless, long-term studies adjusting for other pollutants are still lacking with only one study, which did not show effects in the multi-pollutant model.

Table 8: Summary table of associations for all included studies.

Outcome	Single pollutant effect	Consistency of general pattern	Multi-pollutant effect	Consistency of general pattern
Short-term	49/79*	21/49	18/32	7/18
Mortality	5/7	2/5	4/6	1/4
Morbidity	3/7	0/3	-	-
Hospital admission	4/10	2/4	0/5	-
Subclinical	37/55	17/37	14/21	6/14
Long-term	8/10	1/1	0/1	-
Mortality	1/1	1/1	-	-
Morbidity	3/4	-	-	-
Hospital admission	-	-	-	-
Subclinical	4/5	-	0/1	-

*the number of short-term studies exceed 75, as three studies used different outcome types.

Discussion

Literature search

We conducted a systematic comprehensive search of relevant epidemiological studies on ultrafine and quasi-ultrafine particles for the period from 01.01.2011 until 11.05.2017. The different strategies of our search consisted of a MEDLINE search, using two alternative strategies, a search in the specialized data base LUDOK, and a hand search in review articles and reference lists of identified publications. Overall, the additional yield of the alternative MEDLINE search strategy, and of the complementary search strategies (LUDOK and hand search), and of the repeated search was substantial, with altogether 15 additional references added to the final analysis. Additionaly 13 articles were identified per MEDLINE and hand search in February 2018. This relatively high yield reflects the lag in indexing newly published studies in large literature data bases as well as the fast development of an emerging scientific field. More specialized data bases such as the dedicated LUDOK literature data base are therefore very useful for targeted and timely research.

Evaluation of health relevance of ultrafine particles

Our evaluation of the health relevance of UFPs is based on the above described epidemiologic studies and how they add to the available the evidence since the comprehensive review conducted by the HEI, published in 2013. Overall, the epidemiological evidence is quickly increasing and it can be expected, that the next few years will bring a substantial increase in relevant studies. Currently, we are still in the beginnings of health-related research of UFPs, which is in part due to the still developing methods (see sections below on exposure assessment).

The HEI concluded in its review that "the current database of experimental and epidemiologic studies does not support strong and consistent conclusions about the independent effects of UFPs on human health" (Health Effects Institute, 2013). Major reasons for this lack of evidence, specifically for epidemiologic studies, lie in the difficulty of assessing population-based exposure to UFPs for short-term as well as for long-term studies. Due to the specific properties of UFPs with a high temporal and spatial variability, common exposure assessment strategies, which have been developed for the more homogeneously distributed larger particle fractions, will lead to larger exposure misclassification when applied to UFPs. Nevertheless, HEI does not conclude that independent effects of UFPs can be ruled out, but rather recommends the exploration of alternative exposure metrics, spatial modeling techniques, and statistical methods.

In this review, we use similar design- and outcome-specific categories as in the HEI review to be able to integrate our findings with the prior evidence. Since independence of effects is the key question regarding the health relevance of UFPs, we specifically focus on studies with co-pollutant adjustment.

Inconcistency of results by endpoint

Previous evaluations have concluded, that the combined results for respiratory as well as for cardiovascular endpoints are still inconsistent (Health Effects Institute, 2013). When considering the newly acquired evidence during the years from 2011 to 2017, this picture has not changed substantially. Even though there is a growing number of specifically designed studies to investigate health effects of UFPs, we cannot identify a consistent pattern of health effects on either respiratory or cardiovascular disease across the different endpoints including mortality, morbidity, emergency department visits/hospital admissions or subclinical endpoints. For other outcomes such as mental disorders, neurocognitive function or birth outcomes, the evidence base is still too small to derive firm conclusions.

Even though results are not consistent across different outcomes types, the majority of the 11 studies investigating short-term effects on BP, the major risk factor for cardiovascular disease, indicate an association with increased blood pressure. Once again, evidence from the three co-pollutant-adjusted studies is mixed, which underscores the necessity of further studies with co-pollutant adjustments.

The lack of consistent findings can be explained by a number of factors. These include differences in exposure assessment (see below), endpoint assessment, study design and size, and different confounder control, specifically differences in the adjustment for co-pollutants (see below).

Long-term exposure and health effects

In contrast to the last prior comprehensive review by HEI (2013), ten studies have been published investigating long-term effects of UFPs on various health outcomes. While most of these studies found elevated point estimates for associations of UFPs with adverse health outcomes, only one study adjusted for co-pollutants, including NO_2. Adjustment with NO_2 led to a decrease in the effect estimate to an inverse association.

While the current evidence base does not support an independent effect of UFPs on health outcomes, this should by no means be mistaken for a proof of the absence of such an effect. As will be discussed below, current exposure assessment techniques are not well suited to describe and investigate long-term exposure to UFPs. More studies applying novel methods for individual-level exposure to UFPs are therefore urgently needed. Important applications are next to road traffic-related exposures also the emerging problem regarding exposure to UFPs in the vicinity of airports, which has only recently been described (Hudda & Fruin, 2016).

Exposure assessment

Overall, the number of studies including the assessment of exposure to and the investigation of health effects of UFPs is rapidly increasing. One important factor contributing to this rapid increase is the development of new instrumentation, which enables a less expensive assessment of UFP/quasi-UFP for example with condensation particle counters. However, research is still at the beginning and new exposure assessment methods need to be defined and employed in epidemiological studies.

Challenges of exposure assessment for UFPs include the high spatial and temporal variability of UFP/quasi-UFP, which necessitate different exposure assessment designs than the “classical” air pollutants like $PM_{2.5}$ and PM_{10} with a much more homogeneous spatial distribution. This high spatial variability is of concern not only for long-term health effects studies, which are based on long-term spatial differences in exposure, but also for short-term studies with a central-site measurement. These studies assume that the temporal changes from day to day are evenly distributed across the sometimes very large study areas; an assumption that might not hold true for UFPs. Given the possibility of a larger exposure estimation error for UFPs compared to other pollutants, a systematic bias towards the null in single-pollutant studies and in multi-pollutant studies is probable (Dionisio et al., 2014).

In the future, the development of enhanced spatiotemporal models can contribute to a more precise exposure assessment across larger areas. Current models such as the German EURopean Air Pollution Dispersion (EURAD) model need to be adapted to incorporate specific sources, validation measurements and increase the spatial resolution.

A further challenge of UFP/quasi-UFP exposure assessment is the non-standardized equipment and the non-standardized use of size fractions in the studies. The commonly used measurement devices have different lower cutpoints for the particle size. Since the majority of particles are located in the nucleation mode (< 20 nm) of the particle size distribution, even small differences in the lower cutpoint between 1 and 20 nm can lead to substantial differences in particle number concentration. Futhermore, the reporting of the exposure assessment often does not include the exact size range of particles, which prevents direct comparisons of exposure between studies.

Independence of effects

Even though several studies across the investigated endpoints have observed positive associations of UFP/quasi-UFP with various health effects, the overall evidence for independent effects is still insufficient. We noticed, that specifically the newer studies conduct multi-pollutant models with a higher frequency than the older studies, which is a positive development (e.g., Aguilera et al., 2016; Croft et al., 2017; Lanzinger et al., 2016a; Samoli et al., 2016; Stafoggia et al., 2017). However, the type of adjustment still varies substantially between studies and there is no standard strategy for co-pollutant adjustment yet. At the moment, adjustment for NO_2 generally seems to exert a greater effect on the point estimate than other co-pollutants (e.g., Lanzinger et al., 2016a&b; Su et al., 2015; Samoli, Andersen et al., 2016, Zhang et al., 2013). One reason for this is the overlap in sources and spatial/temporal distribution of UFPs and NO_2, which can lead to instability in the models and biased effect estimates in two-exposure models.

Transferability of results to the situation in Germany

The transferability of the above reported results to the situation in Germany will be judged according to the following criteria: Localizations of identified studies and level of exposure to UFPs, level of exposure to airborne co-pollutants, baseline prevalence of investigated diseases and selection of study populations.

Exposure to UFPs

The vast majority of the identified studies are located in North America (n=37, 43.5%) or Western Europe (n=27, 31.8%) and 5 studies (6%) located in more than one world region. When examining the study sites of studies with multiple study centers, we can observe that the majority of study sites are located in Western and Southern Europe (44 of 101 study sites, 43.6%). The concentrations of UFPs vary considerably in time and space and direct comparisons of single center measurements are subject to large variation depending on hour, day and season of measurement as well as exact placement of the measurement site (traffic, urban background, regional background site)(Birmili et al., 2016; UFIPOLNET, 2008). In the German Ultrafine Aerosol Network (GUAN), long-term measurements of ultrafine and fine particles have been conducted at 17 sites across Germany, including alpine sites (Zugspiptze), rural sites, urban background and roadside measurement sites (Birmili et al. 2016). Of note, the size of the measured particles ranges from 20 to 800 nm, thereby not encompassing the nucleation mode of particles and including the accumulation mode particles. Preliminary results of GUAN measurements indicate a range of hourly median concentrations of particle number (sized 20-800 nm) between 900/ml (Zugspitze) and 9000/ml at the roadside in Leipzig. Hourly mean concentrations are higher with 1120/ml at the Zugspitze and 10.500/ml in Leipzig. The 95 percentile of the distribution of hourly values reaches 22.400/ml in Leipzig-Mitte. All three roadside measurement sites had

P95 values above 19.900/ml, while the urban background sites ranged between 10.000 and 20.000/ml. GUAN also demonstrates the substantial variation in particle size distribution during the course of a week at six mainly urban sites.

The identified studies conducted in Western Europe typically have similar or higher mean total particle counts. A direct comparison is not possible with the available information, since instruments for measurements differ and have different lower cutpoints. 16 out of the 27 studies in Western Europe report the lower cutpoint of their measurement device as 10 nm or lower. Some devices go down as far as 3 nm as their lower cutpoint. Since the majority of particles is sized below 20 nm (nucleation mode) (HEI perspectives, 2013), small differences in the lower cutpoint leads to substantial differences in mean exposures. In addition, the upper cutpoint also varies considerably, with only few studies examining ultrafine particles in the more strict sense (<100 nm), but rather use the surrogate of total particle number concentration as the exposure of interest. This, however, presents a minor problem as total particle number is dominated by the size fraction below 100 nm (HEI perspectives, 2013).

For the benefit of this review, GUAN primarily demonstrates the large variability of exposures within Germany, but it is not well suited to compare absolute values with other studies, which used different measurement devices. The five studies from Germany included in the review are based on central-site or personal measurements (n=4) with lower cutpoints ranging between 3 and 10 nm. These studies yield mean exposures between 10.000/ml and 20.000/ml, which is comparable to other studies in this review. In comparison, the 13 studies located in the Western Pacific region or in South-East-Asia, in the metropolitan areas of China, South Korea or Taiwan, report measured mean particle number concentrations in similar or slightly higher ranges. The only German study based on modelled exposures applying the EURAD CTM yielded substantially higher mean exposures due to the modelling process, which included the complete nucleation mode and therefore also encompasses short-lived particles sized below 3 nm. We therefore conclude that the level of exposure in the identified studies, while very variable across time and space, is generally comparable to the German situation.

The development of population exposure to ultrafine and quasi-ultrafine particles in Germany in the coming years depends on several factors: (1) the formation and emission of these particles, (2) the spatial distribution of the population, and (3) the concentration of fine particles in ambient air.

According to a size-resolved pan-European anthropogenic particle number inventory, the most important sources of emissions are road traffic in urban areas and alongside highly trafficked roads (Health Effects Insitute, 2013). Traffic-related emitters of primary UFP are direct injection engines in vehicles, which have increased in number during the last decade and will probably increase further (Köllner, 2016). On the other hand, vehicles with Diesel-powered engines, which also emit particles in the ultrafine and quasi-ultrafine size range, have been equipped with particle filters. This has reduced the emission of fine particles substantially (according to EURO5a less than 5 mg/km). For UFP, the EURO5b norm for the first time sets a limit at 6×10^{11} (European Union, 2007). Overall, with increasing traffic and a rising number of city dwellers expected in the future (Vallance et al., 2010), exposure to road traffic-related UFPs is likely to increase in Germany in the next decade.

A further source of mostly ultrafine particles is aircraft traffic. Several exposure studies have documented increased UFP exposure downwind of airports around the world (Hudda et al., 2014; Keuken et al., 2015; Masiol et al., 2017; Shirmohammadi et al., 2017; Stafoggia et al., 2016). The increased short-term exposure is correlated with aircraft movements over time and

reach concentrations up to 50,000 particles/ml (Keuken et al., 2015) 7 km downwind of the airport in Amsterdam and up to 75,000 particles/ml (Hudda et al., 2014) 8 km downwind in Los Angeles. The same studies show that long-term concentrations are elevated up to 3-fold 7 km downwind with more than 200,000 exposed inhabitants close to Schipohl airport, Amsterdam (Keuken et al., 2015) and up to 4-5-fold in Los Angeles, 8-10 km downwind (Hudda et al., 2014). Similar exposure studies are ongoiong in Germany and will yield first information about the exposure of residents close to German airports. Given the increase in air travel, the exposure due to aircraft emissions is likely to play an increasing role in the future.

Moreover, the concentration of fine particles in ambient air is a determinant of UFP in a way that UFP will collide and coagulate with larger particles. A high concentration of ambient fine particles will therefore support the clearance of UFP in ambient air. With the reduction of fine particles, UFP will likely stay longer airborne than in an environment with high PM concentrations.

Exposure to co-pollutants

The level of airborne co-pollutants are important, as most of these co-pollutants have own effects on the outcomes of interest. 78 of the 85 identified studies (92%) assessed the level of at least one other air pollutant; however, only 34 studies adjusted for at least one co-pollutants in their analysis. Assessment of and adjustment for airborne co-pollutants is therefore not conducted in a comparable way across the identified studies.

Analysis of the multi-pollutant models revealed, that $PM_{2.5}$ and NO_2 are the co-pollutants which tend to influence the UFP/quasi-UFP estimate the most. Often, but not always, does the adjustment for NO_2 lead to an attenuation of the association of UFP/quasi-UFP with the health outcome (Leitte et al. 2012; Meng et al. 2012; Stafoggia et al., 2017; Su et al. 2015; Iskandar et al., 2012; Lanzinger et al. 2016; Rosenthal et al. 2013; Gong et al., 2014; Janssen et al. 2015; Steenhof et al., 2013). Adjustment for PM_{10} and $PM_{2.5}$ also attenuates the UFP/quasi-UFP association in several studies, but in most cases less than the NO_2 adjustment.

The level of co-pollutants, and specifically $PM_{2.5}$ and NO_2, can be compared across Europe using the "Air quality in Europe - 2017 report" by the European Environmental Agency (European Environmental Agency, 2017). According to this report, Germany ranks top among the 28 member states regarding the annual mean of NO_2 at the included monitoring sites (European Environmental Agency, 2017; Fig 6.1). Similar to UFP/quasi-UFP, the annual mean at selected monitoring sites is not able to give a comprehensive overview of the exposures of the study populations in the included studies, as NO_2 concentrations are subject to a high variability across time and space. Of the 34 studies that adjusted for co-pollutants, 15 were conducted in Western Europe. Of those, three were conducted in Germany, Augsburg, and all other studies were conducted in mostly major cities in Switzerland, the Netherlands, Sweden, Denmark and Finland with comparable traffic exposures.

We therefore conclude that the findings of an at least partial overlap of effects between UFPs and NO_2, which we observe in the Western European studies included in this review (Iskandar et al., 2012; Janssen et al., 2015; Rosenthal et al., 2013; Stafoggia et al., 2017; Steenhof et al., 2013), hold true for Germany as well.

Disease prevalence

The majority of the studies identified in this review is located in Western/Southern Europe and North America. The cause-specific age-adjusted death rates for all non-communicable diseases

and for respiratory diseases for 2015 are similar for the WHO Region of the Americas (including South America, which is not included in this review) and the WHO European Region (World Health Organization, 2016b). On the other hand, the annual cause-specific age-adjusted death rates for cardiovascular diseases differ, with a substantially lower age-specific death rate in the Americas (211/10,000) compared to the European Region (344/10.000). This difference is primarily due to the combination of both Americas in this statistic. Compared to other European countries and the USA included in this review, Germany has a similar distribution of causes of premature deaths as the Netherlands with ischemic heart disease, lung cancer, Alzheimer disease, cerebrovascular disease and chronic obstructive pulmonary disease (COPD) ranking 1 to 5 in both countries. This ranking is very similar in the UK, Denmark, Sweden, Spain and the USA.

Moreover, the majority of studies investigate short-term sublinical outcomes and of those, cardiovascular, respiratory and biomarker outcomes present the focus of the included studies. The outcome assessment of these studies is not subject to country-specific ICD-coding conventions. Unless baseline differences in physiological markers exist between the populations included in this review and the German population, which we have evidence for, transferability on results for Germany can be inferred.

Study population

Most studies included in this review are based on selected study populations (n=62, 72.9%) and only 10 (11.8%), respectively 13 (15.3%) studies were deemed representative or at least somewhat representative of the general population. The studies deemed to be completely representative of the target population are the time-series studies, which are based on general populations of the city of study. One of these time-series studies (Diaz-Robles et al., 2014) targeted selected age-groups within the general population. Of the other studies, 13 (15%) studies include at least one random sample of the source population. Almost all identified articles describe the study population well. The 10 studies investigating long-term effects are mostly analyses based on existing cohorts of several hundreds to thousands of participants, exclusively located in Western Europe or North America. Of these, 6 studies target the adult population of either sex or limited to one sex (Ostro et al., 2015), and 4 studies target children (Laurent et al. 2014, 2016a and 2016b; Sunyer et al., 2015). Among the short-term studies, the study populations are mostly highly selected small groups of either healthy (younger) adults or participants with a respiratory or cardiovascular disorder such as asthma, COPD, coronary artery disease, etc.

Transferability – conclusions

Based on the above descriptions of exposure level, co-pollutant exposure, baseline disease prevalence and included study populations we conclude that the overall results of this review can be transferred with the appropriate caution to the German situation.

Important limitations are (1) the paucity of studies with co-pollutant adjustment, which is specifically important because of the high NO_2 exposures in Germany, and (2) the use of highly selected groups in short-term studies, as these often do not include specifically vulnerable populations such as patients with badly controlled disease, newborns and children.

Overall conclusions

The investigation of health effects in epidemiological studies is a rapidly increasing field of research and substantial developments have been made during the last seven years, tackling two

of the most urgent open questions of research: First, several studies on long-term health effects of UFPs have been conducted and published. Second, specifically the more recent studies have undertaken efforts to control for co-pollutants to identify independent effects of UFPs.

Despite the obvious development in the field, the overall conclusions have not changed substantially over the time period investigated in this study.

First, the evidence on health effects remains inconclusive or insufficient for most of the studied outcomes. Specifically, while a number of studies have investigated mortality and emergency department/hospital admission outcomes, the relatively few studies with co-pollutant adjustment reveal mixed and, up to now, inconclusive evidence. In terms of number of studies, most evidence is available from studies investigating subclinical outcomes. Within this group of studies, cardiovascular outcomes and outcomes of pulmonary and systemic inflammation show the most consistent patterns with associations generally pointing into the direction of the adverse health outcome. Nevertheless, the evidence for independence of effects remains limited here as well, as only few studies have adjusted for co-pollutants.

Second, exposure assessment in the population remains difficult, due to the specific characteristics of UFPs. Studies using central-site exposure assessment probably miss a large part of the variability. Studies using classical spatial modeling techniques need to incorporate the very high spatial and temporal variability. Null findings or reductions in UFP/quasi-UFP effect estimates upon co-pollutant adjustment can at least in part be explained by exposure misclassification and measurement error. Exposure assessment has to devote special attention to measurement techniques, size-fractions and localisations of monitor placement. Reporting needs to be standardized to make studies more easily comparable.

Third, the independence of UFPs cannot be evaluated at the moment, due to the low number of studies with adjustment and the above mentioned limitations to exposure assessment for UFPs. A positive development is the increase in studies paying attention to this issue.

Fourth, there is still an urgent need for long-term studies on health effects of UFPs. This will require the development of modeling techniques. Furthermore, specific high-exposure situations need to be identified and described in more detail to be able to assess long-term health effects. Specifically, while near road exposures have already been recognized as important factors, airport-related exposures, which have recently been shown to be substantially above background concentrations, have not been included in health effects studies yet.

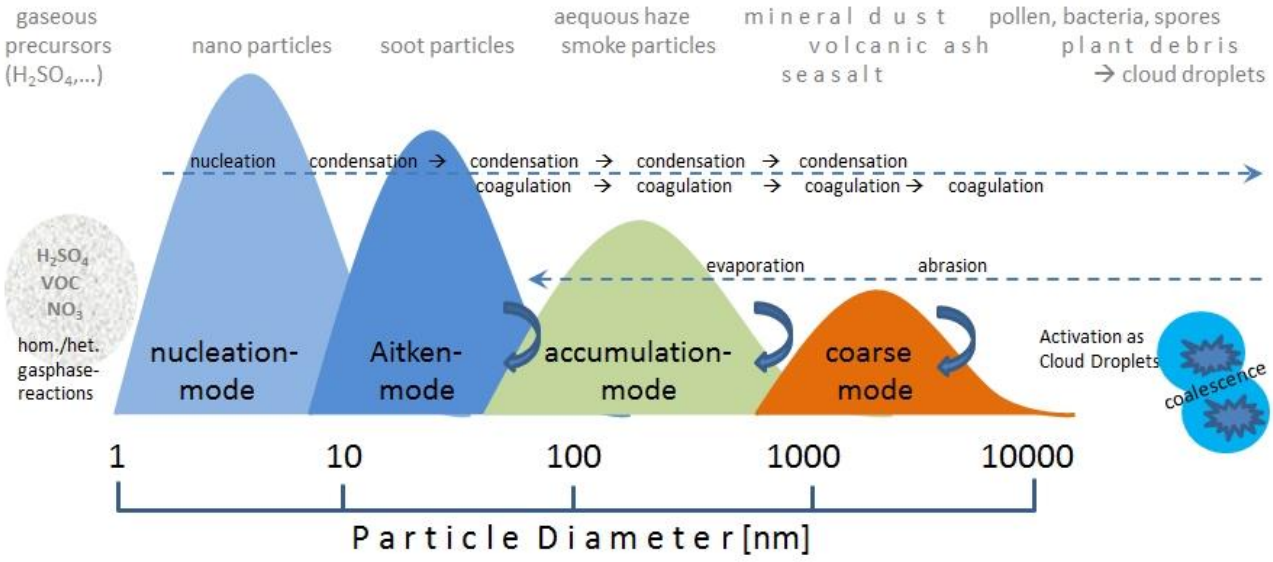
In addition to these general conclusions, we conclude that the overall results of this review can be transferred with the appropriate caution to the German situation. Important limitations are (1) the paucity of studies with co-pollutant adjustment, which is specifically important because of the high NO_2 exposures in Germany, and (2) the use of highly selected groups in short-term studies.

Zusammenfassung

Hintergrund

Ultrafeine Partikel (UFP) bzw. Ultrafeinstäube sind Partikel, welche einen aerodynamischen Durchmesser von maximal 100 Nanometer haben. Da UFP nur einen geringen Anteil zur Partikelmasse der Umgebungsluft beitragen, werden UFP meist als Partikelanzahl pro ml erfasst. Darüber hinaus werden in der epidemiologischen Forschung Partikelfraktionen unterschiedlicher Größenfraktionen genutzt. Dazu zählen nucleation-mode Partikel (Durchmesser von bis zu ca. 20 nm), Aitken-mode Partikel (Kondensationspartikel mit einem Durchmesser von ca. 10 bis 80 nm) sowie accumulation-mode Partikel (Partikel aus Kondensation und Koagulation mit einem Durchmesser von ca. 50 bis 1.000 nm). UFP unterscheiden sich aufgrund physikalischer und chemischer Eigenschaften von größeren Partikeln und werden direkt emittiert oder aus Vorläufersubstanzen im Rahmen sekundärer atmosphärischer Prozesse gebildet. In städtischen Gebieten stammen die UFP vor allem aus Verbrennungsprozessen durch motorisierte Fahrzeuge, insbesondere in Straßennähe (Health Effects Insitute, 2013; Kelly et al., 2012).

Abbildung 1: Größenfraktionen luftgetragener Partikel (Deutscher Wetterdienst, 2018)



Aufgrund ihrer geringen Größe können UFP nach Inhalation bis in den Alveolarbereich eindringen und sogar Zellmembranen durchdringen. Hierdurch können sie in die Blutbahn übergehen sowie letztlich in alle Körperorgane inklusive des Gehirns und des Nervensystems gelangen. Experimentelle Studien deuten auf einen Zusammenhang von UFP mit kardiovaskulärer und respiratorischer Morbidität und Morbidität sowie der Entstehung von lokalen und systemischen Entzündungsprozessen sowie adverse Effekte auf Gehirn und Stoffwechsel hin (Health Effects Insitute, 2013). Mehrere Expertenkommissionen haben in den vergangenen Jahren eine kritische Interpretation unter anderem der epidemiologischen Evidenz der zu UFP vorliegenden Erkenntnisse vorgenommen (Health Effects Insitute, 2013; World Health Organization, 2013). Die vom Health Effects Institute (HEI) , Boston und der Weltgesundheitsorgansiation (WHO) eingesetzten Kommissionen stellten im Jahr 2013 fest, dass es zwar wissenschaftliche Hinweise auf gesundheitsschädigende Wirkungen von UFPs gibt, wobei aber speziell für epidemiologische

Studien die Evidenz insgesamt noch nicht ausreicht, um eine gesetzliche Regulierung von UFPs zu empfehlen.

Unsere Literaturrecherche und das bisherige Wissen basieren auf einigen relevanten Übersichtsarbeiten, die in den letzten Jahren veröffentlicht wurden. Als erstes ist der HEI-Review zu nennen, welcher die bis dahin umfangreichste und vollständigste Datenbasis zu einer möglichen Assoziation zwischen UFP und verschiedensten Gesundheitsendpunkten liefert. Adverse Gesundheitseffekte durch UFP werden als möglich, jedoch nicht eindeutig erwiesen bewertet. Gründe für diese uneindeutige Lage sind unterschiedliche Gesundheitsendpunkte und verwendete Studiendesigns, die einen direkten Vergleich von Studienergebnissen verhindern, unterschiedliche und möglicherweise verzerrte Expositionserfassungen sowie fehlende Studien, welche für weitere Luftschadstoffe adjustiert haben. Darüber hinaus hat die HEI-Suche keine Langzeitstudie identifiziert, so dass die Evidenzlage insgesamt nicht ausreichend war, um Regulierungsmaßnahmen bezüglich UFP zu empfehlen.

Eine weitere Übersichtssarbeit dokumentiert Ergebnisse einer Expertenkonferenz zu Gesundheitseffekten durch die Exposition gegenüber UFP (Baldauf et al., 2016). Die Teilnehmerinnen und Teilnehmer resümieren, dass epidemiologische Kurzzeitstudien auf einen Zusammenhang zwischen verkehrsinduzierten Feinstaub (welcher reich an UFP ist) und adversen kardiovaskulären Gesundheitsendpunkten hinweisen. Jedoch können beobachtete adverse Gesundheitseffekte durch UFP nicht zuverlässig von einer potentiellen Mitwirkung weiterer Luftschadstoffe separiert werden. Ähnlich wie das HEI fassen Baldauf et al. (2016) zusammen, dass der aktuelle Forschungsstand keine ausreichende Evidenz liefert, dass UFP toxischer sind als andere Partikelfraktionen. Nichtsdestotrotz liefern toxikologische Erkenntnisse Hinweise bezüglich potentieller Gesundheitseffekte durch UFP, was es nötig macht, die Partikelgröße bei der Erfassung adverser Effekte durch Feinstaubexpositionen zu berücksichtigen.

Chen et al. (2016) betrachten umfassend Artikel zur Zusammensetzung von UFP, deren Quellen, typische Eigenschaften, oxidative Effekte und potentielle Expositionswege mit einem Hauptfokus auf toxikologischen Studien. Des Weiteren berücksichtigen sie die Evidenz aus dem Bereich der Nanotechnologie, was das Verständnis bezüglich toxischer Mechanismen luftgetragener UFP erweitert. Die Autoren resumieren, dass UFP einen bedeutenden Einfluss auf die menschliche Gesundheit haben.

Eine amerikanische Arbeitsgruppe (Li et al., 2016) nehmen eine Neubewertung die Schlussfolgerungen des HEI-Berichts vor und untersuchen experimentelle, epidemiologische und klinische Studien, welche 2014 und 2015 publiziert wurden. Die Autorinnen und Autoren benennen eine kritische Wissenslücke in Bezug auf Effekte von UFP auf die menschliche Gesundheit. Neuere Studien, insbesondere experimenteller und toxikologischer Art, stellen die Validität der HEI-Schlussfolgerungen in Frage, dass die Evidenz in Bezug auf Gesundheitseffekte durch UFP im Vergleich zu $PM_{2.5}$ keine radikalen Unterschiede belegt. In Bezug auf epidemiologische Studien sehen Li et al. (2016) keine neuen Erkenntnisse.

Heinzerling et al. (2016) untersuchen UFP-bedingte respiratorische Gesundheitheitseffekte bei Kindern anhand 12 relevanter Artikel. In Ein-Schadstoff-Modellen waren UFP mit inzidenter keuchender Atmung („wheezing“), bestehendem Asthma, eingeschränkter Lungenfunktion und durch Asthmaanfälle ausgelösten Besuchen von Notfallambulanzen assoziiert. Nur eine der Studien (Halonen et al., 2008) adjustierte für weitere Luftschadstoffe, woraufhin die Effektschätzer nicht länger signifikant waren. Die Autoren schlussfolgern, dass trotz einer Zunahme der Evidenz bezüglich UFP und der respiratorischen Gesundheit bei Kindern, die Evidenzlage uneindeutig bleibt.

Zusätzlich publizierten Clark et al. (2016) eine Studie, die auf biologische Mechanismen kardiovaskulärer Effekte über die alveoläre Barriere hinaus im Körper oder Gewebeproben, welche UFP und quasi-UFP mit einer Größe bis 500 nm exponiert waren. Die Autoren schlussfolgeren, dass eine mögliche bis hin zu starker Evidenz für verschiedene kardiovaskuläre Gesundheitsendpunkte besteht.

Die in den letzten Jahren deutlich zugenommene Anzahl an wissenschaftlichen Publikationen macht nun eine Neubewertung der Evidenzlage notwendig.

Ziele der Studie

Die Ziele dieses Projekts sind die Durchführung einer systematischen Literaturrecherche zu den gesundheitlichen Effekten von Ultrafeinstaub, eine Bewertung der identifizierten Literatur und eine Bewertung der Übertragbarkeit der Ergebnisse auf die Situation in Deutschland. Zu diesem Zweck sollten folgende Fragen beantwortet werden:

a. Systematische Literaturrecherche
    a. Zu gesundheitlichen Effekten von Ultrafeinstaub
    b. Fokus auf epidemiologischen Studien und quantitativen Effektmaßen (z. B. Relative Risiken, Konzentrations-Wirkungsfunktionen)
    c. Dokumentation der Suche und Archivierung der berücksichtigten Artikel in einem Literaturverwaltungsprogramm (vorzugsweise Endnote)
b. Bewertung der identifizierten Literatur
    a. Bewertung der Studienqualität anhand festzulegender Kriterien
    b. Bewertung der Übertragbarkeit der identifizierten Erkenntnisse auf die Verhältnisse in Deutschland
c. Bewertung der gesundheitlichen Relevanz von Ultrafeinstaub
    a. Mit Bezugnahme auf weitere Luftschadstoffe (z. B. PM_{10}, $PM_{2,5}$, Ozon, Stickstoffdioxid)
    b. Im Hinblick auf die Situation in Deutschland
    c. Unter Berücksichtigung der wahrscheinlichen Entwicklung von Ultrafeinstäuben in Deutschland

Methoden

Es wurde eine systematische Literaturrecherche nach Studien zu Gesundheitseffekten von Außenluft-bezogenen UFP in der MEDLINE-Datenbank (Medical Literature Analysis and Retrieval System Online) durchgeführt. Die Suche umfasste alle im Zeitraum vom 01.01.2011 bis zum 11.05.2017 veröffentlichten Studien. Zusätzlich suchten wir in der vom Schweizer Tropen- und Public Health Institut (Swiss TPH) zur Verfügung gestellten LUDOK (Dokumentationsstelle Luftverschmutzung und Gesundheit)-Datenbank. Diese Datenbank umfasst Fachliteratur zu den Effekten von Luftverschmutzung auf die menschliche Gesundheit.

Der Schwerpunkt der Recherche lag auf epidemiologischen Studien zu gesundheitlichen Effekten von Ultrafeinstäuben mit quantitativen Effektmaßen (Arbeitspaket 1). Die Studien sollten zudem mindestens eines der folgenden UFP-Maße enthalten: Anzahl (PNC) für Partikel mit einer Größe von maximal 100 nm, $PM_{0,1}$, nucleation-mode, Aitken-mode sowie quasi-UFPs: PNC für Partikel mit einer maximalen Größe von über 100 nm, $PM_{0.25}$, surface area concentrations und accumulation mode. Als gesundheitliche Endpunkte wurden neben Mortalität, Morbidität und Krankenhauseinweisungen/ Ambulanzbesuchen auch präklinische Endpunkte berücksichtigt.

Toxikologische Studien wurden lediglich im Hinblick auf unterstützende Evidenz für die im Arbeitspaket 3 (Bewertung der gesundheitlichen Relevanz von Ultrafeinstaub) durchzuführende Bewertung gesichtet. Ebenso wurden Studien gesichtet, welche die Erfassung von bevölkerungsbezogenen Expositionen zum Thema haben, da diese für die Bearbeitung der Teilaspekte 2b) (Bewertung der Übertragbarkeit der identifizierten Erkenntnisse auf die Verhältnisse in Deutschland) und der Teilaspekte 3b) und 3c) (Bewertung der gesundheitlichen Relevanz von Ultrafeinstäuben im Hinblick auf die Situation in Deutschland und unter Berücksichtigung der wahrscheinlichen Entwicklung von Ultrafeinstäuben in Deutschland) notwendig sind.

Suchstrategie

Wir führten eine kombinierte Suche in MEDLINE, LUDOK und eine Handrecherche durch. Die MEDLINE-Suche basierte auf dem letzten umfassenden Review zu Gesundheitseffekten ambienter UFP, welcher das HEI im Jahr 2013 veröffentlicht hat. Der Review umfasste Suchergebnisse der Literaturdatenbanken MEDLINE und Web of Science bis Mai 2011. Innerhalb des hier vorliegenden Projekts wurde die für den HEI-Review gewählte Suchstrategie in MEDLINE repliziert und einzelne Aspekte angepasst (Verweis auf die Suchstrategie im Anhang). Der Startpunkt unserer Suche wurde ein halbes Jahr vor dem Endpunkt der HEI-Suche gesetzt, um auch Publikationen zu erfassen, die zum Zeitpunkt der HEI-Suche noch nicht indiziert waren.

Daneben wurde eine alternative Suchstrategie in MEDLINE mit spezifischen Gesundheitsendpunkten eingesetzt. Statt der allgemeinen Suchtermini „health“ und „epidemiology/ic/ical“ enthielt diese spezifische Krankheitsendpunkte.

Die LUDOK-Datenbank umfasst epidemiologische und experimentelle Originalarbeiten über die Auswirkungen der „klassischen“ Aussenluftschadstoffe auf Menschen, sowie von weiteren Schadstoffen, die via Luft auf die Allgemeinbevölkerung einwirken (d. h. keine alleinig arbeitsmedizinisch relevanten Stoffe) In LUDOK wird monatlich eine Recherche über PubMed mit gleich bleibender, sehr breiter Formulierung durchgeführt. Zusätzlich zur regelmäßigen Suche wird eine intensive Handsuche in über 20 relevanten Fachzeitschriften, allgemein wichtigen Journals sowie den Referenzlisten aus Publikationen durchgeführt. Die Suchstrategie innerhalb des hier vorliegenden Projekts bestand aus einer modifizierten HEI Suchstrategie, ergänzt um eine Suche innerhalb der LUDOK-Datenbank sowie Handssuchen. Die Suchtermini wurden im Vergleich zu den HEI-Termini in Anlehung an die breite LUDOK-Suchstrategie erweitert.

Weitere Zugangswege zu Publikationen bot die Handrecherche in den vorhandenen Übersichtsarbeiten der letzten sechs Jahre sowie Übersichtsarbeiten, die im Rahmen unserer Literaturrecherche identifiziert wurden.

Studienselektion

Zwei Reviewer prüften Titel, Abstracts, sowie nach Bedarf Volltexte der Studien auf die Ein- und Ausschlusskriterien (s. u.) hin. 10 % der Studien wurden doppelt bewertet. Falls Unsicherheiten bzgl. der Selektion einer Studie bestanden, wurden diese im Team besprochen, evaluiert und bei Bedarf die Ein- und Ausschlusskriterien angepasst. Der Prozess der Studienselektion wurde in einem angepassten „Preferred Reporting Items for Systematic Reviews and Meta-Analyses“ (PRISMA)-Diagramm dokumentiert.

Einschlusskriterien

- Epidemiologische Studien mit einem geeignete Studiendesigns: Querschnitt-, Fall-Kontroll-, Kohorten, Zeitreihen-, Panel-, und Case-crossover-Studien, scripted-exposure Studien
- Quantifizierbare Assoziationsmaße mit mindestens einem UFP-Maß: Anzahl (PNC) oder größenfraktionierte PNC für Partikel <100 nm, $PM_{0.1}$, nucleation-mode, Aitken-mode oder einem quasi-UFP-Maß: $PM_{0.25}$, surface area concentrations, accumulation-mode
- Gesundheitliche Endpunkte: Mortalität, Morbidität, Krankenhauseinweisungen/ Notfalleinweisungen, präklinische Endpunkte
- Quantifizierbare Effekte mit mindestens einem der folgenden Maße: Odds ratio, Relatives Risiko, Hazard ratio, β-Schätzer, prozentuale Veränderung oder Expositions-Wirkungsfunktionen.
- Sprachen: Englisch, Deutsch
- Zeitrahmen: Studien, die zwischen dem 01.01.2011 bis zum 11.05.2017 publiziert wurden und nicht im HEI-Review enthalten sind. Studien, die nach diesem Zeitraum publiziert wurden, sind im Anhang gelistet

Ausschlusskriterien

- Toxikologische Studien, kontrollierte Expositionsstudien, (Tier-)Experimente, in-vitro Studien
- Exposition ggü. Nanopartikeln, die über industriell gefertigte Produkte in die Umwelt gelangen.
- Exposition ggü. UFP oder Nanopartikeln am Arbeitsplatz.
- Exposition ggü. Innenraum generierte UFP mit Quellenbezug
- Expositionen beschränkt auf Dieselpartikel, BC, EC
- Expositionen beschränkt auf Entfernungsmessungen
- Gesundheitliche Endpunkte unklarer gesundheitlicher Bedeutung wie Epigenetik, Metabolomics, Methylierung

Alle Referenzen werden in einer Bibliothek des Literaturverwaltungsprogramms Endnote verwaltet.

Datenextraktion

Die identifizierten Studien wurden hinsichtlich ihrer Qualität in Bezug auf die Berichterstattung, der Rigorosität/Aussagekraft und der inhaltlichen Aussage, und bezüglich ihrer Übertragbarkeit auf die Verhältnisse in Deutschland bewertet. Die entwickelten Qualitätskriterien sind angelehnt an das Quality Assessment Tools des National Heart, Lung and Blood Institute des National Institute of Health (2014). Besondere Aufmerksamkeit bei der Entwicklung des Erhebungsinstrumentes zur Datenextraktion wurden der Erfassung der Exposition gewidmet. Insbesondere wurden Kriterien für die Beurteilung der angewandten Messtechnik, der Repräsentativität der Messorte für die Exposition der Zielbevölkerung, die Modellgüte von genutzten Expositionsmodellen sowie für die Erfassung/ Modellierung mehrerer Luftschadstoffe entwickelt.

Ergebnisse

Literatursuche

Die Anwendung der Haupt-Suchstrategie in MEDLINE ergab 1.114 Referenzen. Die alternative Suchstrategie ergab 992 Referenzen, von welchen 332 Referenzen noch nicht in den Ergebnissen

der Haup-Suchstrategie enthalten waren (Abbildung II). Insgesamt erzeugten die beiden MEDLINE Suchen 1.446 Treffer. Die Suche in der LUDOK-Datenbank ergab 106 Treffer, von welchen 30 nicht in der MEDLINE-Suche enthalten waren. Weitere acht Referenzen wurden durch Handsuchen in weiteren Quellen generiert. Insgesamt ergab die kombinierte Suchstrategie 1.484 Referenzen.

Von der finalen Anzahl von 85 Studien wurden 70 über die Haupt-Suchstrategie und 3 Studien über die alternative Suchstrategie in MEDLINE generiert. Hinzu kamen acht weitere Studien aus der LUDOK-Datenbank und vier Studien aus zusätzlichen Quellen.

Eine Replikation der MEDLINE Suchstrategie am 23.02.1018 für den Zeitraum nach der initialen Suche ergab weitere 13 Studien, die im Anhang gelistet sind.

Studiencharakteristika

Die meisten eingeschlossenen Studien wurden in Nordamerika (n=37) oder Westeuropa (n=27) durchgeführt. Weitere 12 Studien fanden in der West-Pazifik-Region statt. Nur sehr wenige Studien wurden in Mittel-/Südamerika (n=1), Osteuropa (n=2) und Südostasien (n=1) durchgeführt. Drei von fünf multizentrische Studien schlossen Studien ein, die in verschiedenen westeuropäischen Ländern (Karakatsani et al., 2012; Manney et al., 2012; Samoli, Andersen, et al., 2016) durchgeführt wurden, zwei multizentrische Studien beinhalteten Studienstandorte in West- und Osteuropa (Lanzinger et al., 2016a, 2016b).

Die Mehrzahl der eingeschlossenen Studien bezogen sich auf die Untersuchung von Kurzzeit-Effekten (n=75) mit Gesundheitsendpunkten, die innerhalb von Stunden bis hin zu Wochen nach der Exposition gemessen wurden. Die Kurzzeitstudien waren dominiert von Panelstudien (31 mit wiederholten Messungen in eine in einem Querschnittsdesign), „Scripted exposure"-Studien (n=16) und Zeitreihenstudien (n=11). Weitere Studienarten der Kurzzeitstudien waren „case-crossover"-Studien (n=8), Kohortenstudien (n=4) und Querschnittsstudien (n=4). Zehn Studien untersuchten Langzeit-Assoziationen und nutzten hierbei Expositionsschätzungen für Zeiträume von Monaten bis Jahren. Die Studien mit einem Langzeitstudiendesign bestanden aus Kohortenstudien (n=4), Querschnittsstudien (n=4), jeweils einer Fall-Kohorten und Fall-Kontrollstudie (Tabelle 1).

Tabelle 1: Studiendesigns unterschieden nach Langzeit-/Kurzzeitstudien

Design	Studienanzahl	%
Langzeit	gesamt=10	
Fall-Kohortenstudie	1	1,2%
Fall-Kontrollstudie	1	1,2%
Kohortenstudie	4	4,7%
Querschnittsstudie	4	4,7%
Kurzzeit	gesamt =75	
Kohortenstudie	4	4,7%
Querschnittsstudie	4	4,7%
Panelstudie (Querschnitt)	1	1,2%
Panelstudie (wiederholte Messungen)	31	36,5%

Case-crossover	8	9,4%
Scripted exposure	16	18,8%
Zeitreihenstudien	11	12,9%
Gesamt	**85**	**100,0%**

Insgesamt nutzten die meisten Studien messbasierte Expositionserfassungen (87,1%). Modellbasierte Expositionen wurden in 10,6 % der Studien genutzt. In Langzeitstudien wurden zumeist modellbasierte Expositionen genutzt (9 von 10 Studien), wohingegen die Mehrzahl der Kurzzeitstudien messbasierte Expositionen nutzte (71 von 75). Dieses Muster ist darauf zurückzuführen, dass modellbsierte Expositionen notwenig sind, um die räumliche Variation der Exposition zu erfassen, welche den notwendigen Expositionskontrast für die Erfassung von Langzeiteffekten widergibt.

Die Mehrheit der Studien verwendete zentrale Messstationen zur Erfassung der Expositionen (n=45), gefolgt von mobilen Messtechniken (n=17) sowie Kombinationen verschiedener Modelle bzw. Messtechniken (n=10), z. B. zentrale Messstationen in Kombination mit Landnutzungsmodellen, Messungen im Wohngebiet oder kleinräumige individuelle Expositionsmodelle (Tabelle 2).

Tabelle 2: Art der Expositionsmodelle bzw. Messungen in den Studien

Expositionsmodell/Messung	Studienanzahl	%
Chemie-Transport-Modell	3	3,5%
Landnutzungsmodell	1	1,2%
Dispersionsmodell	1	1,2%
Messung: zentrale Station	45	52,9%
Messung: Wohngebiet	2	2,4%
Messung: Mobil	17	20,0%
Kleinräumiges personales Expositionsmodell	2	2,4%
Weitere	4	4,7%
Kombination verschiedener Modelle	10	11,8%
Gesamt	**85**	**100,0%**

In den meisten Studien wurden UFP als Partikelanzahlkonzentrationen (PNC) pro Volumen bestimmt. In etwa einem Drittel der Studien wurden PNC mit einer Größe von bis zu 100 nm verwendet (29 von 95[6] Studien). In 66 Sudien wurden quasi-UFP mit PNC-Fraktionen von bis zu

[6] Da viele Studien mehrere Größenfraktionen nutzten, übersteigt die Summe der Studien hier 85.

3.000 nm Größe genutzt. In Bezug auf die verschiedenen Größenmodi werwendeten nur wenige Studien nucleation-mode Partikel (n=1), Aitken-mode Partikel (n=1), oder accumulation-mode Partikel. Elf Studien nutzten Partikelmassen-Konzentrationen pro Kubikmeter. Sechs Studien schätzten submikrone $PM_{0.1}$-Partikel, sieben Studien erfassten quasi-UFP $PM_{0.25}$ oder $PM_{0.1}$ Partikel. Die Oberflächenkonzentration, gemessen als *„lung deposited surface area"* (LDSA) wurde nur in zwei Studien verwendet.

Tabelle 3: Arten von Gesundheitsendpunkten in Langzeit- und Kurzzeitstudien

	Studienanzahl	%
Langzeit	gesamt=10	
Mortalität	1	1,1%
Morbidität	4	4,5%
Krankenhauseinweisung	0	0,0%
Subklinisch[7]	5	5,7%
Kurzzeit	gesamt=78	
Mortalität	7	8,0%
Morbidität	5	5,7%
Krankenhauseinweisung	11	12,5%
Subklinisch	55	62,5%
Total	88	100,0%

Acht Studien analysierten UFP in Zusammenhang mit Gesamtmortalität, kardiovaskulärer oder respiratorischer Mortalität. Neun Studien analysierten Effekte auf kardiovaskuläre, respiratorische oder weitere Morbiditätsbezogene Gesundheitsendpunkte. Elf Studien untersuchten Effekte von UFP auf Krankenhauseinweisungen/Ambulanzkontakte aufgrund von kardiovaskulärer oder respiratorischer Erkrankungen. Die große Mehrheit der Studien untersuchte zahlreiche subklinische Messungen als Gesundheitsendpunkte. Unterteilt nach Organsystemen, untersuchte die Mehrheit der Studien kardiovaskuläre Gesundheitsendpunkte, gefolgt von Entzündungsmarkern und respiratorischen/atopischen Gesundheitsendpunkten. Insgesamt untersuchten nur wenige Studien Gesamtmortalität und oxidativen Stress.

Qualitätsindikatoren

In mehr als der Hälfte der Studien (n=49) wurden „Convenience"-Stichproben genutzt, sechs Studien nutzten zufällig gezogene Stichproben und weitere sieben Studien nutzte eine Kombination beider Stichprobenarten. In 13 Studien repräsentierten die Studienteilnehmerinnen und -teilnehmer die allgemeine Bevölkerung, die Mehrheit der Studien (n=62, 72,9%) verwendete ausgewählte Gruppen, welche nicht die Allgemeinbevölkerung repräsentieren. Eine Berechnung der Stichprobengröße war selten angegeben (n=3).

[7] Subklinische Endpunkte bezeichnen biologische Messungen wie z.B. der Lungenfunktion, des Blutdrucks, der Herzratenvariabilität, der Atherosklerose und Arrhythmie sowie die Untersuchung von Körperflüssigkeiten, beispielsweise zu systemischen oder lungenspezifischen Entzündungsmarkern.

Die Mehrheit der Studien (n=66, 77,6%) nannte das Größenspektrum der gemessenen UFP. Nahezu alle Studien (n=79, 92,9%) nannte das Messgerät zur Partikelmessung. Weniger als die Hälfte (n=34) der Studien, welche weitere Luftschadstoffe erfassten (n=79), adjustierte für weitere Schadstoffe in Mehrschadstoff-Modellen. Diese Studienwurden im Rahmen der Qualitätsbewertungmit einem hohen Risiko für Verzerrung bewertet. Mit Ausnahme von einer Studie wurden die zugewiesenen Expositionswerte vor oder parallel zur Erfassung der Endpunkte gemessen. In fünf der eingeschlossenen Langzeitstudien wurde dies durch die Anwendung von Chemietransport-Modellen erreicht, welche die Abschätzung von täglichen Schadstoffkonzentrationen in spezifischen Zeitperioden ermöglichen. Des Weiteren wurden die Endpunkte in allen bis auf eine Studie (n=84) klar beschrieben und definiert. In 68 Studien konnte eine Verblindung der Erfasserinnen und Erfasser der gesundheitlichen Effekteangenommen werden.

Akute Gesundheitseffekte

Mortalität

Im Vergleich zur bisherigen Evidenzbasis wurden sieben zusätzliche Studien zur **Gesamtmortalität** mit unterschiedlichen Ergebnissen durchgeführt (Tabelle 4). Im Bereich der Gesamtmortalität fanden nur zwei von vier Studien positive Schätzer in Einschadstoff-Modellen. Von diesen zeigte nach Adjustierung für weitere Schadstoffe nur eine Studie positive Assoziationen für quasi-UFP, wohingegen in der anderen Studie die erhöhten Effektschätzer gegen null tendierten.

Die Evidenz bezüglich **respiratorischer Mortalität** ist ebenfalls sehr begrenzt und inkonsistent: Von den fünf Studien zu respiratorischer Mortalität beobachteten vier Studien positve, wenn auch zumeist nicht-signifikante Assoziationen für UFP oder quasi-UFP. Drei der Studien adjustierte für weitere Luftschadstoffe, wobei die Adjustierung für NO_2 gegensätzliche Effekte zeigte: Teilweise wurden die Schätzer erhöht, teilweise verringert. Die Studien präsentierten lediglich die Mehrschadstoff-Modelle für die jeweiligen Modelle/ Zeitfenster/ Größenfraktionen mit der stärksten Assoziationen. Daher können die unterschiedlichen Effektschätzer nur sehr eingeschränkt verglichen werden bzw. die Konsistenz der Ergebnisse nur sehr eingeschränkt bewertet werden.

Die Evidenz bezüglich **kardiovaskulärer Mortalität** ist ähnlich inkonsistent. Die sechs Studien zu diesem Endpunkt zeigten sowohl positive (n=3) als auch entgegengesetzte Assoziationen (n=3). In den zwei Studien mit Mehrschadstoffmodellen führte die Adjustierung für NO_2 zu verringerten Effektschätzern, was in einer Studie zu einem Verlust der Signifikanz führte und in einer anderen Studie zu einer signifikant inversen Assoziation. Adjustierungen für $PM_{2.5}$ führte nur zu geringen Veränderungen der UFP-Schätzer.

Die Evidenz aus dieser wie auch vorheriger Reviews weist darauf hin, dass die Effekte in der warmen Jahreszeit größer sind; daher sollte in zukünftigen Kurzzeitstudien eine mögliche Effektmodifikation durch die Jahreszeit unbedingt berücksichtigt werden. Darüberhinaus überlappen die beobachteten Effekte zumindest teilweise mit den Effekten weiterer Luftschadstoffe, was am deutlichsten für NO_2 beobachtet werden kann. Aufgrund von Unterschieden bei den erfassten Partikelfraktionen kann keine Aussage zu den relevantesten Fraktionen gemacht werden.

Tabelle 4: Zusammenfassung der Analysen in sieben Studien zur Mortaliät

Studie	Alle Ursachen	Ein-Schadstoff-	Zwei-Schadstoff-	Respiratorisch	Ein-Schadstoff-	Zwei-Schadstoff-	Kardiovaskulär	Ein-Schadstoff-	Zwei-Schadstoff-

		Assoziation	Assoziation		Assoziation	Assoziation		Assoziation	Assoziation
Lanzinger et al., 2016a	✓	0	0	✓	(+)	+	✓	(-)	-
Leitte et al., 2012				✓	UFP: (+), quasi-UFP: +	UFP: 0 quasi-UFP: (+)			
Meng et al., 2013, (only quasi-UFP)	✓	+	+	✓	(+)	nc	✓	+	nc
Samoli et al., 2016	✓	0	0	✓	-	-	✓	(-)	nc
Stafoggia et al.,2017	✓	(+)	(-)	✓	+	nc	✓	(-)/(+)*	nc
Su et al., 2015							✓	+	(+)
Wolf et al., 2015							✓	(+)	nc

0 bezeichnet keine Assoziation. (+) und (-) bezeichnen primär nicht-signifikante Assoziationen, + und - bezeichnen signifikante Assoziationen. Nc: nicht durchgeführt. *variieren je nach Zeitfenster.

Morbidität

Von den wenigen Studien, die Kurzzeiteffekte von UFP/quasi-UFP in Bezug auf verschiedene Maße der **Morbidität** untersuchten, beobachteten ledigleich zwei Studien erhöhte Schätzer in Zusammenhang mit einem Indikator zu wahrgenommenem Stress und mit unterschiedlichen Symptomen. Da keine dieser Studien für weitere Schadstoffe adjustiert hat, konnten die Effekte verschiedener Komponente des Luftverschmutzungsgemisches nicht separiert werden. Daher ist ein Fazit zu unabhängigen Effekten von UFP/quasi-UFP auf die Morbidität nicht möglich. Die Evidenz für kardiovaskuläre Morbidität ist eingeschränkt mit nur zwei Studien zu unterschiedlichen Endpunkten. Diese Studien weisen darauf hin, dass Teilnehmende mit bestehender kardiovaskulären Erkrankung möglicherweise empfindlicher gegenüber UFP/quasi-UFP sind. Auch wenn beide Studien generell auf positive Assoziationen hinweisenn, kann aufgrund fehlender Mehrschadstoffmodelle keine Aussage zu unabhängigen Effekten von UFP/quasi-UFP gemacht werden. Die Evidenz zu Assoziationen mit akuten Veränderungen von Symptomen psychischer Gesundheit ist unzureichend.

Die Evidenzbasis für UFP-bedingte Effekte auf die **Inanspruchnahme der ambulanten und stationären Gesundheitsversorgung (Krankenhausaufnahmen, Notfallambulanzen)** aufgrund respiratorischer Symptome ist limitiert (Tabelle 5). Mögliche Assoziationen scheinen am wahrscheinlichsten bei Kindern als vulnerable Subgruppe. Während Einschadstoff-Modell-Assoziationen in einigen Studien beobachtet werden konnten, war dies nicht der Fall für Mehrschadstoff-Modelle. Insbesondere die Adjustierung für NO_2 führte zu verringerten Schätzern bis hin zum Nulleffekt.

Die meisten Studien untersuchten kardiovaskulär-bedingte Krankenhausaufnahmen. Diese weisen auf einen stärkeren Zusammenhang für kürzere Zeitfenster bis zu 24 Stunden hin. Diese Assoziationen wurden nach Adjustierung für weitere Schadstoffe schwächer und zeigten keine klare Evidenz mehr für UFP/quasi-UFP-Assoziationen.

Tabelle 5: Gesamttabelle zu Analysen in 11 Studien zu Krankenhauseaufnahmen/Ambulanzkontakte

Studie	Respiratorisch	Ein-Schad-	Zwei-Schad-	Kardiovaskulär	Ein-Schad-	Zwei-Schad-

		stoff-Assoziation	stoff-Assoziation		stoff-Assoziation	stoff-Assoziation
Evans et al., 2014	✓	(+)	(+) (no NO_2 adjustment)			
Gardner et al., 2015				✓	(+)/0	nr
Iskandar et al., 2012	✓	(+)	0			
Rosenthal et al., 2013				✓	(+)/+	0
Wichmann et al., 2013				✓	(+)/0	nc
Delfino et al., 2014	✓	nr	nc			
Diaz-Robles et al., 2014	✓	+				
Lanzinger et al., 2016	✓	(+)	0	✓	(+)/0	0
Samoli UK, 2016	✓	(+)/(-)	(+)	✓	(+)	(-)/(+)
Samoli EU, 2016	✓	(+)/(-)	(-)/-			
Liu et al., 2013				✓	+/(+)	nc

0 bezeichnet keine Assoziation. (+) und (-) bezeichnen primär nicht-signifikante Assoziationen, + und - bezeichnen signifikante Assoziationen. Nc: nicht durchgeführt, nr: nicht berichtet.

Subklinische Endpunkte

Die Mehrzahl der 11 Studien zu **subklinischen respiratorischen Endpunkten** (Tabelle 6) verfügen nur über eine geringe Stichprobengröße (15 bis 84 Teilnehmende). Darüber hinaus waren die Stichproben zumeist selektiv und repräsentierten entweder junge, gesunde Erwachsene oder Personen die unter Atopie/Asthma leiden. Die untersuchten Zeitfenster und Mittelungsperioden variieren je nach Studie, wobei die meisten Assoziationen in einem Zeitfenster von 0 bis 48 Stunden im Zusammenhang mit erhöhter Exposition beobachtet wurden. Letztendlich waren die Ergebnisse der meisten Studien inkonsistent in Bezug auf die einzelnen respiratorischen Endpunkte. Im Hinblick auf Peak-flow Endpunkte sind Messfehler aufgrund von selbst-erfassten Messwerten möglich, insbesondere in der Studie von Cole-Hunter et al. (2013), die nicht verblindet werden konnte. Trotz beobachteter Assoziationen in Einschadstoffmodelle in vier der 11 Studien kann aufgrund von fehlender Adjustierung für weitere Schadstoffe in Bezug auf unabhängige Effekte in den meisten Studien keine Schlussfolgerung gezogen werden. Studien mit Zwei-Schadstoffmodellen weisen auf eine zumindest teilweise Überlappung von UFP, bzw. PNC Effekten mit NO_2-Effekten hin.

13 Studien beobachtete adverse Assoziationen zwischen UFP/quasi-UFP-Expositionen und **Blutdruck-Indizes**, d. h. sie wiesen auf erhöhte Blutdruckwerte hin. Die Ergebnisse variierten je nach Endpunkt (systolischer, diastolischer, Pulsdruck), nach Größenfraktionen und untersuchten Zeitfenstern. Abgesehen von einer Studie mit mehr als 1.000 Teilnehmenden, bestanden die Studien aus kleineren, überwiegend selektiven Studienpopulationen. Die Evidenz aus Studien mit Mehrschadstoff-Modellen ist zu gering um Schlussfolgerungen im Hinblick auf unabhängige UFP Effekte auf Blutdruck-Indizes zu ziehen.

Für **Herzratenvariabilität** (HRV) ist eine relativ große Zahl von 16 Studien verfügbar, von denen 12 Studien für mindestens einen HRV-Endpunkt einen Zusammenhang beobachteten. Nach

Adjustierung für weitere Luftschadstoffe veränderten sich die Assoziationen je nach Studie in beide Richtungen. Die einzelnen Studien nutzten verschiedene Zeitfenster und unterschiedliche Luftschadstoffe in den Mehrschadstoff-Modellen, so dass keine eindeutigen Muster beobachtet werden konnten.

In Anbetracht der begrenzten Studienanzahl zu **Arrhythmie**-Endpunkten mit nur einer vorliegenden Studie, ist die Evidenz weiterhin unzureichend.

Die Mehrzahl der sieben Studien, welche Assoziationen zwischen UFP/quasi-UFP und **vaskulärer Funktion** untersucht haben, weisen auf eine mögliche Assoziation hin. Auch hier verhindern jedoch fehlende Konsistenz im Studiendesign und insbesondere bei den Parametern des Endpunkts sowie fehlende Mehrschadstoffmodelle Schlussfolgerungen zu den gesundheitlichen Effekten.

Alle der 12 durchgeführten Studien zu **pulmonalen Entzündungsprozessen** weisen auf Assoziationen zwischen UFP und adversen Veränderungen der pulmonalen Inflammationsmarker hin, insbesondere sofort nach der Exposition. Die Evidenzbasis für pulmonale Entzündungsprozesse in Folge von UFP-Exposition ist dennoch weiterhin begrenzt, da unterschiedliche Subgruppen, Expositionsmetriken, Endpunktmessungen und Zeitfenster genutzt wurden. Die zwei Studien, welche Mehrschadstoff-Modelle angewandt haben, beobachteten insgesamt robuste Effektschätzer.

Die Mehrzahl der 18 Studien, welche UFP-Effekte auf **systemische Entzündungsprozesse** untersucht haben, deutet auf inkonsistente Assoziationen hin. Effekte auf das hochsensitive C-reaktive Protein (hs-CRP), Fibrinogen, Anzahl weißer Blutzellen, Myeloperoxidase variierten, was auf unterschiedliche Zusammensetzungen der Teilnehmenden, erfasste PNC Fraktionen und Expositionserfassung zurückzuführen ist. Die meisten Studien zeigen deutlichere Effekte für kürzere Zeitfenster zwischen Exposition und dem Auftreten systematischer Entzündungen. Eine begrenzte Anzahl an Mehrschadstoff-Modellen lässt keinen Rückschluss auf unabhängige Effekte von UFP/quasi-UFP zu, da nur zwei der fünf durchgeführten Studien mit Mehrschadstoffmodellen robuste Ergebnisse zeigten.

Tabelle 6: Gesamttabelle der durchgeführten Analysen zu subklinischen Endpunkten in 55 Studien.

Endpunkt	Studien anzahl	Studienanzahl mit Ein-Schadstoff-Assoziationen in der erwarteten Richtung	Studienanzahl mit Mehr-Schadstoff-Assoziationen in der erwarteten Richtung	Kommentare (z.B. zu Studien mit signifikanten Schätzern in der nicht erwarteten Richtung)
Respiratorische Indizes	11	4/11	3/3	Li et al. (2016) beobachtete signifikant positive Assoziationen zwischen UFP und FEV_1 und FVC
Blutdruck	13	9/13	2/4[8]	Zwei der neun Studien mit Assoziationen zeigten inkonsistente Ergebnisse je nach Zeitfenster.
HRV	16	12/16	3/5	In Zhang et al. (2013), sanken die Effektschätzer nach Adjustierung für NO_2 und stiegen nach Adjustierung für O_3.

[8] Eine von vier Studien zeigte keine Assozioationen in Einschadstoffmodellen. Eine weitere Studie (Rich et al., 2012) zeigte nicht alle Ergebnisse und wurden daher hier als nicht assoziiert dargestellt.

Arrhythmie	1	1/1	nc	Starke Assoziationen mit $PM_{0.25}$, nahezu protective Assoziationen zwischen Partikelanzahl und stündlich gemessener nächtlicher Tachykardie.
Vaskuläre Funktion	7	4/7	1/2	
Entzündungsprozesse in der Lunge	12	12/12	2/2	Die Mehrzahl der Studien untersuchten den Endpunkt FeNO
Systemische Entzündungsprozesse (inkl. Fibriongen	18	7/18[9]	2/5	Signifikant inverse Assoziationen zwischen Fibrinogen und PNC nach Adjustierung für NO_2 (Strak et al., 2013)
Neurokognitive Endpunkte	2	1	nr	-

HRV: Herzratenvariabilität, FEV_1: Forciertes Lungenvolumen in einer Sekunke, FVC: forcierte Vitalkapazität

Chronische Gesundheitseffekte

Aufgrund einer geringen Anzahl an Studien, unterschiedlichen Endpunkte und Expositionserfassungsmethoden sowie fehlende Mehrschadstoffmodellen ist es nicht möglich, finale Schlussfolgerungen zu chronischen Gesundheitseffekten von UFPs zu ziehen. Eine Übersicht der Studienergebnisse ist in Tabelle 7 dargestellt.

[9] Die Mehrzahl der positiven Assoziationen bezieht sich auf Fibrinogen.

Tabelle 7: Gesamttabelle der zehn Langzeitstudien

Endpunkt/ Studie	Endpunkt	Ein-Schadstoff-Assoziationen	Mehr-Schadstoff-Assoziationen
Mortalität/ Ostro et al., 2015	- Gesamtmortalität - kardiovaskular/ IHD - pulmonal	0 (+)/0 0	nc nc nc
Morbidität / Li et al., 2017 Laurent et al., 2014/2016b Laurent et al., 2016a	- kardiometabolisch - geringes Geburtsgewicht - vorzeitige Geburt	(+) +/(+) -/+	nc nc nc
Subklinisch/ Aguilera et al., 2016 Viehmann et al., 2015 Lane et al., 2015 Lane et al., 2016 Sunyer et al., 2016	- karotid-intima-media Dicke (PNC/LDSA) - hs-CRP/ Fibrinogen/ WBC - hs-CRP/ IL-6 - hs-CRP/ IL-6/ TNRFIII/ Fibrinogen - Arbeitsgedächtnis, - übergeordnetes Arbeitsgedächtnis - Unaufmerksamkeit	+/+ (+)/+/(+) (+)/(+) (+)/(+)/(+)/(-) (+) + +	-/(+) nc nc nc nc

0 bezeichnet keine Assoziation. (+) und (-) bezeichnen primär nicht-signifikante Assoziationen, + und - bezeichnen signifikante Assoziationen. Nc: nicht durchgeführt.

Zusammenfassung der akuten und chronischen Gesundheitseffekte

Eine Übersicht über alle eingeschlossenen Kurz- und Langzeitstudien spiegelt die Inkonsistenz der Ergebnisse wider (Tabelle 8). Mehr als die Hälfte (n=49) der Studien zu Kurzzeiteffekten (n=79) berichtete zumindest einen signifikanten Effekt in Einschadstoff-Modellen, insbesondere Studien zu Mortalität oder subklinischen Endpunkten. Bei mehr als der Hälfte der Einschadstoff-Assoziationen (21 von 49) war das generelle Muster der Assoziationen konsistent - unabhängig vom Signifikanziveau. 18 von 32 Studien, die Mehrschadstoff-Modelle angewendet haben, beobachteten zumindest einen signifikanten Schätzer. Von diesen war in etwa der Hälfte der Studien (7 von 18) das Muster generell konsistent. Die Assoziationen zwischen Langzeitexpositionen gegenüber UFP mit Gesundheitsendpunkten waren in Einschadstoffmodellen konsistenter (8 von 10), auch wenn es wesentlich weniger Studien waren. Es fehlen jedoch Langzeitstudien, welche für weitere Luftschadstoffe adjustieren - es wurde nur eine Studie identifiziert, welche im Mehrschadstoff-Modell keine Assoziationen zeigte.

Tabelle 8: Gesamttabelle zu den Resultaten aller eingeschlossenen Studien

Endpunkt	Einschadstoff-Modell-Assoziationen	Konsistenz des generellen Musters	Mehrschadstoff-Modell-Assoziationen	Konsistenz des generellen Musters
Kurzzeit	49/79	21/49	18/32	7/18
Mortalität	5/7	2/5	4/6	1/4
Morbidität	3/7	0/3	nc	nc
Krankenhaus-einweisungen	4/10	2/4	0/5	nc
Subklinisch	37/55	17/37	14/21	6/14

Langzeit	8/10	1/1	0/1	nc
Mortalität	1/1	1/1	nc	nc
Morbidität	3/4	nc	nc	nc
Krankenhaus-einweisungen	nc	nc	nc	nc
Subklinisch	4/5	nc	0/1	nc

Nc: nicht durchgeführt.

Diskussion

Literatursuche

Wir haben eine umfassende systematische Suche zu relevanten epidemiologischen Studien zu UFP und quasi-UFP für den Zeitraum vom 01.01.2011 bis 11.05.2017 durchgeführt. Unsere Suchstrategie setzte sich aus einer MEDLINE-Suche mittels einer Haupt- und einer alternativen Suchstrategie, einer Suche in der spezialisierten LUDOK-Datenbank sowie einer Handsuche in Reviewartikeln und Referenzlisten der identifizierten Publikationen zusammen.Die alternative Suche in MEDLINE sowie die LUDOK-Suche mit 15 von insgesamt 85 Publikationen hat einen beträchtlichen Zugewinn an Studien bedeutet. Ebenso ergab die Replikation der Suchstrategie im Februar 2018 mit 13 zusätzlichen Artikeln eine erhebliche Anzahl an Treffern. Diese relativ hohen Zahlen spiegeln die rasante Entwicklung des neu entstehenden Forschungsfeldes wider sowie den Wert einer spezialisierten Datenbank - in diesem Fall LUDOK - für eine zielgerichtete und zeitnahe Suche.

Bewertung der Relevanz von UFP für die Gesundheit

Unsere Bewertung der Relevanz von UFP auf die Gesundheit basiert auf den oben beschriebenen epidemiologischen Studien. Dabei wird zusätzlich berücksichtigt, inwiefern diese neu publizierten Studien die Evidenz des letzten umfassenden HEI-Berichts von 2013 erweitern. Insgesamt hat die epidemiologische Evidenz in den letzten Jahren erheblich zugenommen und es ist in den nächsten Jahren ein weiterer deutlicher Zuwachs an relevanten Studien zu erwarten. Derzeit befinden wir uns noch in den Anfängen der gesundheitsbezogenen Forschung zu UFP, was teilweise an den sich noch entwickelnden Methoden liegt (siehe Abschnitt unten zu Expositionserfassung).

Das HEI schlussfolgerte in seiner Übersichtsarbeit, dass die derzeitige Datenbasis an experimentellen und epidemiologischen Studien keine starken und konsistenten Rückschlüsse zu den unabhängigen Effekten von UFP auf die menschliche Gesundheit zulässt. Wesentliche Gründe für diese fehlende Evidenz, insbesondere der epidemiologischen Studien, liegen in der Schwierigkeit, die bevölkerungsbezogene Exposition gegenüber UFP sowohl für Kurz- als auch für Langzeitstudien zu erfassen. Aufgrund der ausgeprägten zeitlichen und räumlichen Variabilität von UFP führen gängige Expositionserfassungsstrategien, welche für homogener verteilte größere Partikelfraktionen entwickelt worden sind, bei der Anwendung auf UFP zu größeren Fehlern bei der Expositionserfassung. Im Hinblick darauf folgert das HEI, dass unabhängige UFP-Effekte nicht ausgeschlossenen werden können, und empfiehlt die Erforschung alternativer Expositionsmetriken, räumlicher Modellierungstechniken und statistischer Methoden.

In dieser Übersichtsarbeit werden ähnliche Studiendesign- und Endpunkt-spezifische Kategorien verwendet wie im HEI-Review, um neue Erkenntnisse in den bisherigen Wissensstand integrieren zu können. Da die Unabhängigkeit der Effekte von anderen Luftschadstoffen ein

Kernthema bezüglich der Relevanz von UFP auf die Gesundheit darstellt, fokussieren wir insbesondere auf Studien mit Mehrschadstoffmodellen.

Inkonsistenz der Ergebnisse nach Endpunkt

Vorherige Bewertungen folgerten, dass die kombinierten Ergebnisse für respiratorische wie auch kardiovaskuläre Endpunkte noch inkonsistent sind (Health Effects Institute, 2013). Wenn man die neu gewonnene Evidenz aus den Jahren 2011 bis 2017 betrachtet, hat sich dieses Bild nicht wesentlich verändert. Trotz einer wachsenden Zahl an Studien können wir kein eindeutiges Muster für respiratorische oder kardiovaskuläre Gesundheitseffekte in Bezug auf die Endpunkte Mortalität, Morbidität, Krankenhaus-/Notfalleinweisungen oder subklinische Endpunkte identifizieren. Für weitere Endpunkte wie z. B. psychische Erkrankungen, neurokognitive Funktionen oder Geburtsergebnisse ist die Evidenzbasis noch zu gering um sichere Schlüsse zu ziehen.

Auch wenn die Ergebnisse bezogen auf die unterschiedlichen Endpunktarten nicht konsistent sind, weist die Mehrzahl der elf Studien zu UFP-Kurzzeiteffekten auf einen Zusammenhang mit erhöhtem Blutdruck hin, dem Hauptrisikofaktor für kardiovaskuläre Erkrankungen. Die Evidenz der drei für weitere Luftschadstoffe adjustierten Studien ist gemischt, was die Notwendigkeit weiterer Studien mit Mehrschadstoffmodelln unterstreicht.

Das Fehlen von konsistenten Ergebnissen kann durch mehrere Faktoren erklärt werden. Diese beinhalten Unterschiede in der Expositionserfassung (siehe unten), in der Erfassung des Endpunkts, dem Studiendesign und -größe sowie der unterschiedliche Umgang mit Störfaktoren, insbesondere unterschielliche Korrekturen für weitere Luftschadstoffe (siehe unten).

Expositionserfassung

Insgesamt nimmt die Anzahl der Studien zur Exposition und Erforschung von Gesundheiteffekten durch UFP rasant zu. Ein wesentlicher Faktor, der zu dessen rasanten Zuwachs beiträgt, ist die Entwicklung neuer Messinstrumente, welche eine kostengünstigere Erfassung von UFP, z. B. mit Kondensationspartikelzählern, ermöglichen. Die Forschung ist jedoch noch in den Anfängen und neue Expositionserfassungsmethoden in epidemiologischen Studien müssen noch entwickelt und evaluiert werden.

Herausforderungen bei der Erfassung von UFP beinhalten deren hohe Variabilität in Zeit und Raum, was andere Erfassungsdesigns benötigt als für die „klassischen“ Methoden, mit welchen die räumlich homogener verteilten größeren Luftschadstoffe wie $PM_{2.5}$ und PM_{10} erfasst werden. Die hohe räumliche Variabilität ist nicht nur in Langzeitstudien zu Gesundheitseffekten, welche auf Langzeitunterschieden in der Exposition beruhen von Bedeutung, sondern ebenfalls für Kurzzeitstudien mit zentralen Messstationen. Diese Studien setzen voraus, dass zeitliche Veränderungen von Tag zu Tag in relativ großen Studiengebieten gleichmäßig stattfinden, das heißt wenn an der zentralen Messstation die Schadstoffkonzentration um einen bestimmten Wert steigt, so steigt die Schadstoffkonzentration an anderer Stelle um einen ähnlichen Wert. Diese Annahme muss für UFP nicht zutreffen, da die lokale UFP-Konzentration stärker von örtlichen Quellen abhängt, als die Feinstaubkonzentration. Wenn man davon ausgeht, dass bei der Erfassung von UFP daher im Vergleich zu anderen Luftschadstoffen größere Messfehler zu erwarten sind, ist eine systematische Verzerrung der Schätzer gegen Null in Zusammenhangsanalysen mit Gesundheitseffekten wahrscheinlich (Dionisio et al., 2014).

In Zukunft kann die Entwicklung von erweiterten Luftschadstoffmodellen, die räumliche und zeitliche Faktoren integrieren, zu einer präziseren Expositionserfassung in größeren Gebieten

beitragen. Aktuelle Chemie-Transport-Modelle wie z. B. das deutsche EURAD-Modell, benötigen eine Anpassung der Methodik mit Aufnahme spezifischer UFP-Quellen, eine Validierung von Modelloutput mit Messungen und eine erhöhte räumliche Auflösung.

Eine zukünftige Herausforderung bezüglich der UFP-Expositionserfassung sind die bisher nicht standardisierten Messgeräte und der nicht standardisierte Gebrauch von Partikelfraktionen in den Studien. Die üblicherweise verwendeten Messgeräte haben unterschiedliche untere Messgrenzen bezüglich der Partikelgröße. Da die Mehrheit der Partikel dem nucleation-mode (< 20 nm) der Partikelgrößenverteilung zugeordnet werden kann, können bereits geringe Unterschiede der unteren Messgrenze zwischen 1 und 20 nm zu erheblichen Unterschieden in der Partikelanzahlkonzentration führen. Des Weiteren beinhaltet die Beschreibung der Expositionserfassung nicht immer das exakte Größenspektrum der Partikel, was einen direkten Vergleich der Exposition der unterschiedlichen Studien behindert.

Langzeitexpositionen und Gesundheitseffekte

Im Gegensatz zur letzten umfassenden Übersichtsarbeit durch das HEI wurden zehn Studien veröffentlicht, die Langzeiteffekte von UFP auf verschiedene gesundheitliche Endpunkte untersuchen. Während die meisten dieser Studien erhöhte Punktschätzer für Assoziationen zwischen UFP und adversen Gesundheitsendpunkten fanden, adjustierte nur eine Studie für weitere Luftschadstoffe einschließlich NO_2. Adjustierung für andere Luftschadstoffe führte zu verringerten Effektschätzern bis hin zu Effektschätzern in die entgegengesetzte Richtung.

Auch wenn die gegenwärtige Evidenz keine unabhängigen Langzeit-Effekte von UFP auf Gesundheitsendpunkte zeigen, sollte dies auf keinen Fall als ein Beweis für das Fehlen eines solchen Effekts missverstanden werden. Gegenwärtige Methoden zur Erfassung von UFP-Langzeitbelastungen sind nicht gut geeignet, um die räumliche Varianz von UFP zu erfassen. Daher sind dringend weitere Studien nötig, welche innovative Methoden zur Erfassung individueller UFP-Expositionen anwenden und evaluieren. Bedeutende Anwendungsfelder für neu zu entwickelnde Methoden zur Erfassung der Langzeitexposition gegenüber UFP sind verkehrsnahe Expositionen. Dabei sollten Erhebungen der Langzeitexposition auch das neu auftretende Problem in Bezug auf Expositionen gegenüber UFP in der Umgebung von Flughäfen angehen (Hudda, Simon, et al., 2016).

Unabhängigkeit von Effekten

Die Evidenz zu unabhängigen Effekten von UFP ist insgesamt weiterhin als unzureichend einzustufen. Wir haben festgestellt, dass insbesondere neuere Studien verstärkt Mehrschadstoffmodelle durchgeführt haben, was eine positive Entwicklung darstellt (z. B.., Aguilera et al., 2016; Croft et al., 2017; Lanzinger et al., 2016a; Samoli et al., 2016; Stafoggia et al., 2017). Die verschiedenen Studien nutzen jedoch verschiedene Adjustierungen und es gibt noch keine Standardstrategie zur Adjustierung für weitere Luftschadstoffe. Derzeit scheint NO_2 einen größeren Effekt auf den UFP-Punktschätzer zu haben als andere Luftschadstoffe (z. B. Lanzinger et al., 2016a&b; Su et al., 2015; Samoli, Andersen et al., 2016, Zhang et al., 2013). Gründe hierfür sind die Überlappung der Quellen sowie die höhere Übereinstimmung der räumlichen und zeitlichen Varianz von UFP und NO_2, was zu instabilden Modellen und verzerrten Effektschätzern in Mehrschadstoffmodellen führen kann.

Übertragbarkeit der Ergebnisse auf die Situation in Deutschland

Die Übertragbarkeit der oben beschriebenen Ergebnisse auf die Situation in Deutschland wird nach folgenden Kriterien bewertet: Lokalität der identifizierten Studien, Expositionslevel gegenüber UFP und anderen luftgetragenen Schadstoffen, der Prävalenz der untersuchten Endpunkte sowie die Auswahl der Studienpopulation.

Lokalität und Exposition

Die große Mehrheit der identifizierten Studien wurde in Nordamerika (n=37, 43,5%) oder Westeuropa (n=27, 31,8%) sowie in mehreren Weltregionen (n=5, 6%) durchgeführt. Wenn wir die Studienorte der Studien mit mehreren Studienzentren berücksichtigt, beobachten wir, dass die Mehrzahl der Studienstandorte in West- und Südeuropa zu verorten sind (n=44 von 101 Studienstandorten, 43,6%). Die Konzentration der UFP variieren beträchtlich in Raum und Zeit, so dass direkte Vergleiche der Messungen zwischen einzelnen Studienstandorten großer Variabilität in Bezug auf Stunde, Tag und Jahreszeit der Messungen sowie der exakten Platzierung der Messstandorte (Verkehr, urbaner Hintergrund, regionaler Hintergrund) (Birmili et al., 2016; UFIPOLNET, 2008) unterliegen.

Im deutschen Messnetz für Ultrafeinstaub (German Ultrafine Aerosol Network; GUAN), wurden Langzeitmessungen ultrafeiner und feiner Partikel an 17 Standorten innerhalb Deutschlands, einschließlich alpiner Standorte (Zugspitze), ländlicher Standorte, Standorte urbanen Hintergrunds sowie straßennahen Messstellen durchgeführt (Birmili et al., 2016). Zu beachten ist, dass Partikel mit einer Größe von 20 bis 800 nm gemessen wurden und somit nicht die nucleation-mode Fraktionen, jedoch die accumulation mode Partikel umfassten. Vorläufige Ergebnisse der GUAN-Messungen ergaben, dass stündliche mediane Partikelanzahlkonzentrationen zwischen 900/ ml (Zugspitze) und 9.000/ml an der straßennahen Messtation in Leipzig rangieren. Stündliche durchschnittliche Konzentrationen sind etwas höher mit 1.120/ml an der Zugspitze und 10.500/ml in Leipzig. Das 95. Perzentil der Verteilung stündlicher Werte erreicht 22.400/ml in Leipzig-Mitte. Alle drei straßennahen Messstationen hatten Maximalwerte oberhalb 19.900/ml, während die Werte der urbanen Hintergrund-Standorte zwischen 10.000 und 20.000/ml rangierten. GUAN demonstiert ebenfalls die substantielle Variation der Partikelgrößenverteilung im Laufe einer Woche an sechs hauptsächlich urbanen Standorten.

Die in Westeuropa durchgeführten eingeschlossenen Studien messen typischerweise ähnliche oder höhere durchschnittliche Gesamtpartikelanzahlen. Mit der verfügbaren Information sind direkte Vergleiche nicht möglich, da die Messinstrumente unterschiedlich sind und verschiedene untere Messgrenzen haben. 16 von 27 Studien aus Westeuropa gaben als untere Messgrenze ihrer Messgeräte 10 nm oder geringer an. Einige Geräte weisen bis zu 3 nm als unterste Grenze auf. Da die Mehrzahl der Partikel eine Größe von 20 nm unterschreitet (nucleation-mode) (HEI perspectives, 2013), können geringe Unterschiede des unteren Messgrenzwertes zu beträchtlichen Unterschieden der durchschnittlichen Exposition führen. Zusätzlich variiert der obere Grenzwert beträchtlich, wobei nur wenige Studien UFP in engerem Sinne (<100 nm) untersuchen, sondern eher die Gesamtpartikelanzahl als Surrogat für UFP-Exposition nutzen. Dies stellt jedoch ein geringeres Problem dar, da die Gesamtpartikelanzahl von den Größenfraktionen unter 100 nm dominiert wird (HEI perspectives, 2013).

Im Rahmen von GUAN konnte die große Variabilität der Expositionen innerhalb Deutschlands dokumentiert werden. Ein direkter Vergleich absoluter Werte mit denen anderer Studien ist jedoch wegen der unterschiedlichen eingesetzten Messinstrumente schwierig. Die fünf Studien aus Deutschland basieren auf zentralen oder personenbezogenen Messungen (n=4) mit unteren Messgrenzen zwischen 3 und 10 nm. Diese Studien ergaben mittlere Expositionen zwischen

10.000/ml und 20.000/ml, was mit anderen Studien in diesem Review vergleichbar ist. Im Vergleich hierzu berichten die 13 Studien, welche in der westlichen Pazifik-Region oder Süd-Ost-Asien in den Metropol-Regionen von China, Südkorea oder Taiwan durchgeführt wurden, ähnliche oder minimal höhere Werte der gemessenen Partikelanzahlkonzentrationen. Die einzige Modellbasierte deutsche Studie wandte das EURAD-Chemietransportmodell an und erhielt wesentlich höhere mittlere Expositionen. Dies ist auf den Modellierungsprozess zurückzuführen, welcher den vollständigen nucleation-mode und damit auch kurzlebige Partikel kleiner 3 nm umfasst. Aus den Messwerten des GUAN Netzwerkes und den in Deutschland durchgeführten Gesundheitsstudien zu UFP schließen wir, dass das Expositionsniveau der in dieser Übersichtsarbeit betrachteten Studien zwar sehr variabel in Raum und Zeit ist, jedoch generell vergleichbar mit der Situation in Deutschland ist.

Erwartete Entwicklung der UFP-Belastung in Deutschland

Die Entwicklung der UFP-Belastung der Bevölkerung in den nächsten Jahren hängt von mehreren Faktoren ab: (1) Der Bildung und Emission dieser Partikel, (2) der räumlichen Verteilung der Bevölkerung und (3) der Konzentration von feinen Partikel in der Umgebungsluft.

Gemäß eines pan-europäischen Inventars anthropogener Partikelanzahlen ist der Straßenverkehr die bedeutendste Ursache von Emissionen in städtischen Gebieten und entlang stark befahrener Straßen (Health Effects Insitute, 2013).Verkehrsbezogene Emittenten von primären UFP sind Direkteinspritzer in Fahrzeugen, deren Anzahl in der letzten Dekade angestiegen ist und wahrscheinlich weiter ansteigen wird (Köllner, 2016). Andererseits wurden Diesel-Motoren, welche ebenfalls Partikel der ultrafeinen Größenbereiche ausstoßen, mit Partikelfiltern ausgestattet. Dadurch wurde der Ausstoß feiner Partikel beträchtlich reduziert (gemäß EURO5a auf weniger als 5 mg/km). Die EURO5b-Norm setzte erstmalig ein Limit für UFP, und zwar auf 6 x 10^{11} (European Union, 2007). Insgesamt ist aufgrund des wachsenden Verkehrs und der ansteigenden Anzahl von Stadtbewohnerinnen und Stadtbewohnern (Vallance et al., 2010) in Zufkunft mit einer zunehmenden Exposition von Bevölkerungsanteilen gegenüber verkehrsbezogenen UFP zu rechnen.

Eine weitere Quelle hauptsächlich UFPs ist der Luftverkehr. Mehrere Expositionsstudien haben erhöhte UFP-Expositionen in Windrichtung von Flughäfen weltweit berichtet (Hudda et al., 2014; Keuken et al., 2015; Masiol et al., 2017; Shirmohammadi et al., 2017). Die zunehmenden Kurzzeit-Belastungen sind zeitlich korreliert mit Flugzeugbewegungen und erreichen Konzentrationen von bis zu 50.000 Partikeln/ml (Keuken et al.. 2015) sieben km in Windrichtung des Amsterdamer Flughafens und bis zu 75.000 Partikeln/ml (Hudda et al.. 2014) acht km in Windrichtung des Flughafens in Los Angeles. Dieselben Studien zeigen, dass Langzeit-Belastungen sieben km in Windrichtung mit mehr als 200,000 betroffenen Einwohnern in der Nähe des Flughafens Schipohl/ Amsterdamum bis zu dreifach erhöht sind (Keuken et al., 2015) und bis um das vier- bis fünffache erhöht acht bis zehn km in Windrichtung in Los Angeles (Hudda et al., 2014). Ähnliche Expositionsstudien laufen in Deutschland und werden erste Informationen zur Belastung der Anwohner deutscher Flughäfen bringen. Angesichts des wachsenden Luftverkehrs werden Expositionen aufgrund von Flugzeugemissionen wahrscheinlich in Zukunft eine zunehmende Rolle spielen.

Des Weiteren beeinflusst die Konzentrationen feiner Partikel in der Umgebungsluft die UFP Konzentrationen insofern, als dass UFP mit größeren Partikeln zusammentreffen und dabei koagulieren. Eine höhere Konzentration feiner Partikel in der Umgebung wird daher die Entfernung von UFP aus der Umgebungsluft unterstützen. Bei einer Reduktion feiner Partikel werden UFP

wahrscheinlich länger in der Luft zirkulieren als in einer Umgebung mit höherer Feinstaubkonzentration.

Exposition gegenüber weiteren Luftschadstoffen

Die Höhe weiterer Luftschadstoffe ist von Bedeutung, da die meisten dieser Schadstoffe eigene Effekte auf den untersuchten Endpunkt haben. 78 der 85 identifizierten Studien (92%) erfassten die Höhe von mindestens einem weiteren Luftschadstoff, wenn auch nur 34 der Studien in ihren Analysen für mindestens einen Luftschadstoff adjustiert haben. Die Erfassung von und die Adjustierung für weitere Luftschadstoffe in den Studien ist daher nicht auf vergleichbare Art und Weise durchgeführt worden.

Die Analyse der Mehrschadstoffmodelle zeigte, dass $PM_{2.5}$ and NO_2 den größten Einfluss auf die UFP-Schätzer zu haben scheinen. Oft - jedoch nicht immer - führt die Adjustierung für NO_2 zu einer Schwächung der Assoziation zwischen UFP und dem Gesundheitsendpunkt (Leitte et al., 2012; Meng et al., 2012; Stafoggia et al., 2017; Su et al., 2015; Iskandar et al., 2012; Lanzinger et al., 2016; Rosenthal et al., 2013; Gong et al., 2014; Janssen et al., 2015; Steenhof et al., 2013). Die Adjustierung für PM_{10} and $PM_{2.5}$ schwächt die UFP-Assoziation ebenfalls in mehreren Studien, in den meisten Studien jedoch in geringerem Maße als die Adjustierung für NO_2.

Die Höhe der weiteren Luftschadstoffe, dabei insbesondere $PM_{2.5}$ and NO_2, kann innerhalb Europas mit dem Bericht der europäischen Umweltagentur "Air quality in Europe — 2017 report" (European Environmental Agency, 2017) verglichen werden. Gemäß dieses Berichts rangiert Deutschland unter den 28 Mitgliedsstaaten mit der höchsten durchschnittlichen Belastung an NO_2 (European Environmental Agency, 2017; Fig 6.1).

Ähnlich wie für UFP, können die jährlichen Durchschnittswerte der ausgewählten Messstationen keinen umfassenden Überblick über die Belastungen der Studienpopulationen der eingeschlossenen Studien widergeben, da NO_2-Konzentrationen hoher Variabilität in Raum und Zeit unterliegen. Von den 34 Studien, welche für weitere Luftschadstoffe adjustierten, wurden 15 in Westeuropa durchgeführt. Von diesen wurden drei in Augsburg/ Deutschland durchgeführt. Die übrigen Studien wurden hauptsächlich in größeren Städten in der Schweiz, den Niederlanden, Schweden und Finnland durchgeführt, welche vergleichbare Verkehrsexpositionen haben.

Daraus schließen wir, dass die Ergebnisse bezüglich einer teilweisen Überlappung der Effekte zwischen UFP und NO_2, die wir in den westeuropäischen Studien dieser Übersichtsarbeit beobachten (Iskandar et al., 2012; Janssen et al., 2015; Rosenthal et al., 2013; Stafoggia et al., 2017; Steenhof et al., 2013), ebenfalls für Deutschland zutreffen.

Prävalenz der Erkrankungen

Die Mehrzahl der in diesem Review identifizierten Studien ist in West-/Südeuropa und Nordamerika zu verorten. Die ursachenspezifischen altersadjustierten Sterberaten für alle nichtübertragbaren Erkrankungen und für respiratorische Erkrankungen im Jahr 2015 ähneln sich innerhalb der WHO-Region Amerika (inklusive Südamerika, was nicht in diesem Review eingeschlossen ist) und der WHO-Region Europa (World Health Organization, 2016b). Auf der anderen Seite unterscheiden sich die jährlichen ursachenspezifischen altersadjustierten Sterberaten für kardiovaskuläre Erkrankungen, mit einer erheblich geringeren altersspezifischen Sterberate für untere Altersklassen in Amerika (211/10,000) verglichen mit der europäischen Region (344/10.000). Der Unterschied in dieser Statistik ist primär auf die Kombination beider amerikanischer Kontinente zurückzuführen. Verglichen mit anderen europäischen Ländern und den

USA, die in diesem Review eingeschlossen wurden, hat Deutschland eine vergleichbare Verteilung der Ursachen für vorzeitige Todesfälle wie die Niederlande bei ischämischen Herzerkrankungen, Lungenkrebs, Alzheimer-Erkrankung, zerebrovaskulären Erkrankungen und chronisch-obstruktiver Lungenerkrankung (COPD). Dieses Ranking ähnelt der Krankheitsverteilung in UK, Dänemark, Schweden, Spanien und den USA stark. Darüber hinaus erforscht die Mehrheit der Studien kurzzeitige subklinische Endpunkte und innerhalb dieser Kategorie kardiovaskuläre, respiratorische und Biomarker-bezogene Endpunkte. Die Erfassung der Endpunkte in diesen Studien betrifft nicht die länderspezifischen ICD-Kodierungsrichtlinien.

Wegen der Ähnlichkeit der Verteilung von Krankheiten in Deutschland und den Ländern, in denen Studien zu Mortalitäts- und Morbiditätseffekten von UFP durchgeführt wurden, kann eine Übertragbarkeit der Ergebnisse auf Deutschland angenommen werden. Dies gilt auch für Studien mit subklinischen Endpunkten, solange keine Unterschiede in den physiologischen Markern zwischen der Studienpopulation und der deutschen Bevölkerungen zum Zeitpunkt der Ersterhebung bestehen.

Studienpopulation

Die meisten Studien dieses Reviews basieren auf selektiven Studienpopulationen (n=62, 72,9%), und nur zehn (11,8%) bzw. 13 (15,3%) Studien wurden als repräsentativ oder zumindest teilweise repräsentativ für die Allgemeinbevölkerung erachtet. Bei den Studien, welche als komplett repräsentativ für die Zielbevölkerung erachtet wurden, handelt es sich um Zeitreihenstudien, welche auf der Allgemeinbevölkerung der jeweiligen Studienregion basieren. Eine dieser Zeitreihenstudien (Diaz-Robles et al., 2014) zielte auf ausgewählte Altersgruppen innerhalb der Allgemeinbevölkerung ab. Von den übrigen Studien wählten 13 (15%) eine zufällige Stichprobe der Bevölkerung. Von den zehn Studien, die Langzeit-Effekte eruieren, basiert die Mehrzahl der Analysen auf mehreren Hunderten oder Tausenden von ausschließlich in Westeuropa oder Nordamerika lokalisierten Teilnehmenden. Von diesen zielen sechs Studien auf die Erwachsenen-Populationen eines oder beiderlei Geschlechts ab (Ostro et al., 2015), und weitere vier Studien wählten Kinder als Zielpopulation (Laurent et al., 2014, 2016a and 2016b; Sunyer et al., 2015). Unter den Kurzzeitstudien sind die Studienpopulationen zumeist hochgradig selektierte kleine Gruppen von entweder gesunden (jüngeren) Erwachsenen oder Teilnehmenden mit respiratorischen oder kardiovaskulären Erkrankungen wie Asthma, COPD, Erkrankungen der Koronararterien etc.

Schlussfolgerungen - Übertragbarkeit

Basierend auf den oben beschriebenen Kriterien, Belastungshöhen, Exposition gegenüber weiteren Schadstoffen, Basisprävalenz der Erkrankungen sowie Repräsentativität der eingeschlossenen Studienpopulationen folgern wir, dass die Gesamtergebnisse dieses Reviews mit angemessenem Vorbehalt auf die Situation in Deutschland übertragen werden können.

Wichtige Einschränkungen sind (1) der Mangel an Studien mit Adjustierung für weitere Luftschadstoffe, was insbesondere angesichts der hohen NO_2-Belastungen in Deutschland relevant ist und (2) die Wahl hochradig selektierter Gruppen in den Kurzzeitstudien, da diese oft keine spezifischen vulnerablen Bevölkerungsgruppen wie Personen mit unzureichend therapierten Erkrankungen, Neugeborenen und Kindern berücksichtigen.

<u>Gesamtfazit</u>

Die Erforschung von UFP-Gesundheitseffekten in epidemiologischen Studien nimmt schnell zu. In den letzten sieben Jahren wurden erhebliche Fortschritte gemacht, welche zwei der dringendsten offenen Forschungsfragen betreffen: Es wurden mehrere Studien zu Langzeit-Gesundheitseffekten von UFP publiziert. Zweitens wurden insbesondere in den neueren Studien Bemühungen unternommenfür weitere Luftschadstoffe zu adjustieren und unabhängige Effekte der UFP zu identifizieren.

Trotz der offensichtlichen Weiterentwicklungen in den oben genannten Bereichen hat sich das Gesamtfazit für den erforschten Zeitraum nicht erheblich von vorherigen Bewertungen verändert.

Zunächst bleibt die Evidenz zu Gesundheitseffekten für die meisten der untersuchten Endpunkte uneindeutig oder unzureichend. Von den Studien zu Mortalität und Krankenhausaufnahmen/Ambulanzkontakten ergaben die relativ wenigen Studien mit Adjustierung für weitere Luftschadstoffe gemischte Ergebnisse. Dies führt zu einer derzeit uneindeutige Evidenzlage. Was die Anzahl der Studien betrifft, ist die größte Evidenz für Studien verfügbar, die subklinische Endpunkte erforschen. Innerhalb dieser Studiengruppe zeigen Studien mit kardiovaskulären Endpunkten sowie Endpunkten zu pulmonalen und systemischen Entzündungsprozessen die konsistentesten Muster mit Assoziationen, die im Allgemeinen auf adverse Gesundheitseffekte hinweisen. Nichtsdestotrotz bleibt die Evidenz für die Unabhängigkeit der Effekte für diese Endpunkte ebenfalls limitiert, da nur wenige Studien für weitere Luftschadstoffe adjustiert haben und dies häufig zu einer Reduktion der Effekte führt.

Zweitens bleibt die Expositionserfassung der Bevölkerung aufgrund der spezifischen Eigenschaften der UFP schwierig. Die Studien, welche die Belastung mittels zentralen Messstationen erfassen, verpassen wahrscheinlich einen großen Teil der UFP-Variabilität, da räumliche Varianz nicht berücksichtigt wird. Studien, die klassische räumliche Modellierungsmodelle anwenden, benötigen die Integration von Techniken, die räumliche und zeitliche Variabilität genauer erfassen. Null-Effekte oder die Abnahme von UFP-Effekten nach Adjustierung für weitere Luftschadstoffe können zumindest teilweise mit Expositions-Missklassifizierung und Messfehlern erklärt werden. Bei der Erfassung der Exposition sollte der Messtechniken, der Größenfraktionen und der Lokalisierung der Messstationen besondere Aufmerksamkeit gewidmet werden. Auch die Berichterstattung sollte standardisierter werden, um Studienergebnisse besser vergleichen zu können.

Drittens kann die Unabhängigkeit von UFP-Effekten derzeit aufgrund der geringen Anzahlen an Studien mit Adjustierung und der oben erwähnten Einschränkungen bezüglich der Expositionserfassung für UFP nicht bewertet werden. Eine positive Weiterentwicklung ist der Zuwachs an Studien, die diesem Thema Beachtung schenken und die zeitgleiche Expositionen mit anderen Luftschadstoffen berücksichtigen.

Viertens besteht weiterhin ein dringender Bedarf an Langzeitstudien zu den Gesundheitseffekten von UFP. Die Durchführung von qualitativ hochwertigen Langzeitstudien wird eine Weiterentwicklung von Modellierungstechniken erfordern, welche sowohl räumliche als auch zeitliche Varianz berücksichtigen. Des Weiteren sollten spezifische Situationen mit Spitzenbelastungen identifiziert und detaillierter beschrieben werden um die Erfassung von Langzeitbezogenen Gesundheitseffekten zu ermöglichen. Während straßennahe Expositionen bereits als wesentliche Faktoren erkannt wurden, fehlen insbesondere Studien zu Flughafenbezogenen Belastungen, welche kürzlich in Verbindung mit erheblichen Konzentrationsanstiegen im Vergleich zur Hintergrundbelastung gebracht wurden.

Zusätzlich zu diesen allgemeinen Schlussfolgerungen folgern wir, dass die Gesamtergebnisse dieser Übersichtsarbeit mit angemessenem Vorbehalt auf die Situation in Deutschland übertragen werden kann. Wichtige Einschränkungen sind (1) der Mangel an Studien mit Adjustierung für weitere Luftschadstoffe, die angesichts hoher Belastung von NO_2 in Deutschland insbesondere von Bedeutung ist, sowie (2) der Nutzung hochgradig selektierter Bevölkerungsgruppen in den Kurzzeitstudien.

1 Background

1.1 Scientific Background

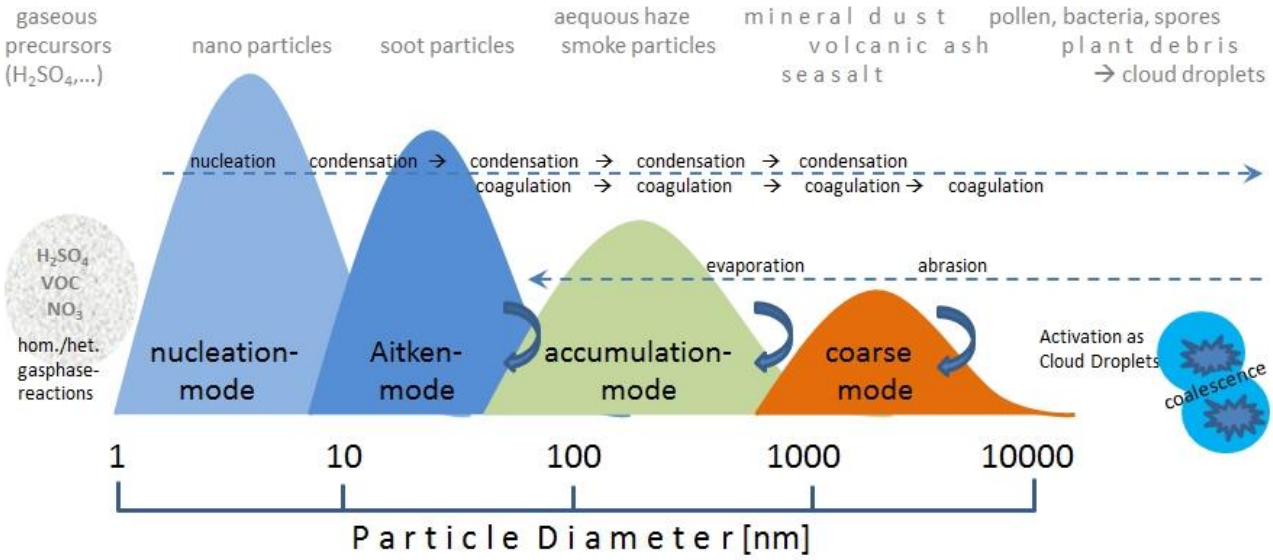
Environmental risks are important determinants of health and healthy ageing. Even if environmental risks are minor risks in relation to individual risk factors, their ubiquitous exposures may lead to a high attributable burden of disease at the population level (Cohen et al., 2017; Forouzanfar et al., 2016; World Health Organization, 2016a).

The pollution of ambient air, amongst others by particulate matter (PM, determined by different size fractions), increases all-cause mortality and has negative health impacts particularly on respiratory functions (e.g., asthma, chronic obstructive pulmonary disease (COPD), lung cancer), the cardiovascular system (e.g. myocardial infarction, stroke, elevated blood pressure) as well as on metabolic changes and early childhood development (Thurston et al., 2017). It was estimated that ambient air pollution (AAP) accounts for more than four million premature deaths annually (Cohen et al., 2017) and thus is the most important environmental risk factors for mortality by chronic diseases worldwide (Forouzanfar et al., 2016). At the EU-level, AAP is estimated to account for about 400,000 premature deaths annually and to reduce life expectancy by nearly one year (European Environmental Agency, 2017).

AAP is a complex mixture of particulate and gaseous components. These include airborne particulate matter (PM), which can be divided by size and includes soot, and gaseous pollutants such as ozone (O_3), nitrogen oxides (NO_2, NO), sulphur dioxide (SO_2), volatile organic compounds (VOCs), and carbon monoxide (CO). Pollutants can be emitted from their source as primary pollutants, such as diesel soot and NO_2 from diesel powered combustion engines (referred to as primary pollutants), or they can be formed in the atmosphere from precursor substances (referred to as secondary pollutants). Ambient particulate matter in a specific city comes from a variety of sources, and can therefore differ widely in composition and extent. A detailed characterization of AAP is therefore useful when examining the consequences and potential abatement strategies for exposure reduction. By convention, airborne particles are often classified into three major groups by their size, irrespective of their sources or chemical composition, and measured as mass concentration. Particulate matter 10 (PM_{10}) is the mass of all particles with an aerodynamic diameter of <10 µm, $PM_{2.5}$ includes particles with an aerodynamic diameter of <2.5 µm. The mass of particles between 10 and 2.5 µm ($PM_{10\text{-}2.5}$) is usually called coarse PM. Ultrafine particles[10] are defined as 100 nm or less and measured mostly as particle number concentration, since they contribute only little to particle mass (Health Effects Insitute, 2013). Specifically in atmospheric modeling, further size fractions are defined according to the mode of generation (Figure 1). These modes include the nucleation mode (smallest particles with a diameter of up to approximately 20 nm) of particles that are formed from precursor substances, the so-called Aitken mode of particles formed by condensation (size range between approximately 10 and 80 nm), the accumulation mode particles which are formed by condensation and coagulation (size range approximately 50-1,000 nm) and the coarse mode, also formed by condensation and coagulation with a size range approximately 500-10,000 nm) (Baldauf et al., 2016). The coarse mode from atmospheric modeling is therefore not the same as the so-called PM_{coarse} (PM_{10}-$PM_{2.5}$), which is derived from mass measurements and contains mostly larger particles from mechanistic processes (earth crustal material, break and tire wear).

[10] Nanoparticles which are often used in the context of ultrafine particles, represent industrially produced particles (N. Li et al., 2016)

Figure 1: Size fractions of airborne particles (Deutscher Wetterdienst, 2018)



In order to protect the human population against adverse health effects, mass concentrations of PM_{10} and $PM_{2.5}$ (particles with a diameter of less than 10 and 2.5 µm/m^3, respectively) are measured and evaluated in relation to air quality guidelines in many regions of the world. Up to now, routine measurements and regulations of the concentration of ultrafine particles (UFPs; particles with a diameter of less than 100 nm, see Figure 1) are lacking.

UFPs vary with regard to their chemical composition and physical reactivity. They are emitted directly or are formed from precursors in atmospheric processes. In urban areas, UFPs mostly originate from combustion processes through motorized vehicles, particularly alongside roads (Health Effects Insitute, 2013; Kelly et al., 2012).

Due to their small size, inhaled UFPs may enter into alveoli and are even capable to penetrate cell membranes. Consequently, UFPs may pass into the blood system, overcome the placental barrier, and finally diffuse into all organ systems including the brain and nervous system. Toxicological studies suggest that UFPs contribute to the development and progression of various diseases (Health Effects Insitute, 2013).

Epidemiological evidence for health effects of UFPs is scarce in comparison to that of larger particles. Nevertheless, an increasing number of epidemiological studies examining the exposure of the population and health effects of UFPs have been published in the last decade. Hypothesized health effects of UFP include cardiovascular and respiratory morbidity and mortality, the elicitation of local pulmonary and systemic inflammation and oxidative stress, and adverse actions on the brain and the metabolism. Two expert committees have reviewed and interpreted the epidemiological evidence base concerning UFPs (Health Effects Insitute, 2013; World Health Organization, 2013). The expert commissions of the Health Effects Institute (HEI) and of the World Health Organization (WHO) both concluded that scientific studies point towards adverse effects of UFPs on health. However, the evidence base on epidemiologic studies was not sufficient to recommend regulations on UFP concentrations.

Recently published epidemiologic studies now make it necessary to reevaluate the evidence base on the health effects of UFPs.

2 Hypotheses and Aims of the Study

The aims of this project were to systematically review the literature on the effects of UFPs on health, to evaluate the selected studies and to assess the transferability of the results to the situation in Germany. For this purpose, we focus on the following objectives:

(2) Conducting a systematic literature review
   a. Focus on health effects associated with ultrafine particles
   b. Emphasis on epidemiologic studies and quantitative effect measures (e.g., relative risks, dose-response relationships)
   c. Documentation of the literature search results and storage of all considered articles using a literature management database (EndNote).

(3) Evaluation of the identified literature
   a. Evaluation of individual study quality based on defined criteria
   b. Evaluation of the transferability of the identified findings to the present conditions in Germany

(4) Evaluation of the health relevance of ultrafine particles, specifically:
   a. Within the context of other air pollution exposures (e.g., PM_{10}, $PM_{2.5}$, ozone, nitrogen dioxide)
   b. With regard to the current German situation
   c. When considering the projected trajectory of ultrafine particle exposure in Germany

3 Methods

3.1 Selection Criteria for systematic review

We conducted a systematic literature review with a focus on epidemiologic studies that explore health effects of UFPs including quantitative effect measures *(work package 1 (a) search literature systematically in terms of health effects of UFP and (b) focusing on epidemiologic studies and quantitative effect measures (e.g., relative risk, dose-response-functions))*.

We included not only traditional/classic epidemiologic study designs such as cross-sectional studies, cohort studies and case-control studies, but also study designs often applied in environmental epidemiology such as time-series studies, panel studies, case-crossover studies, crossover studies and scripted exposure studies (novel study design in which participants are assigned to prespecified exposures, e.g. specific bike routes through a city). Furthermore, the studies had to comprise at least one of the following UFPs-measures: Particle numbers (PNC) for particles with a diameter of less than 100nm, $PM_{0.1}$, nucleation-mode, Aitken-mode particles as well as quasi-UFPs-measures: PNC for particles with a maximum diameter exceeding 100 nm, $PM_{0.25}$, surface area concentrations and accumulation mode particles. In terms of health outcomes, mortality, (international classification of diseases (ICD)-code determined) morbidity including symptoms, emergency/hospital admissions and subclinical outcomes were considered.

Toxicological studies were assessed only with regard to supporting evidence of the evaluation of UFP-related health relevance as stated in work package 3. Toxicological studies were not specifically considered in the search strategy.

Studies which investigate population related exposure to UFPs were assessed in order to evaluate the transferability of the reviewed results to the situation in Germany (work package 2b) and

to evaluate the health related relevance of UFPs with regard to the situation in Germany (work package 3b) and in consideration of the potential trends of UFP exposure in Germany (work package 3c).

Studies focusing on occupational exposures to UFPs or to industrially engineered nanoparticles were excluded.

3.2 Databases

We systematically searched MEDLINE (Medical Literature Analysis and Retrieval System Online) and LUDOK for eligible studies investigating health effects of AAP related UFPs. The period included in the search was 01.01.2011 until 11.05.2017.

MEDLINE is a comprehensive database providing international literature in the fields of medicine, psychology and the Public Health system. Currently, MEDLINE contains more than 5,600 scientific journals. The search in MEDLINE was carried out through the search engine „PubMed", published by the provider NLM (US National Library of Medicine).

In addition, we searched the LUDOK-database, which is provided by the Swiss Tropical and Public Health institute (Swiss TPH) on behalf of the Swiss Federal Office for the Environment. This database contains scientific literature on the effects of AAP on human health.

3.3 Search Strategies

3.3.1 HEI Search Strategy

In 2013, the Health-Effects Institute (HEI) published a comprehensive review on the health effects of ambient UFPs. The review was based on a literature search in the databases MEDLINE and Web of Science up to Mai 2011. In September 2011, the search was re-run and updated (Annex I, part 1).

The search of the HEI was performed by Dr. Stephanie Ebelt-Sarnat, Assistant Professor of Environmental Health at Rollins School of Public Health in Atlanta, Georgia. In the framework of our project, we replicated the search of the HEI. Divergences in the results were clarified in a phone call with Dr. Ebelt-Sarnat on 24.04.2017. The following issues were discussed:

- Applied keywords -> The HEI did not apply any truncations, only keywords documented in the search protocol (Annex I, part 1) were used.
- Eventually used field tags -> The HEI search did not include any special field-tags.
- As the HEI (1) imported the references from the Integrated Science Assessment – Particulate Matter (PM ISA) of the US EPA[11] in the Endnote database, followed by (2) references from the Web of Science and finally (3) imported the references from MEDLINE, we could not retrace in how far the references from the three different sources overlap.

The absolute number of retrieved references in the replicated search was about 2-3% higher, depending on the keyword. The search of the HEI could not be retraced thoroughly, as in the months following the HEI search, some references may not have been indexed in MEDLINE for

[11] The National Center for Environmental Assessment of the United States Environmental Protection Agency (EPA) develops Integrated Science Assessments (ISAs) that summarize the science related to the health and ecological effects caused by these pollutants

the selected search period. When we replicated the search for the same time period, these references had been included. In order to tackle this problem in the current project, we set the starting point of our search on the 01.01.2011, i.e. about half a year earlier than the end point of the search period of the HEI.

3.3.2 LUDOK Search Strategy

The LUDOK database provides epidemiological and experimental original works studying the effects of „classical"/traditional ambient air particles on humans, as well as effects of further air pollutants that have an effect on the general population (i.e. excluding agents merely relevant in occupational settings). Additionally, meta-analyses and methodological work in this context is provided.

LUDOK performs a monthly search with a constant, very broad search strategy in PubMed. LUDOK uses the following keywords and field-tags: "Air Pollutants/adverse effects" [Mesh] OR "Air Pollution/adverse effects" [Mesh[12]] OR "Air Pollutants" [Pharmacological Action] OR "Environmental Exposure/adverse effects" [Mesh] OR "air pollutants" OR "air pollution" OR "air pollutant".

Besides the regular search in PubMed, an intensive hand search was performed in over 20 relevant scientific and general journals as well as within the reference lists of publications (original works and reviews). Furthermore, LUDOK pursues notices from different sources as e.g., the Swiss TPH internal, the Bundesamt für Umwelt/ Schweiz (BAFU), the WHO and other research committees/teams.

A detailed description of the search strategy including a list of searched journals is provided in Annex 1, part 2.

3.3.3 Combined UKD Search Strategy

Our search strategy included a modified MEDLINE search of the HEI, a search in LUDOK and hand searches (Annex I, part 3).

MEDLINE Search

In the UKD search strategy, the keywords were extended in comparison to the HEI search keywords, following the very general search strategy of the LUDOK database (figure 2).

[12] Medical Subject Headings, MeSH is the National Library of Medicine's controlled vocabulary thesaurus. It consists of sets of terms naming descriptors in a hierarchical structure that permits searching at various levels of specificity.

Figure 2: Search strategies of HEI, LUDOK and UKD

	HEI	LUDOK	UKD
Database	**MEDLINE, Web of Science**	**Pubmed (Embase was replaced by hand searches)**	**MEDLINE**
time period	until 09.05.2011	since 1929	01.01.2011 - 11.05.2017
Language	English	English, German, French, Italian	English, German
#1	"air pollution"	"air pollution" "air pollutant" "air pollutants" "environmental Exposure/adverse effects" [Mesh] "air Pollutants/adverse effects" [Mesh] "air Pollutants" [Pharmacological Action] [Mesh] "air Pollution/adverse effects" [Mesh]	"air pollution" "air pollutant" "air Pollutants" "environmental exposure" "particulate matter" "environmental Exposure/adverse effects" [Mesh] "air Pollutants/adverse effects" [Mesh] "air Pollution/adverse effects" [Mesh]
#2	"surface area" "number count" "number concentration" "particle count" "ultrafine"		"surface area" PNC "particle number" "ultrafine particle" "ultrafine particles" ultrafine "nano particle" "nano particles" nanoparticle nanoparticles PM0.1 PM0.25 "accumulation mode" "Aitken mode" submicron*
#3	epidemiology health		epidemiology epidemiological epidemiologic health

The keyword „air pollution", applied by the HEI, was expanded by the keywords „air pollutant", „air pollutants", „environmental exposure", "particulate matter", "air pollutants/adverse effects [Mesh], "environmental Exposure/adverse effects [Mesh], "air pollution/adverse effects" [Mesh] were complemented in our search.

In addition to the keywords for ultrafine particles used by the HEI search, the following keywords were added: "PNC", "particle number", „ultrafine particle", nano particle", „nanoparticle", „PM0.1", „PM0.25", „accumulation mode", "nucleation mode", „Aitken mode", submicron*. The keywords „surface area" and "ultrafine", which were applied by the HEI, were retained. The HEI keywords "particle count", "number count" und "number concentration" were already represented in our search by the keywords „PNC" und „particle number". A test search including the above named HEI-keywords did not result in further additional references.

The keywords related to health outcomes „health" und „epidemiology" applied by the HEI were extended by the adjectives „epidemiological" and „epidemiologic". In order to increase the specificity of our search (i.e. to reduce false-positive retrievals), we did not use the truncated keyword epidemiolog*). A truncation would have resulted in keywords that are not relevant or sensible for our project. Field-tags as [tw] were not used, as their usage did not influence the number of matches remarkably

Alternative search strategy in MEDLINE including health-specific outcomes

Based on the above described search strategy, an alternative search strategy including specific health outcomes has been applied. Instead of using the general keywords "health" and "epidemiology/ic/ical", specific disease related keywords were used. A list with keywords provided by

the UBA was extended (by keywords as allergi*, depression, dementia, vascular, asthma, COPD, inflammation, metabolic etc.) and reduced by keywords which did not yield further matches.

Further hand searches in reviews

Reviews of the last six years presented further sources of studies. The following reviews, which were known to the investigators, were searched a priori (Baldauf et al., 2016; Cassee et al., 2013; Health Effects Insitute, 2013; Henschel et al., 2013; Rückerl et al., 2011; Stone et al., 2016; World Health Organization, 2013):

- Rückerl R, Schneider A, Breitner S, Cyrys J, Peters A. 2011. Health effects of particulate air pollution: A review of epidemiological evidence. Inhal. Toxicol. 23:555–592; doi:10.3109/08958378.2011.593587.
- WHO, Regional Office for Europe. 2013. Review of evidence on health aspects of air pollution – REVIHAAP Project, Technical Report.
- Henschel S & Chen G, WHO, Regional Office for Europe. 2013. Health risks of air pollution in Europe – HRAPIE project. New emerging risks to health from air pollution – results from the survey of experts. World Health Organ. 65.
- HEI Review Panel. 2013. Understanding the health effects of ambient ultrafine particles. Heal. Eff. Inst. 122.
- Cassee FR, Héroux M-E, Gerlofs-Nijland ME, Kelly FJ. 2013. Particulate matter beyond mass: recent health evidence on the role of fractions, chemical constituents and sources of emission. Inhal. Toxicol. 25:802–812; doi:10.3109/08958378.2013.850127.
- Stone V, Miller MR, Clift MJD, Elder A, Mills NL, Møller P, et al. 2016. Nanomaterials vs Ambient Ultrafine Particles: an opportunity to exchange toxicology knowledge. Environ. Health Perspect.; doi:10.1289/EHP424.
- Li N, Georas S, Alexis N, Fritz P, Xia T, Williams MA, et al. 2016. A work group report on ultrafine particles (American Academy of Allergy, Asthma & Immunology): Why ambient ultrafine and engineered nanoparticles should receive special attention for possible adverse health outcomes in human subjects. J. Allergy Clin. Immunol. 138:386–396; doi:10.1016/j.jaci.2016.02.023.
- Baldauf RW, Devlin RB, Gehr P, Giannelli R, Hassett-Sipple B, Jung H, et al. 2016. Ultrafine particle metrics and research considerations: Review of the 2015 UFP workshop. Int. J. Environ. Res. Public Health 13:1–21; doi:10.3390/ijerph13111054.

Besides these reviews, further reviews identified in our MEDLINE-search were screened for references.

Published abstracts from conference proceedings

As being a young area of research, published abstract bands from the following relevant conferences and symposia were searched.

- UFP-Symposium 2016 of the TU Berlin und Umweltbundesamt on 22. and 23. September 2016 in Berlin. URL: http://www.tu-berlin.de/?167019
- ETH-conference „Combustion-generated nano-particles“, 1997-2017. URL: http://www.nanoparticles.ch/conference_bibliography.html
- 6th International Symposium on Ultrafine Particles Air Quality and Climate Brussels, Belgium May 10 and 11, 2017. URL: http://ufp.efca.net/

- 20th Meeting of the Task Force on the Health Effects of Long-range Transboundary Air Pollution on 16–17 May in Bonn 2017. URL: http://www.euro.who.int/en/health-topics/environment-and-health/air-quality/news/news/2017/05/historic-20th-meeting-of-the-joint-task-force-on-the-health-aspects-of-air-pollution
- 3rd NANOAPP (Nanomaterials&Applications) 2017 (http://nanoapp.ios.si/) is a scientific meeting of acknowledged and renowned researchers, scientists and experts in the field of synthesis of various nanomaterials and their applications in Energy, Environment, Human Health, Sensors, Textiles, Medicine.
- 21st ETH-Conference on Combustion Generated Nanoparticles, June 19th to 22nd 2017, Zürich, Switzerland. URL: http://www.nanoparticles.ch/

3.4 Study selection by Inclusion and Exclusion Criteria

Ron Kappeler (RK) and Simone Ohlwein (SO) screened title, abstracts and – if needed – full texts of the studies with regard to the inclusion and exclusion criteria (see below). 10 % of the studies were screened by both reviewers. In case of uncertainties concerning the selection of a study the case was discussed by the whole team. If necessary, inclusion and exclusion criteria were clarified and extended. The process of the study selection is illustrated in a Flowchart (Annex I, part 4) and documented in a chart adapted to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) method (Figure 4).

Inclusion criteria

- Epidemiologic studies with an adequate study design, i.e.: cohort, case-control, cross-sectional, case-crossover, panel-studies, scripted exposures, time-series studies.
- Quantifiable measures of association containing at least one UFP measure/metric: Number (PNC) or size-fractioned PNC for particles < 100 nm, $PM_{0.1}$, nucleation-mode particles (NucMP) and Aitken-mode particles (AitMP) or containing at least one quasi-UFP effect measures: PNC < 3000, $PM_{0.25}$, $PM_{0.1}$, surface-area concentration or accumulation mode particles (AccMP).
- Quantifiable measures of association including at least one measure: Odds ratio, relative risk, hazard ratio, β-estimates of percent change or exposure-response functions.
- Health outcomes including mortality or ICD-coded diseases, symptoms, emergency/hospital admissions/visits, preclinical outcomes.
- Languages: English, German.
- Year: Studies published from 2011 onward until 11.5.2017 which were not included in the HEI review; studies published after the deadline are listed in the appendix (annex I, part 5)

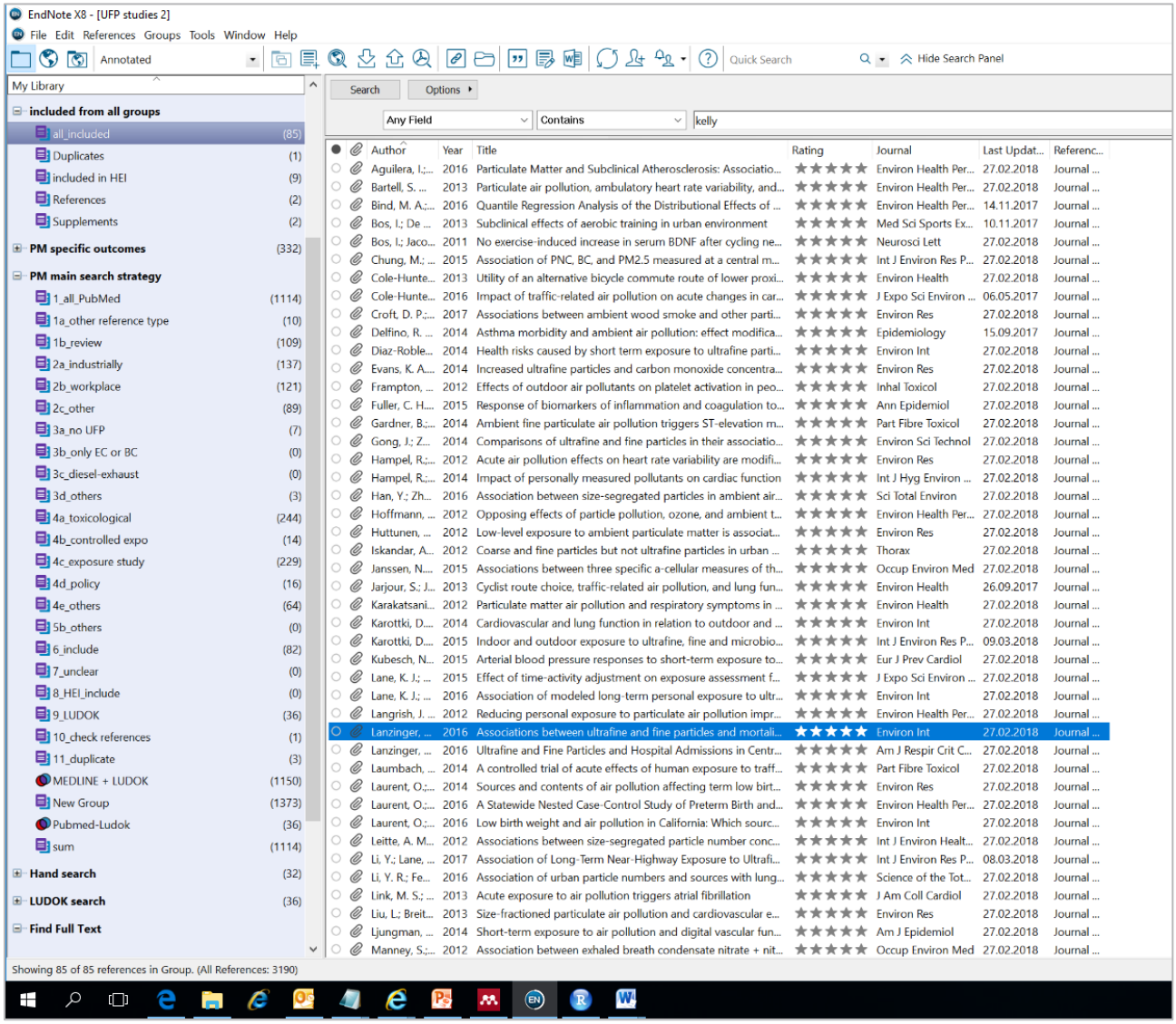
Exclusion criteria

- Toxicological studies, controlled exposure studies, animal experiments, in-vitro studies,
- Exposure to industrially engineered nanoparticles,
- Exposure to nanoparticles/ UFPs in occupational settings,
- Exposure to source-related indoor nanoparticles/ UFPs,
- Exposure to diesel particles, BC or EC only,
- Distance measures in substitution of exposure measurements
- Health outcomes of unclear health relevance, e.g. epigenetics, metabolomics, methylation.

3.4.1 Organization of the References

All references were organized within a library of the reference management program “Endnote”. Access was provided for all project team members. Group sets have been created for the four different sources “MEDLINE main search strategy”, “MEDLINE outcome specific search strategy”, “Hand Search” (including search by author) and “LUDOK”. Within each group set, separate groups have been created to document the assignment of the references analogous to the exclusion criteria (Figure 3). Furthermore, duplicates were documented within separate groups. To prevent mismatches, duplicates were not discarded automatically

Figure 3: Example: Organization of the References in Endnote in separate groups



3.5 Data extraction to evaluate the studies quantitatively and qualitatively (WP II)

The identified articles were evaluated concerning their quality of report, significance and contents as well as their transferability to the German context. The established quality criteria (Annex I, part 6) are adapted from the Quality Assessment Tools of the National Heart, Lung and

Blood Institute of the National Institute of Health (2014). When developing the different criteria, special attention was paid to exposure assessment. The available standardized instruments do not include this crucial element of studies in environmental epidemiology to the extent necessary and in the required depth. Therefore, new criteria had to be developed.

In particular, criteria to evaluate the applied measurement devices, the representativeness of the measurement sites for the exposure of the target population, the validity of used exposure models and for the assessment/modeling of several air pollutants.

Assignment of study designs

We assigned repeated measure analyses embedded within a cohort study as short-term cohort study. Scripted exposure studies are a relatively novel study design in which participants are assigned to prespecified exposures, e.g. specific bike routes through a city. Scripted exposure studies also contain so-called crossover studies, in which participants are exposed to different prespecified exposure scenarios.

There were single studies that measured outcomes in a weekly timeframe (Bind et al., 2016; Bos et al., 2013). By practical reasons, we decided to assign medium-term studies either to short- or long-term studies, depending whether their exposure assessment primarily relied on temporal variability (short-term studies) or whether it was based on spatial variability (long-term studies). This was done because these two design aspects determine the choice of the model in a major way.

4 Results

4.1 Literature search

Literature Research

The application of the main search strategy in MEDLINE yielded 1,114 references, the application of the alternative outcome-specific MEDLINE search strategy yielded 992 references, of which 332 were not included in the main search strategy (Figure 4). Together, the MEDLINE search yielded 1,446 references. The yield of the two MEDLINE searches was different regarding the focus of the resulting references: While the main search yielded many exposure studies, the alternative search strategy with specific health outcomes yielded many toxicological and animal experimental studies.

The search in the LUDOK database yielded 106 references, of which 30 were additional to the MEDLINE search. Another eight additional references were identified through hand search in other sources such as the 142 identified reviews, original article reference lists, conference proceedings, and author search, yielding an overall total of 1,484 unique references that were examined for in- and exclusion criteria.

Exclusion of articles

Out of the 1,484 identified unique references, 1,399 were excluded by title and abstract and - if necessary by full-text - according to our predefined exclusion criteria, leaving an overall number of 85 orginal references for our further evaluation (Figure 4).

Exclusions were mainly due to type of study (review, toxicological, exposure, policy, other publication type). In detail, 744 studies were excluded due to study type. Of those, 142 references were excluded because they were reviews. 14 studies did not include appropriate UFP measures and 11 studies were already included in the HEI review. Further 475 studies were excluded due to source of investigated particles (industrial, occupational, indoor).

Yield of articles by search strategy

The final number of 85 original references included in this systematic review was achieved from the following sources: Of the 1,114 unique references identified by the main MEDLINE search strategy, 70 references were included in the analysis. Of the 332 unique references identified by the alternative outcome-specific MEDLINE search strategy, 3 additional references were identified for the review.

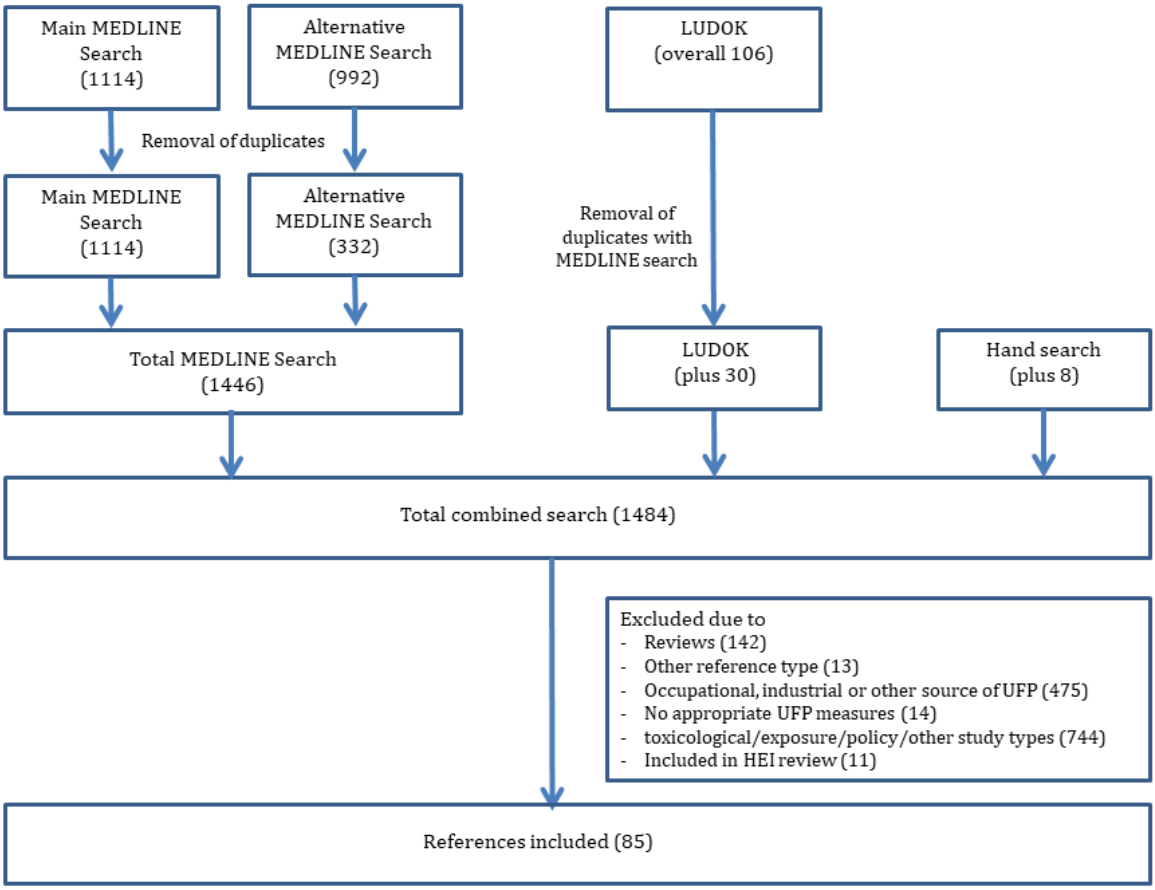
Of the 106 LUDOK references, 76 references were duplicates of already identified references through the combined MEDLINE search. 30 unique references that had not been identified through the MEDLINE search were further investigated: Of those, 14 references were assigned to the group of reviews and eight references were excluded due to the other predefined exclusion criteria. Finally, altogether eight relevant studies were identified additionally by the LUDOK database.

Of the eight studies identified through hand search, four studies met our inclusion criteria and were added to the final analysis database.

Repeated search

In a repeated search on 23.02.2018, limited to articles published or accepted after the closing date of the full search, we identified another 13 articles, which are listed in the appendix (Annex I, part 5). These articles are not included in the detailed analysis of this report, but are added for the benefit of future evaluations.

Figure 4: Study selection process adapted to the PRISMA method



Evidence base from previous reviews

Our literature research and knowledge draws upon some relevant reviews published recently. The HEI provides the most thorough and complete information on the relationships between UFPs and various health effects. This report reviewed 79 primary research articles that examined the effects of UFPs and quasi-UFPs (>1,000nm) on health published after the U.S. EPA's 2009 Integrated Science Assessment for Particulate Matter until 2011. HEI found a growing number of studies assessing the health effects of UFPs as their main focus or as one of several pollutants of interest. They found 25 epidemiological studies assessing the short-term health effects of ambient UFPs characterized by particle number, which were not included in the 2009 PM ISA. However, HEI stated that "the evidence to date continues to lack consistency and coherence (…) whether ambient UFP's affect human health differently or independently from the effects of other particle or gaseous co-pollutants" (Health Effects Institute, 2013, p.63). The body of research, which has provided a suggestive but not definitive answer on the adverse health effects of UFPs on respiratory and cardiovascular outcomes was facing three issues according to HEI:

1) Inconsistency of Outcomes

Some studies on respiratory and cardiovascular outcomes report associations with UFP exposure (e.g. (Song et al., 2011) while others do not (e.g. De Hartog et al., 2010). Various factors such

as different study designs, populations examined and UFP metric utilized or differences in pollutant composition might contribute to those inconsistencies.

2) Exposure assessment

There is a lack of larger epidemiologic studies of air pollution health effects because UFP monitoring data are scarce and when they are routinely assessed they are measured in different ways. Studies depending on only one monitor might miss the high spatial variability of UFP concentrations (Fanning et al., 2009; Terzano et al., 2010; U.S. Environmental Protection Agency, 2009). Exposure misclassification may be at least partly responsible for null findings for health effects of UFPs.

3) Independence of UFP Effects

If UFP data is available, the high covariation with other combustion-related pollutants makes it difficult to disentangle the independent effects of UFPs from other pollutants. Therefore, studies adjusting their models for expected co-pollutant effects are conducted rarely. Even in studies where other metrics or gases were measured, co-pollutant exposures were not addressed in the analysis.

On top of those issues, HEI couldn't find any studies on long-term exposure effects of UFPs. Therefore the evidence base in 2013 on epidemiologic studies was not sufficient to recommend regulations on UFP exposure concentrations.

In February 2015 the United States Environmental Protection Agency invited experts from around the world to discuss and present evidence of health effects associated with UFP exposure, which has been summarized in 2016 (Baldauf et al., 2016). According to that workshop, short-term epidemiological studies provided evidence that exposure to traffic pollution (rich in UFPs) was associated with adverse cardiovascular outcomes, however, the effects still couldn't be reliably disentangled from other PM fractions or other gaseous pollutants. The scarce UFP monitoring networks still had not allowed for a comprehensive examination of long-term UFP exposures and adverse health outcomes in more locations. Similar to HEI's conclusion, epidemiological studies did not provide enough evidence that UFPs are more potent than other PM size fractions. Nevertheless, toxicological concerns about health effects of UFPs suggested that particle size may need to be considered in assessing potential adverse effects of exposures to PM (Baldauf et al., 2016).

Chen et al. (2016) thoroughly reviewed articles on composition of UFPs, their sources, typical characters, oxidative effects and potential exposure routes with a main focus on toxicology. Furthermore they also considered evidences emerging from nanotoxicology, as this research field contributes to the understanding of toxicity mechanisms of airborne UFPs in air pollution. They concluded that UFPs play a major role in adverse impacts on human health, but further investigations are required and efforts have to be made to raise awareness of the critical hazardous potential of UFPs among the public and authorities.

An American working group (Li et al., 2016) reevaluated the conclusions made by the HEI report by assessing experimental, epidemiological and clinical trial studies published in 2014 and 2015. The authors mentioned a critical knowledge gap in clearly identifying the impact of exposure to the nano-scale pollutants on human health. However, due to new evidence, especially from experimental and toxicological studies, they questioned the validity of HEI's conclusion that there is no evidence that the adverse health effects of UFP were dramatically different from those of $PM_{2.5}$. E.g., toxicological studies suggest that UFPs promote allergic lung inflammation and are capable of inhibiting the immune response to infectious pathways. Nevertheless, the authors

concluded that the issues of epidemiological studies assessing health effects of UFPs reported by the HEI Panel still remained.

Heinzerling et al. (2016), examining respiratory health effects of UFPs in children, identified 12 relevant articles from which four were not included in the HEI-report. In single pollutant models, exposure to UFPs were associated with incident wheezing, current asthma, lung function and emergency department visits due to exacerbation of asthma. Despite the recommendations from the HEI report, there were no long-term studies conducted since the publication of the report and only one study that reported a significant association between asthma emergency department visits and UFPs, also adjusted for co-pollutants (Halonen et al., 2008). In this study, the association was no longer significant after adjusting for NO_2 exposure. Even though the evidence between UFPs and children's respiratory health is accumulating, the authors concluded for the same reasons stated by the HEI Panel that the evidence remains inconclusive.

In addition, Clark et al., published in 2016 a study focusing on biological mechanisms of cardiovascular effects beyond the alveolar barrier within the body or in vitro tissues exposed to UFPs and quasi-UFPs of up to 500 nm size. They concluded that there is some (e.g. altered autonomic modulation with increases of heart rate in animal models) up to strong evidence (e.g. vasoconstriction induced by endothelium-dependent and independent pathways mediated through UFPs) for various cardiovascular outcomes (heart rate, vasoactivity, atherosclerotic advancement, oxidative stress, coagulability, inflammatory changes). The authors state that oxidative stress is important in mediating downstream cardiovascular outcomes such as vasoactivity, heart rate etc., and therefore this might be a good target to mitigate outcomes associated with UFP exposure.

Table 1: Previously conducted reviews including search period (ordered chronologically)

Reference	Title	Comments
HEI Perspectives 3, (2013)	Understanding the Health Effects of Ambient Ultrafine Particles	79 primary research articles, published after PM ISA until December 2011
Baldauf, R. et al. (2016)	Ultrafine Particle Metrics and Research Considerations: Review of the 2015 UFP Workshop	Summary of a workshop from February 2015
Chen, R. et al. (2016)	Beyond $PM_{2.5}$: The Role of Ultrafine Particles on Adverse Health Effects of Air Pollution	Thorough review of toxicity mechanisms of airborne UFPs
Li, N. et al. (2016)	A Work Group Report on Ultrafine Particles: Why Ambient Ultrafine and Engineered Nanoparticles Should Receive Special Attention for Possible Adverse Health Outcomes in Human Subjects	34 (mostly toxicological) studies that are not included in HEI
Heinzerling. et al. (2016)	Respiratory Health Effects of Ultrafine Particles in Children: A Literature Review	4 out of 12 epidemiological studies are published after HEI (until February 2015)
Clark et al. (2016)	The Biological Effects upon the Cardiovascular System consequent to Exposure to particulates	Focusing on biological mechanisms of cardiovascular effects beyond

	of less than 500 nm in size	the alveolar barrier (studies until January 2013)

4.2 Study characteristics

Location

Overall, 85 studies published between 29.06.2011 and 26.04.2017 were identified. Most of these studies were conducted in North America (n=37) or Western Europe (n=27) (see Tables 2 and 3). Further 12 studies were performed in the Western-Pacific region. Only very few studies were conducted in Middle/ South America (n=1), Eastern Europe (n=2) and South-East-Asia (n=1). Three out of five multi-center studies included studies conducted in several Western Europe countries (Karakatsani et al., 2012; Manney et al., 2012; Samoli, Andersen, et al., 2016), two multi-center studies included study sites located both in Western and Eastern Europe countries (Lanzinger et al., 2016a, 2016b).

Table 2: World regions of studies

World region	Number of studies	%
Africa	0	0.0%
North America	37	43.5%
Middle/ South America	1	1.2%
Western Europe	27	31.8%
Eastern Europe	2	2.4%
South-East-Asia	1	1.2%
Western-Pacific	12	14.1%
Multiple study regions	5	5.9%
Total	**85**	**100.0%**

Table 3: World regions of studies, with multi-center studies assigned to multiple study locations

World region	Number of studies	%
Africa	0	0.0%
North America	37	36.6%
Middle/ South America	1	1.0%
Western Europe	44	43.6%
Eastern Europe	6	5.9%
South-East-Asia	1	1.0%
Western-Pacific	12	11.9%
Total	**101**	**100.0%**

Time frame and study design

The majority of the studies were related to the investigation of short-term effects (n=75) measuring outcomes during hours to weeks after exposure. Ten studies investigated long-term associations using exposure estimates averaged over a period of months to years. Among the included long-term studies, most studies used exposure time windows of one year. The study with the largest exposure window covered seven years (Ostro et al., 2015). Short-term studies are dominated by panel studies - 31 as repeated measures and one in a cross-sectional design, scripted exposure studies (n=16), and time-series studies (n=11). Further studies investigating short-term associations were case-crossover (n=8), cohort (n=4) and cross-sectional studies (n=4). The studies with a long-term study design consisted of cohort studies (n=4), cross-sectional studies (n=4), one case-cohort and case-control study, respectively (Table 4).

Table 4: Study design by long-term/ short-term studies

Design	Number of studies	%
Long-term	all=10	
Case-cohort study	1	1.2%
Case-control study	1	1.2%
Cohort study	4	4.7%
Cross-sectional study	4	4.7%
Short-Term	all=75	
Cohort study	4	4.7%
Cross-Sectional study	4	4.7%
Panel (cross-sectional)	1	1.2%
Panel (repeated measure)	31	36.5%
Case-crossover	8	9.4%
Scripted exposure	16	18.8%
Time-series	11	12.9%
Total	**85**	**100.0%**

% numbers are related to the sum of the long-/medium and short-term studies, respectively

Exposure assessment

Overall, most studies used measurement-based exposure assessments (87.1%) (Table 5). Model-based exposures were used in 10.6% of the studies. In long-term studies, mostly model-based exposure were used (9 out of 10), whereas the majority of short-term studies used measurement-based exposures (71 out of 75). This pattern is attributable to the fact, that model-based exposures are necessary to capture the spatial variation in exposure, which is the required exposure contrast for the assessment of long-term effects in cohorts studies (Aguilera et al., 2016; Ostro et al., 2015; Viehmann et al., 2015), cross-sectional studies and case-control/ case-cohort studies and less used in typically short-term study designs as time-series or scripted exposure studies. Among the identified studies, only two short-term studies applied model-based expo-

sures, one cross-sectional study (Fuller et al., 2015) and one time-series study (Delfino et al., 2014).

Table 5: Exposure assessment technique of medium-/long-term and short-term-studies

Exposure	Number of studies	%
Long-term	**10**	
Model based	9	10.6%
Measurements only	1	1.2%
Short-Term	**75**	
Model-based	2	2.4%
Measurements only	73	85.9%
Total	**85**	**100.0%**

The majority of the studies applied central-site measurements (n=45), followed by mobile measurement techniques (n=17) and combination of different modeling/ measurements (n=10), e.g central-site measurements in combination with spatio-temporal land-use regression models, residential measurements or microscale personal exposure models (Table 6).

Table 6: Type of expose models/ measurements used in the studies

Exposure model/measurement	Number of studies	%
Chemical-transport model	3	3.5%
Land-use regression model	1	1.2%
Dispersion model	1	1.2%
Measurement: Central site	45	52.,9%
Measurement: Residential	2	2.4%
Measurement: Mobile	17	20.0%
Microscale personal exposure model	2	2.4%
Other	4	4.7%
Combination of different types	10	11.8%
Total	**85**	**100.0%**

In most studies, UFPs were assessed as particle number concentrations (PNCs) per volume (Table 7). In about one third of the studies, PNCs sized up to 100 nm were used (29 out of 92[13]). In 63 studies, quasi-UFPs sized PNC up to 3,000 nm were used. In relation to different size modes, only one study used nucleation mode particles (n=1), representing particles with a diameter of less than 10 nm and Aitken-mode particles (n=1), representing particles with a diameter of 10-100 nm. In 14 studies, accumulation mode particles (AccMPs) were used, representing particles

[13] As many studies used various size-fractioned PNCs, the number of analyses using PNCs with a size up to 100 nm (n=29) and/or up to 3,000 nm (n=66) exceed the number of 75 included studies that assessed PNCs.

with a diameter of 100-1,000 nm[14] (see figure 1, p.13). Particles measured as mass per m^3 were used in 13 studies: In 6 studies, submicron $PM_{0.1}$ particles were assessed, in 7 studies, quasi-UFP $PM_{0.25}$ or $PM_{0.1}$ particles were assessed. Lung-deposited surface area (LDSA) was only used in two studies, of which one was long-term and one short-term.

Table 7: Particle metrics used in the studies

	Number of studies	%
PNC < 100 nm	29	23.6%
PNC < 3,000 nm	63	51.2%
NucMP	1	0.8%
AitMP	1	0.8%
AccMP	14	11.4%
$PM_{0.1}$	6	4.9%
$PM_{0.25}$	3	2.4%
PM_1	4	3.3%
LDSA	2	1.6%
Total	**123**[15]	**100.0%**

Type of outcome

We analyzed the number of studies according to the main type of outcome (mortality, hospital admissions/emergency, subclinical outcome measure, table 8) and by dividing outcomes according to major organ systems (Table 9). Eight studies assessing mortality (7 short-term, 1 long-term) analyzed the effects of UFPs on total, cardiovascular or respiratory mortality. Eleven (7 short-term, 4 long-term) studies analyzed the effects on cardiovascular, respiratory, or other morbidity outcomes. Eleven studies (all short-term) investigated UFP effects on cardiovascular or respiratory disease-related emergency calls/ hospital admissions. The vast majority of studies (n=60, 55 short-term, 5 long-term) used various subclinical measures as health outcomes, e.g. systemic inflammation. Three studies investigated several different types of main outcome types.

Table 8: Health outcome types of long-term and short-term-studies

	Number of studies	%
Long-term	All=10	
Mortality	1	1.1%
Morbidity	4	4.4%
Emergency/hospital call/admission	0	0.0%
Subclinical	5	5.6%

[14] In literature, different cutpoints are used to divide particles in the different modi.

[15] As many studies used various size-fractioned PNCs, the number of analyses using PNCs with a size up to 100 nm (n=29) and/or up to 3,000 nm (n=66) exceed the number of 75 included studies that assessed PNCs.

Short-term	All=80	
Mortality	7	7.8%
Morbidity	7	7.8
Emergency/hospital call/admission	11	12.2%
Subclinical	55	61.1%
Total	90	100.0%

* Three studies analyze several types of outcome measures and were assigned to two different outcome types. Thus, the studies do not sum up to 85 studies.

Most studies measured cardiovascular organ system-related outcomes (4 long-term, 47 short-term), followed by inflammatory (3 long-term, 26 short-term) and respiratory/atopy (1 long-term, 24 short-term) health outcomes (Table 9). Few studies investigated total mortality (1 long-term, 4 short-term), oxidative stress (0 long-term, 4 short-term) and other outcomes (e.g., pre-term birth, term low birthweight, perceived stress, 3 long-term, 2 short-term).

Table 9: Health outcomes according to organ systems of long-term and short-term-studies

	Number of studies	%
Long-term	all=13	
Total mortality	1	0.8%
Cardiovascular*	4	3.3%
Respiratory*	1	0.8%
Inflammation	3	2.4%
Oxidative stress	0	0.0%
Neurocognitive	1	0.8%
Other	3	2.4%
Short-term	all=110	
Total mortality	4	3.3%
Cardiovascular*	47	38.2%
Respiratory + Atopy*	24	19.5%
Inflammation	26	21.1%
Oxidative stress	4	3.3%
Neurocognitive	3	2.4%
Other	2	1.6%
Total	**123**	**100.0%**

* includes mortality

Outcome type by particle metric

Tables 10 to 13 give an overview of the main exposure metrics for ultrafine and quasi-ultrafine particles used for the investigation of different outcomes. Table 10 shows the overall number of studies, while table 11 and 12 differentiate the studies by short- and long-term studies. Table 13

summarizes the studies according to the use of primarily ultrafine and / or quasi-ultrafine particle size.

Most studies (n=66) use total particle number concentrations as a surrogate for ultrafine particle exposure and therefore do not investigate ultrafine particles in the stricter sense (Table 10). A smaller group of studies (n=29) uses particle number concentration measures for "true" ultrafine particle exposure available. Among the other UFP metrics, only accumulation mode particle concentration is frequently assessed, while the other metrics are rarely used. Specifically lung deposited surface area has not been established as a routine metric in epidemiological studies yet.

Across outcome types, the majority of studies investigates short-term subclinical outcomes with total particle number measurements. Only in the relatively small group of mortality studies, about half of the studies use true ultrafine particle exposure metrics (i.e. particles smaller than 100nm). Only few long-term studies are available yet, applying particle number either for total particles or for submicron particles or the mass-based measure $PM_{0.1}$.

Table 10: Number of studies (long and short term) by outcome and exposure assessment.

All studies	PNC<100 nm	PNC<3000 nm	NucM	AitM	AccM	$PM_{0.1}$	$PM_{0.25}$	PM_1	LDSA
Mortality	4	6	0	0	0	1	0	0	0
Morbidity	3	6	0	0	1	3	0	0	0
Emergency	5	10	0	0	3	2	0	1	0
Subclinical	17	44	1	1	10	0	3	3	2
Total	29	66	1	1	14	6	3	4	2

* PNC: particle number concentrations, nm: nanometer, NucM: nucleation-mode particles, AitM: Aitken-mode particles, AccM: Accumulation-mode particles, LDSA: long-deposited surface area, UFP: particles with a diameter sized less than 100 nm.

Table 11: Number of short-term studies by outcome and exposure assessment

Short-term	PNC<100 nm	PNC<3000 nm	NucM	AitM	AccM	$PM_{0.1}$	$PM_{0.25}$	PM_1	LDSA
Mortality	4	6	0	0	0	0	0	0	0
Morbidity	1	5	0	0	1	0	0	0	0
Emergency	5	10	0	0	3	2	0	0	0
Subclinical	16	40	1	1	10	0	3	3	1
Total	26	61	1	1	4	2	3	3	0

Explanations see Table 10

Table 12: Number of long-term studies by outcome and exposure assessment

Long-term	PNC<100 nm	PNC<3000 nm	NucM	AitM	AccM	$PM_{0.1}$	$PM_{0.25}$	PM_1	LDSA
Mortality	0	0	0	0	0	1	0	0	0
Morbidity	2	1	0	0	0	3	0	0	0

Emergency	0	0	0	0	0	0	0	0	0
Subclinical	1	4	0	0	0	0	0	0	1
Total	3	5	0	0	0	4	0	0	1

Explanations see Table 10

In sum, 33 studies used UFP exposure measures and 69 studies used quasi-UFP exposure measures (Table 13). In 19 of the above named studies, both UFP and quasi-UFP measures were used. Among long-term studies, the number of studies measuring UFP and quasi-UFP was equal with five studies each, whereas short-term study authors mostly applied quasi-UFP measures.

Concerning the different outcome types, subclinical outcomes were more frequently related to quasi-UFP measures than to UFP measures, both in long-term and short-term study designs. For mortality and morbidity, the ratio of quasi-UFP versus UFP was nearly balanced, for emergency department/hospital admissions, quasi-UFPs were more frequent (10 versus 6).

Table 13: Number of studies by outcome and UFP versus quasi-UFP measurement

Studies	**short-term**			**long-term**			**total**		
Metric	**UFP**	**quasi-UFP**	**UFP + quasi-UFP**	**UFP**	**quasi-UFP**	**UFP + quasi-UFP**	**UFP**	**quasi-UFP**	**UFP + quasi-UFP**
Mortality	1	3	3	1	0	0	2	3	3
Morbidity	0	4	1	3	1	0	3	5	1
Emergency	1	5	5	0	0	0	1	5	5
Subclinical	7	36	10	1	4	0	8	40	10
Total	**9**	**45**	**19**	**5**	**5**	**0**	**14**	**50**	**19**

UFPs consist of the particle metrics PNC <100 nm, NucMP, AitMP and $PM_{0.1}$, quasi-UFP: particles with a diameter sized < than 3,000 nm, without a cutpoint at 100 nm, even though particles might me dominated by particles <100 nm. Here, quasi-UFPs consist of PNC <3,000 nm, AccMP $PM_{0.25}$ and $PM_{0.1}$.

The average particle numbers across the studies range from 1,646/ml PNC (no size range reported) in the warm season modelled by a dispersion model and assigned to study participants in California (Delfino et al., 2014) and 2,905/ml PNC (10-100 nm) assessed by a central monitor in Rochester (Croft et al., 2017) up to mean averages of 164,464/ml total PNC (10-1,000 nm) assessed by mobile measurement at a highly trafficked site in Barcelona (Kubesch et al., 2015). The highest central site-measured UFP-concentrations measured in Europe were assessed in Rome with 34,046/ml total PNC and and outside Europe in Beijing with 43,900/ml PNC (11.1-101 nm) (Song et al., 2013).

Quality indicators – general aspects and study population

Most studies clearly stated the research question (n=82). In most publications (n=82), the study authors specified the included participants clearly. In more than half of the studies (n=49), convenience samples were used. Six studies recruited participants from random samples of the population. Further seven studies used a combination of random and convenience samples. In

10 studies, other sample types were used, e.g. subsets of cohorts with specified health measurements or cohorts of specified groups.

In 10 studies, which were all time-series studies, the sample was completely representative for the general population as the data assessment of the health endpoint referred to the whole population in the respective study area. In thirteen studies, the study populations were representative samples of population groups, e.g., children or adults above a certain age. In most of the included studies (n=62, 72.9%), the study population was a selected group, not representative for the general population. A sample size justification was rarely provided (n=3). Most of the study participants (exposed and unexposed or cases and controls) were recruited from the same populations and the same time period (n=71 and n=82).

Table 14: Quality criteria of the UFP/quasi-UFP Studies concerning selection bias

Quality aspects – study population	n	%
Was the research question or objective in this paper clearly stated?		
Yes	82	96.5%
Not specified, reference given	2	2.4%
Not specified, no reference given	1	1.2%
	85	100.0%
Was the study population clearly specified and defined?		
Yes	82	96.5%
Not specified/ reference given	2	2.4%
Not specified/ no reference given	1	1.2%
	85	100.0%
Sample Type		
Random	6	7.1%
Convenience	49	57.6%
Random + Conv.	7	8.2%
Other	10	11.8%
NA	13	15.3%
	85	100.0%
Representativeness		
Completely	10	11.8%
Somewhat	13	15.3%
Selected group	62	72.9%
	85	100.0%
Was a sample size justification, power description provided?		

Yes	3	3.5%
Not reported/ reference given	1	1.2%
Not reported/ no reference given	77	90.6%
Not applicable	4	4.7%
	85	100.0%
Were all the subjects selected or recruited from the same or similar populations?		
Yes	71	83.5%
No	12	14.1%
Not reported/ reference given	0	0.0%
Not reported/ no reference given	2	2.4%
	85	100.0%
Were all the subjects selected or recruited from the same time period?		
Yes	82	96.5%
No	0	0.0%
Not reported/ reference given	3	3.5%
Not reported/ no reference given	0	0.0%
	85	100.0%

Quality indicators – Exposure assessment

Second, quality aspects concerning exposure assessment were investigated (Table 15). The majority of the studies (n=66, 77.6%) reported the size-ranges of the measured UFPs. Almost all studies (n=79, 92.9%) reported the technical device used to measure the particles. Less than half (n=34) of the studies assessing other air pollutants (n=78) adjusted for co-pollutants within multi-pollutant-models. Studies without adjustment for co-pollutant were considered as "high risk of bias". 66 studies adjusted for meteorology, from which the majority (n=64) were short-term studies.

Table 15: Quality of the UFP/quasi-UFP Studies concerning exposure assessment

Quality aspects - exposure	n	%
Reporting of UFP size ranges		
Reported	66	77.6%
NR/ reference given	6	7.1%
NR/ no reference given	13	15.3%
	85	100.0%
Reporting of technical device		
Reported	79	92.9%

NR/ reference given	4	4.7%
NR/ no reference given	2	2.4%
	85	100.0%
QA/QC for UFP measures described		
Yes	33	38.8%
Partly	1	1.2%
No	51	60.0%
Assessment of other air pollutants		
Yes	78	91.8%
No	7	8.2%
	85	100.0%
Adjustment for co-pollutants		
Yes	34	40.0%
No	48	56.5%
unclear	3	3.5%
	85	100.0%
Adjustment for meteorology		
Yes	66	77.6%
No	19	22.4%
	85	100.0%

Quality indicators – Outcome assessment

Third, quality aspects concerning the outcome assessment were explored (Table 16). In all but one study (n=84) assigned exposure values were measured or modeled for time periods prior or parallel to the assessment of the outcome or for the time period of follow-up. In five of the included long-term studies, this was achieved by the use of chemical transport modeling, which allows the estimation of daily air pollutant concentrations for specific time periods. Furthermore, all but one study (n=84) defined and described the outcome measures clearly. In 68 of the studies, a blinding of the outcome assessors could be presumed. In 15 studies, no blinding was ensured. 13 of these studies were scripted exposure studies and we assumed that no blinding was conducted in these studies, unless specifically mentioned in the reported methods.

Table 16: Quality of the UFP/quasi-UFP Studies concerning outcome assessment

Quality aspects	n	%
Exposure measured prior or parallel to outcome assessment		
Yes	84	98.8%
No	1	1.2%
	85	100.0%

Outcome measures clearly defined and implemented		
Yes	84	98.8%
No/ reference given	1	1.2%
	85	100.0%
Outcome assessors blinded to exposure status resp. case-control status of participants		
Yes	68	80.0%
Partly	2	2.4%
No	15	17.6%
	85	100.0%

4.3 Health effects

4.3.1 Short-term effects

4.3.1.1 Mortality

Seven short-term time-series-studies (Lanzinger et al., 2016a; Leitte et al., 2012; Meng et al., 2013; Samoli, Atkinson, et al., 2016; Stafoggia et al., 2017; Su et al., 2015; Wolf et al., 2015) applying central-site measurements investigated effects of PNCs (four studies measured UFPs, two quasi-UFPs) on various mortality outcomes (total, cardiovascular, respiratory) (Table A1a). Two of the seven studies were conducted in a multi-center approach covering several countries (Lanzinger et al., 2016a; Stafoggia et al., 2017). Six studies adjusted for co-pollutants or constituents (Samoli, Atkinson, et al., 2016) for at least parts of the examined particle-outcome relationships (Table A3a).

All-cause mortality

For all-cause mortality, the evidence base is currently inconsistent: **Stafoggia** et al. (2017) investigated effects of PNC size ranges on non-accidental mortality in 8 western European cities from 1999 to 2013. The pooled percent changes were strongest for an increase of PNC of 10,000 particles at 7 days before death, with effect estimates of 0.37% (confidence intervals (CI) -0.03%; 0.78%) increase in total non-accidental mortality. These effects were mostly influenced by the Rome estimate. The city-specific non-significant effect estimates varied from about -0.8% (Athens) to 1.9% (Augsburg) (I^2 index for heterogeneity was below 50%). In Shenyang in China, **Meng** et al. (2013) observed consistent and weak significant positive associations between particles of six size fractions between 250 and 500 nm and all-natural-cause mortality with a 2-day moving average interqurartile range (IQR) incremental change (e.g., percent change per 2,600 PNC250–280/ml: 2.41% (1.23%, 3.58%)).

These associations were only present in the larger size fractions up to 10 µm in the warm season and absent in the cold season. Within the framework of the UFIREG[16]-project, **Lanzinger** et al.

[16] UFIREG: Ultrafine Particles - an evidence based contribution to the development of regional and European environmental and health policy

(2016a) did not find associations between UFPs/quasi-UFPs averaged over lag 0-2 and natural mortality (percent changes of relative risks (RRs) per 2,750 PNCs: 0.1% (−2.0%; 2.4%)). Exposures averaged over lag 2-5 (delayed) and averaged over lag 0-5 (cumulated) yielded similar non-significant results. Finally, **Samoli** et al. (2016) investigated single-site measured, 1-day lagged PNC > 6nm related mortality among approximately 9 million Londoners and found estimates close to zero.

Two of the above named studies adjusted for co-pullutants (Meng et al., 2013; Stafoggia et al., 2017): When **Stafoggia** et al. (2017) adjusted their models for NO_2, $PM_{2.5}$ and $PM_{2.5-10}$ for 5-, 6- and 7-lagged exposures, estimates decreased considerably or even turned into a negative direction (e.g., for NO_2 and lag 7: -0.25% (-0.72%; 0.22%)), whereas estimates decreased to a minor extent upon adjustment of PM_{10}, CO or O_3 (e.g., for PM_{10} and lag 7: 0.28% (-0.13%; 0.68%)). The effect estimates for quasi-UFPs in the study by **Meng** et al. (2013) decreased only moderately and remained statistically significant upon adjustment for SO_2, NO_2, $PM_{2.5}$ and PM_{10}, respectively. The lowest two-pollutant estimate was 1.66% (0.14%, 3.17%) for the model including NO_2. Upon adjustment for $PM_{2.5-10}$ effect estimates became stronger.

Respiratory mortality

The up-to-date body of evidence for respiratory mortality is similarly inconsistent. Strong associations were observed only in two studies conducted in China (Leitte et al., 2012; Meng et al., 2013). **Leitte** et al. (2012) explored associations between various size-fractions of PNC and respiratory mortality in about 8,000,000 residents in Beijing, China, from 2004 to 2005. They found slightly negative (1-day lag) to non-significant positive (same day, lag of 2 days, average of 4 days and average of 5 days) changes per IQR of 13,000 particles/ml in the UFP range (PNC3-100nm). Associations with PNC sized 3-1,000 nm ranged from close to zero to significantly positive (percent increase per IQR of 14,000 particles/ml: 9.3 (1.3–17.9)) for a 2-day lag. **Meng** et al. (2013) observed associations slightly above the null effect between particles of eight size fractions between 250 and 1,000 nm and respiratory mortality. These associations increased substantially in the warm season, but did not reach statistical significance. **Lanzinger** et al. (2016a) found positive but non-significant associations of UFPs and quasi-UFPs with respiratory mortality for the 2-day lag, lag 2-5 and lag 0-5. **Stafoggia** et al. (2017) observed inconsistent estimates, ranging from significantly negative (lag 3) to positive associations (e.g., lag 6, lag 10). The study by **Samoli** et al. (Samoli, Atkinson, et al., 2016) indicated slightly inverse percent changes of 2-day lagged quasi-UFPs related respiratory mortality.

Of the above stated studies, two adjusted for co-pollutants (Lanzinger et al., 2016a; Leitte et al., 2012) and one for constituents and total PNC within the source-related estimates (Samoli, Atkinson, et al., 2016). Upon adjustment for SO_2, NO_2 or PM_{10}, the significant associations in the study by **Leitte** et al. (2012) remained positive but lost significance for PNC sized 300-1,000 nm averaged over 4 and 5 days. The decline in effect estimates was strongest after adjustment for NO_2. The association with PNC total as an average of two days didn't change upon adjustment for any of the co-pollutants and remained significantly positive. The effect estimates for 6-day averaged PNC < 100nm in **Lanzinger** et al. (2016a) decreased only slightly after adjusting for $PM_{2.5}$ and increased and became significant upon adjustment for NO_2. In the two-pollutant models by **Samoli, Atkinson** et al. (2016) co-source adjustment was conducted, i.e. estimates for single sources were calculated and these estimates were adjusted for all other sources. Furthermore, co-pollutant models were adjusted for PNC total minus the investigated source. The inverse effect estimate for respiratory mortality was generally robust to co-source adjustment. Mutual adjustment for all sources generally exerted a greater influence on the estimates compared with estimates from two-source models.

Cardiovascular mortality

The six studies investigating short-term effects of UFPs on cardiovascular (CV) mortality indicate inconsistent evidence: **Meng** et al. (2013) observed significant positive associations with a 2-day average IQR incremental change in PNCs fractions between 250 and 650 nm in Shenyang, China. These associations were stronger in the warm season and lost significance in the cold season. **Su** et al. (2015) also found significant positive associations between UFP particles of differend size fractions and cardiovascular mortality with a 1-day lag and cumulated 5-day average in Beijing. For lag 0, associations were weakly positive. A German single-centre time series study analyzed measured PNC sized between 10 and 2,000 nanometers with fatal myocardial infarction (Wolf et al., 2015). The authors found slightly positive to quasi null associations with same day exposure, previous day exposure and with mean exposures of the 4 preceding days. **Lanzinger** et al. (2016a) reported slightly inverse effects of UFPs and quasi-UFPs on cardiovascular mortality for 2-day lag, lag 2-5 and lag 0-5. Likewise, Stafoggia et al. (2017) found non-significant inverse associations of UFPs for the lags 0 to 3, 8 and 9, zero effects at lags 4 and 10 and slightly positive effects at lags 5 and 6 with the strongest effect at lag 7 for CV mortality. Similar to their results for respiratory mortality, **Stafoggia** et al. (2017) observed inconsistent estimates, ranging from negative (lags 0-3, 8,9) to positive associations (lag 7). **Samoli**, **Atkinson** et al. (2016) observed inverse effect estimates for the association between 1-day lagged quasi-UFPs and cardiovascular mortality (RR: -1.86 (-4.50, 0.86) per 5,180/ml).

Of the above stated studies, two studies applied two-pollutant models (Lanzinger et al., 2016a; Su et al., 2015). In **Lanzinger** et al. (2016a), the null association of PNC < 100 nm averaged over day 2-5 with cardiovascular mortality remained close to zero after adjusting for $PM_{2.5}$ and turned to a significant inverse association upon adjustment for NO_2. **Su** et al. (2015) observed only slightly reduced (by 1-2%) and still significant effect estimates for 5-day averaged PNC < 100 nm upon adjustment for PM_{10} and $PM_{2.5}$. However, upon adjustment for NO_2, effect estimates lost significance and decreased by about 5%.

Summary: Mortality

In comparison to the prior evidence, seven additional studies have been conducted with overall mixed results. For all-cause mortality, only two out of four studies found positive estimates in analyses not adjusted for co-pollutants. Of these, only one study showed positive associations for quasi-ultrafine particles after adjustment for other pollutants, while in the other study, elevated point estimates decreased towards the null upon adjustment.

The evidence of respiratory mortality is also scarce and inconsistent. Out of the 5 studies on respiratory mortality, four studies found positive, though mostly non-significant associations for UFPs or quasi-UFPs. Three studies adjusted for co-pollutants, with opposite effects after NO_2 adjustment, leading either to an enhancement or to an attenuation of effect estimates after adjustment for NO_2. The studies presented two-pollutant associations only for those models/ lags/ size fractions showing the strongest associations. Thus, the specific effect estimates are difficult to compare and consistency of the results can't be fully assessed.

Similar to the overall results for respiratory mortality, associations of UFP/quasi-UFP with CV mortality are inconsistent. The six single exposure studies observe positive (three studies) as well as inverse associations (three studies) with CV mortality. In the two multi-pollutant studies, adjustment for NO_2 led to a decrease in effect estimates, causing the loss of significance in one study and a decrease to a significantly inverse relationship in the other study. Adjustment for $PM_{2.5}$ only caused small or no changes in the UFP estimate.

Evidence from this as well as from prior reviews suggests that effects may be larger in the warm season; therefore possible effect modification by season is an important factor to consider in future short-term effect studies. Moreover, the observed effects at least partially overlap with other air pollutant effects, most clearly seen for NO_2. Due to differences in investigated size fractions, no conclusions can be made about the most important fractions.

Table 17: Summary table of conducted analyses in the seven mortality studies

Study	All-Cause	Single pollutant associations	Multi-pollutant associations	Respiratory	Single pollutant associations	Multi-pollutant associations	Cardiovascular	Single pollutant associations	Multi-pollutant associations
Lanzinger et al. 2016a	✓	0	0	✓	(+)	+	✓	(-)	-
Leitte et al. 2012				✓	UFP: (+), quasi-UFP: +	UFP: 0 quasi-UFP: (+)			
Meng et al. 2013, (only quasi-UFP)	✓	+	+	✓	(+)	nc	✓	+	nc
Samoli et al. 2016	✓	0	0	✓	-	-	✓	(-)	nc
Stafoggia et al.,2017	✓	(+)	(-)	✓	+	nc	✓	(-)/(+)*	nc
Su et al. 2015							✓	+	(+)
Wolf et al. 2015							✓	(+)	nc

0 indicates no association. (+) and (-) indicates primarily non-significant associations, + and - indicate significant associations. Nc: not conducted

4.3.1.2 Morbidity

Associations between UFPs/quasi-UFPs and acute morbidity outcomes was assessed by one panel study (Karakatsani et al., 2012), two case-crossover studies (Cole-Hunter et al., 2013; Link et al., 2013), a scripted exposure study (Langrish et al., 2012), a time-series study (Wolf et al., 2015) and two cohort studies (A. J. Mehta et al., 2015; Y. Wang et al., 2014). Exposure measurement was conducted by a central-site monitor, except for the scripted exposure and one case-crossover study (Cole-Hunter et al., 2013, Langrish et al., 2012) which determined personal pollution using monitoring equipment within a backpack. One study used a multi-center approach covering several states within the EU (Karakatsani et al., 2012) (Table A1b). None of the studies additionally adjusted for co-pollutants.

Respiratory morbidity

One panel study within the framework of the "Relationship between Ultrafine and fine Particulate matter in Indoor and Outdoor air and respiratory Health" (RUPIOH)-project explored total PNC related effects on respiratory symptoms (**Karakatsani** et al., 2012). Respiratory health of 136 participants aged 35 and older and with either chronic obstructive pulmonary disease (COPD) or asthma from Amsterdam, Athens, Birmingham and Helsinki was monitored for six months by daily symptoms diary. Daily diary records contained breathing problems after wake up, shortness of breath, wheeze, cough, phlegm as well as limitations of vigorous activities, moderate activities and walking due to breathing problems. Karakatsani et al. (2012) did not find significant associations between PNC and daily respiratory symptoms over lag 0-2 (e.g. lag1 OR

for cough per 10,000 particles/cm³: 1.009 (0.944; 1.079)) and a 6-day moving average (OR 0.894 (0.714; 1.119). They found significant inverse effect estimates for shortness of breath (lag1: OR 0.91 (0.844; 0.982) and limitations in walking (ma0-6: OR 0.804 (0.658; 0.981)).

In contrast, **Cole-Hunter** et al. (2013) found significantly increased respiratory symptoms in healthy adults commuting in an urban environment of high air pollution. 35 participants completed two return trips, one each in a highly polluted area near busy roads and one alternative route of lower proximity to motorized traffic. Participants reported significantly more nose and throat irritation between high and low pollutant trips (mean ± standard deviation: 1.82 ± 0.33 vs. 1.53 ±0.23, $p<0.01$ and 2.00 ± 0.4 vs. 1.56 ± 0.24, $p<0.01$, respectively). An evaluation of independent effects of UFP, however, is not possible.

Langrish et al. (2012) conducted a study with patients suffering from coronary heart disease measuring quasi-UFP PNC by mobile devices. In this scripted exposure study 98 patients walked on a predefined route in central Bejing, once while using a respiratory mask and once while not using the mask. At the beginning of the study day, 2 hours and 24 hours after the walk they were asked to report physical symptoms (headache, dizziness, nausea, tiredness, cough, breathlessness, irritation of the throat or nose, unpleasant smell and bad taste in the mouth). In the presence of the mask, there were significantly lower self-reported symptoms except for dizziness and breathlessness, but since the exposure with the mask was determined based on measurements of mask filter efficacy and the authors only conducted a group comparison (mask vs. no mask), they were not able to disentangle the effect of total PNC from other pollutants.

Cardiovascular morbidity

Utilizing a case-crossover design, **Link** et al. (2013) investigated the onset of atrial fibrillation in patients with dual chamber implantable cardioverter-defibrillators (ICDs) associated with total PNC 2 and 24 hours before the event. The authors observed increased odds for atrial fibrillation with 12% (-19; 56) for a moving average of 24h per 8,400 particles/cm³ and an even higher odds with 24% increase (-4; 61) for a moving average of two hours per 10,900 particles/cm³.

Within the Cooperative Health Research in the Region Augsburg (KORA)-framework **Wolf** et al. (2015) investigated in a registry-based time-series study the effect of PNC (10-2,000 nm) on nonfatal myocardial events in the general population. A total of 8,298 coronary events were recorded and thereof 3,303 were recurrent events. An non-significant increased risk of 2% (-1.5; 5.8) per IQR change of 6,800 particles/cm³ of the four preceding days in nonfatal myocardial events and a significantly increased risk of 6% (0.6; 11.7) in recurrent myocardial events were found. However, while most effect estimates for different lags and outcomes were above the null, only one estimate for recurrent events was significantly elevated.

Mental health

Between 2005 and 2008, adults 65 years of age and older without cognitive impairment were recruited for the MOBILIZE Cohort in Boston and followed until 2010. During two in-home interviews the presence of depressive symptoms using the 20-item CESD-R Scale was assessed by trained staff (**Wang** et al., 2014). There was no association between the presence of depressive symptoms and the proximity to major roadways as an indicator for long-term traffic exposure. They also found no evidence suggestive of a positive association between depressive symptoms and mean PNC exposure in the preceding two weeks. **Mehta** et al. investigated the association between AP and non-specific stress in a cohort of 987 elderly men in the Veterans Administration Normative Aging Study. Stress was quantified with a 14-item Perceived Stress Scale (PSS),

which scores stress experienced in the previous week from 0-56. PNC (7-3,000 nm) at moving averages of 1, 2 and 4 weeks were associated with increased stress. An interquartile range increase of 15,997 particles/ml in one week average PNC was significantly associated with a 3.2 point (2.1; 4.3) increase in perceived stress score.

Summary short-term studies on morbidity outcomes

Of the few studies investigating short-term effects of UFPs/quasi-UFPs on morbidity outcomes, only two studies observed significantly elevated estimates with a marker of perceived stress and with recurrent coronary events. Since none of the above mentioned studies adjusted for co-pollutants or were by design able to disentangle the independent effects of different constituents of the air pollution mixture, we cannot conclude an independent effect of UFPs on morbidity outcomes. The evidence base for CV morbidity outcomes is scarce with only two studies available on different outcomes. This evidence suggests that participants with preexisting cardiovascular disease might be more susceptible to adverse associations with elevated UFP/quasi-UFP concentrations.

However, while both studies show generally positive associations, no inference on the independence on the reported UFPs effect can be made. The evidence for associations with short-term changes in mental health symptoms is insufficient.

4.3.1.3 Emergency department/ hospital call/visit/admission

The use of emergency health care services was investigated by five case-crossover studies (Evans et al., 2014; Gardner et al., 2014; Iskandar et al., 2012; Rosenthal et al., 2013; Wichmann et al., 2013) and six time-series studies (Delfino et al., 2014; Diaz-Robles et al., 2014; Lanzinger et al., 2016b; Liu et al., 2013; Samoli, Andersen, et al., 2016; Samoli, Atkinson, et al., 2016) (Table A1c). Six studies adjusted for co-pollutants (Evans et al., 2014; Iskandar et al., 2012; Lanzinger et al., 2016b; Rosenthal et al., 2013; Samoli, Andersen, et al., 2016; Samoli, Atkinson, et al., 2016) (Table A3c).

Respiratory disease

Three studies (Delfino et al., 2014; Evans et al., 2014; Iskandar et al., 2012) investigated UFP-related effects on asthma symptoms or hospital/emergency department visits in children. **Evans** et al. (2014) found adverse associations between central-site measured PNC<100nm and the occurrence in pediatric asthma visits in 74 asthmatic children aged 3-10 years (OR up to 1.27 (0.90; 1.79), for an average of 4 days per 2,088/ml PNC in Rochester. Other lags and accumulation mode particle associations were null or inverse. A register-based study from Copenhagen (**Iskandar** et al., 2012) explored the effects of central-site measured PNC sized 10-700 nm on hospital admissions for asthma in 8,226 children aged 0-18 years. In this case-cross-over study, the authors observed non-significant associations for an average of 5 days, being strongest for 0-1-year-old infants (OR: 1.08 (0.97; 1.22) per 3,812.86/ml PNC). In a case-crossover analysis of hospital admissions in children with asthma, **Delfino** et al. (2014) investigated, whether high exposure to traffic-related air pollution, measured among others as UFP, modified the air pollution-outcome association. They found that generally, associations were stronger in high traffic exposure situations, especially in the cold season. Direct associations with UFPs/quasi-UFPs were not reported.

One of the four respiratory illness related time-series studies investigated associations between $PM_{0.1}$ and outpatient visits for respiratory illness in the general population of Temuco, Chile (Diaz-Robles et al., 2014). **Diaz-Robles** et al. (2014) reported a significant associations primarily for longer lags of 3,4 and 5 days (e.g., RR for lag 4: 1.07 (1.04; 1.10) per 4.73 µg/m^3 $PM_{0.1}$). The study by **Samoli,** Atkinson et al. (2016) (short description provided in 4.3.1.1) indicated positive associations in 0-14-year old children (percent increase per 5,180/ml 1.86 (-0.28; 4.05) for a two-day lag, but not in the 15+ age groups (15-64-year olds: -1.14 (-2.66; 0.41) in a study in the UK. Moreover, two multi-center time-series studies investigated this issue. A large study by **Samoli**, Andersen et al. (2016) analyzed central site-measured PNCs of various size fractions in relation to hospital admissions in approximately 9 million persons from five Western European cities during 10 years. Inconsistent and non-significant pooled effect estimates were observed across different lags ranging from a percentage change of -0.44 (-1.73; 0.87) per 10,000 particles/ml for lag 7 up to 0.43 (-0.58; 1.45) for lag 5. One aspect of this study was the lack of a harmonized approach to UFP measurements, relying on site-specific measurements that had already been in place, at least partly explaining the heterogeneity of the results. **Lanzinger** et al. (2016b) investigated respiratory UFP effects in 2,582,000 habitants of five Western and Eastern European cities in the UFIREG study. The authors found consistent pooled non-significant increased relative risks up to 3.4% (-3.2; 7.3) in the 6-day average submicron PNC per increment of 2,750 particles/ml. One strength of this multi-center study was the attention given to a harmonized exposure assessment in all study centers.

Four studies investigating UFP-related emergency department visits/hospital admissions for respiratory disease adjusted UFP-associations for co-pollutants and one study adjusted for co-sources (Samoli, Atkinson, et al., 2016) (Table A3c). Two studies found no major effect of adjustment for co-pollutants on the original estimates (Samoli, Andersen et al., 2016; Evans et al. 2014): While single day lags changed slightly in most of the two-pollutant models, the general direction of the inconsistent associations remained inverse in **Samoli, Andersen** et al. (2016). When adjusting for NO_2, the mostly inverse effect estimates remained stable or became closer to the null for the 0 to 2-day lags. However, for 3 to 7-day lags, most effect estimates decreased and turned into a negative direction, specifically when adjusting for NO_2. For the 5-day lagged exposure, the previously positive point estimate turned into a significantly inverse direction upon adjustment for NO_2. Adjustment for $PM_{2.5}$ and PM_{10} led mostly to decreased effect estimates, but most point estimates remained positive. **Evans** et al. (2014) conducted two-pollutant models using the pollutants shown to be associated with asthma exacerbation. In the two-pollutant models including carbon monoxide and O_3, the authors state that the effect estimates in these models did not differ substantially from those in the single-pollutant models without presenting any data. However, no NO_2 or $PM_{2.5}$ adjustments were performed, and the original, un-adjusted estimates were already inconsistent and not significantly different from the null.

The other two studies found decreases to the null after adjustment for co-pollutants: In the study by **Lanzinger** et al. (2016b), adjusting UFPs for $PM_{2.5}$ and NO_2 in the models averaged for 5 days, led to weakened effect estimates which turned negative in Ljubljana and Prague, while they stayed slightly above the null in Augsburg and Dresden. The pooled estimate, however, also reversed into an inverse relationship. In general, the two-pollutant models with NO_2 showed stronger decreases in effect estimates than with $PM_{2.5}$. **Iskandar** et al. (2012) conducted two-pollutant models for 4-day averaged UFP and asthma-related hospital admissions. After adjustment for PM_{10}, $PM_{2.5}$, NO_2 and NO_x, the positive associations disappeared.

Cardiovascular disease

Three case-crossover studies (Gardner et al., 2014, Rosenthal et al., 2013, Wichmann et al., 2013) and three time series studies (Lanzinger et al., 2016b; Liu et al., 2013; Samoli, Atkinson, et al., 2016) explored the effects of central-site measured PNCs of various size fractions on the use of health services due to acute cardiovascular conditions. None of the case-crossover studies adjusted for co-pollutants, while all of the time-series studies did.

In their case-crossover study, **Gardner** et al. (2014) used medical records on cardiac catheterizations from a hospital in New York and analyzed UFP-related effects on 677 myocardial infarctions, classified in ST-elevation[17] myocardial infarction (STEMI) and non-ST-elevation myocardial infarction (NSTEMI). The authors observed non-significant positive effect estimates for STEMI (OR for an average of 24h per 3,284 PNC/ml: 1.06 (0.89; 1.26). Accumulation mode particles were associated with slightly higher point estimates (OR for an average of 24h per 755 PNC/ml: 1.12 (0.92; 1.38). The study did not indicate any association for NSTEMI with PNCs or accumulation mode particles. Two studies investigated effects of PNCs and accumulation mode particles on out-of-hospital cardiac arrests. **Rosenthal** et al. (2013) studied effects of submicron PNCs and accumulation mode particles on 2,134 cases of out-of-hospital cardiac arrests due to all cardiac causes, MI and other cardiac causes in Helsinki/ Finland. Effect estimates across differently lagged and cumulated submicron PNCs for all cardiac arrests were close to zero and mostly positive but never significantly elevated for UFP and for accumulation mode particles (e.g., 2h-lagged OR: 1.04 (0.98; 1.10) per 1,007 accumulation mode particles/ml). For myocardial infarction, ORs were frequently positive in relation to PNCs, being significantly elevated for the average exposure of the previous 24 hours (OR of PNCs < 100 nm: 1.27 (1.05; 1.54) per 10,624 particles/ml and for accumulation mode particles: 1.19 (1.04-1.35) per IQR). Another Scandinavian study (Wichmann et al., 2013), assessing out-of-hospital cardiac arrests in relation to different particle metrics sized 10-700nm (PNC, PAC, PVC) in 4,657 patients from Copenhagen, found non-significant effect estimates close to zero.

Three studies used a time-series design to examine the effects of PNC of different size fractions on cardiovscular hospital admissions (Lanzinger et al., 2016b; Samoli, Atkinson, et al., 2016) and emergency room visits (Liu et al., 2013). In the multi-site study by **Lanzinger** et al. (2016b), the relative risk for cardiovascular hospital admissions decreased slightly for an average of 2-day (percent change: -0.6 (-2.4; 1.1) and 6-day (-0.1 (-2.6; 2.4)) exposure to UFP PNC. A delayed exposure to UFP (mean average: 2-5 days) led to minimally increased RRs. Associations with PNC sized 20 to 800 nm were slightly elevated for delayed (ma 2-5 days) and cumulated (ma 0-5 days) lags. **Liu** et al. (2013) explored different size fractions of PNC in relation to total cardiovascular emergency room visits. The authors found delayed (ma 0-10 days) associations between number concentration of ultrafine particles and cardiovascular emergency room visits, mainly from lag 4 to lag 10, mostly contributed by 10–30 nm and 30–50 nm particles. Increase in 2-day average number concentration of 30–50 nm particles led to a 2.4% (1.5–6.5%) increase in cardiovascular emergency room visits per IQR of 2,269 particles/ml. A study investigating quasi-UFP-related effects on cardiovascular hospital admissions (Samoli, Atkinson, et al., 2016) reported non-significantly increased effect estimates for 15 to 64-year old London residents (percent changes: 0.81 (-0.78; 2.42) and close to zero effect estimates for residents equal/older than 65 years for 1-day lagged exposures.

[17] ST segment on an electrocardiogram normally represents an electrically neutral area of the complex between ventricular depolarization (QRS complex) and repolarization (T wave

Three studies investigated independent effects of UFPs with adjustment for co-pollutants (Lanzinger et al., 2016b; Rosenthal et al., 2013; Samoli, Atkinson, et al., 2016). **Lanzinger** et al. (2016) explored $PM_{2.5}$ and NO_2-adjusted effect estimates for delayed (mean average: 2-5 days) UFP exposure. The close to zero single-pollutant effect estimates decreased to the null upon adjustment for $PM_{2.5}$ or NO_2. **Rosenthal** et al. (2013) conducted multi-pollutant models for UFPs and accumulation mode particles, adjusting for $PM_{2.5}$, in relation to out-of-hospital cardiac arrest due to myocardial infarction and due to other cardiac causes. For accumulation mode particles associated with myocardial infarction, most effect estimates declined upon adjustment for $PM_{2.5}$, strongest for the average of 24 previous hours and the average of the lags 0 to 3. For UFPs, associated with myocardial infarction, the results were inconsistently slightly decreasing or remaining similar for short-term lags of 0, 1, 2 and 3 hours and the average of the 7 previous hours and slightly increasing for 1- to 3-day lagged exposures and the average of the previous three days. Furthermore, Rosenthal et al. (2013) conducted two-pollution models for PNC and accumulation mode particles adjusting for O_3 in relation to other cardiac causes. For accumulation mode particles, most effect estimates declined. For UFPs, the results were again inconsistent across the different lags and averaged time periods. No adjustments for NO_2 were conducted. In the study by **Samoli**, **Atkinson** et al. (2016), effect estimates for immission factors derived from particle size and number distribution decreased for background and nucleation sources and increased for traffic and secondary sources after adjustment for the other factors.

Summary of short-term associations with emergency department visits/hospital admissions

The evidence base for UFP-related effects on utilization of the healthcare system due to respiratory symptoms is scarce (Tables A1c, A3c). Possible associations seem to be most probable for children as a susceptible subgroup. While single-pollutant associations were observed in few studies, multi-pollutant models of the studies could not verify independent associations of UFPs/quasi-UFPs with respiratory hospital admissions/emergency department visits. Specifically adjustment for NO_2 led to a decrease in estimates, which mostly reached the null in co-pollutant models.

Most studies investigating cardiovascular disease-related use of the healthcare system indicate weak associations being stronger for shorter time lags of up to 24 hours. These associations decreased upon adjustment for co-pollutants with no clear evidence for independent associations of UFPs/quasi-UFPs with cardiovascular emergency department visits/hospital admission. .

Table 18: Summary table of conducted analyses in the seven studies on emergency department visits/hospital admissions

Study	Respiratory	Single pollutant associations	Multi-pollutant associations	Cardiovascular	Single pollutant associations	Multi-pollutant associations
Evans et al., 2014	✓	(+)	(+) (no NO_2 adjustment)			
Gardner et al., 2015				✓	(+)/0	nc
Iskandar et al., 2012	✓	(+)	0			
Rosenthal et al., 2013				✓	(+)/+	0
Wichmann et al., 2013				✓	(+)/0	nc

Delfino et al., 2014	✓	nr	nc			
Diaz-Robles et al., 2014	✓	+				
Lanzinger et al., 2016	✓	(+)	0	✓	(+)/0	0
Samoli UK, 2016	✓	(+)/(-)	(+)	✓	(+)	(-)/(+)
Samoli EU, 2016	✓	(+)/(-)	(-)/-			
Liu et al., 2013				✓	+/(+)	nc

0 indicates no association. (+) and (-) indicates primarily non-significant associations, + and - indicate significant associations. Nc: not conducted.

4.3.1.4 Subclinical outcomes

Respiratory markers

Lung function related indices were investigated in one cross-sectional study in Denmark (Karottki et al., 2014), one case-crossover study in Australia (Cole-Hunter et al., 2013), four panel-studies, conducted in South Korea (Song et al., 2013), Taiwan (Y. R. Li et al., 2016), Denmark (Karottki et al., 2015) and Atlanta/ USA (Sarnat et al., 2014) and five scripted exposure studies, conducted in Atlanta/ USA (Mirabelli et al., 2015), California/ USA (Jarjour et al., 2013; Park et al., 2017) and The Netherlands (Janssen et al., 2015; Strak et al., 2012) (Table A1d). Three of these studies applied two-pollutant-models (Table A3d).

In a cross-sectional approach, **Karottki** et al. (2014) found non-significant, slightly elevated effect estimates for the ratio of forced expiratory volume in 1 second (FEV_1) to forced vital capacity (FVC) in relation to central-site measured PNC sized 10 to 280 nm. Using a case-crossover design, **Cole-Hunter et** al. (2013) found non-significant lower peak-flow rates for high versus low inbound traffic exposure in healthy cycling adults.

Four panel studies investigated UFP effects on peak flow rates and spirometry indices in different groups. The study by **Song** et al. (2013) observed lower peak flow rates for central-site measured PNC sized 11 nm to 110 nm for 1-day lagged exposures in children with atopic diseases. For PNC sized 111 to 930 nm. Song et al. (2013) found consistently decreased peak flow rates for 1-day lagged exposures as well as exposures averaged over two to three days among children with atopic disease. Peak flow rates increased in response to 1-day lagged exposures of PNC sized 111 to 930 nm and decreased in response to cumulated exposures in healthy children. Associations were similar for particles sized 11 to 110 nm In contrast, **Li** et al. (2016) observed an increase in lung function indices FVC, FEV_1, and different FEF-values related to increases in UFP and accumulation mode particles of the previous day in children with asthma or allergic rhinitis. When separating different contributors, only secondary aerosol contributors yielded decreased lung function indices (in contrast do diesel vehicle emissions and aged vehicle emissions). **Karottki** et al. (2015) found decreases in the lung function parameter FEV_1:FVC (percent changes: -4.0 (-8.1; 0.5) per 3,000 particles/ml) in a panel of 48 adults for a lag period of 48 hours. **Sarnat** et al. (2014) observed slightly elevated FEV_1 levels relative to baseline levels among asthmatic participants at the 1 h and 2 h post-commute time points.

Five scripted exposure studies investigated effects of PNC with varying size fractions on lung function indices both in vulnerable as well as healthy participants. **Mirabelli** et al. (2015) observed slightly reduced FEV_1 % predicted values per IQR of mobile measured total PNC in young asthmatic adults. This adverse association could not be observed in young non-asthmatic participants. In a study comparing health effects of high versus low traffic routes in healthy regular

cyclists, the authors did not observe significant changes in spirometry indices (Jarjour et al., 2013). On the other hand, a study with healthy cyclists in California/USA (Park et al., 2017) found significant adverse associations between mobile measured PNC > 10 nm and the spirometry indices FVC, FEV_1 and PEF in 32 healthy cyclists immediately after exposure. However, for FEV_1: FVC, the authors did not find any associations. Similarly to Park et al. (2017), a study within the framework of the RAPTES-study in the Netherlands (Janssen et al., 2015) found significant adverse associations between exposure to mobile measured total PNC and percentage change in FEV_1 2 hours after exposure in outdoor sites (-1.5 per 23,000 particles/ml) in healthy students. Another study of the RAPTES project (Strak et al., 2012) indicated significant associations between PNC and FVC immediately after exposure (-1.19 per 23,000 particles/ml).

Three studies adjusted for co-pollutants: The decrease in lung function in response to quasi-UFP in the RAPTES-study by **Janssen** et al. (2015) for FEV_1 remained unchanged upon adjustment for $PM_{2.5}$ and PM_{10} but lost significance upon adjustment of NO_2. **Strak** et al. (2012) observed independent inverse associations between PNC and FVC after adjusting for PM_{10} and $PM_{2.5}$. Adjustment for NO_2 led only to slight decreases of the effect estimate. **Li** et al. (2016) observed a similar protective effect of 1-day lagged UFP exposures on FVC, FEV_1, and different FEF-values upon adjustment for O_3. However, models with accumulation mode particles adjusted for O_3 yielded non-significant inverse associations.

Summary of subclinical respiratory endpoints

Most of the above reviewed studies have only limited sample sizes (15-84 participants). Moreover, study samples were frequently selective, either representing healthy young adults or persons suffering from atopy and/or asthma. The investigated lags and averaging periods differ across studies, but generally, most associations were found in a time range of 0-48 hours after increased exposure. Finally, results of the studies are mostly inconsistent in relation to the specific respiratory endpoints. With regard to peak-flow endpoints, measurement error could be an issue in this self-monitored endpoint, especially in the study by Cole-Hunter et al. (2013) which could not be blinded. Due to the lack of adjustment for co-pollutants, little can be concluded regarding the independence of effects. The scarce evidence on studies with co-pollutant adjustment suggests an at least partial overlap of UFP, respectively PNC effects with NO_2-effects.

Blood pressure indices

Blood pressure indices have been assessed in a panel analysis within a cohort study located in Massachusetts/USA (Bind et al., 2016), seven panel studies located in Massachusetts/USA (Chung et al., 2015; Hoffmann et al., 2012), Beijing (Gong et al., 2014; J. Zhang et al., 2013) (Gong et al., 2014; J. Zhang et al., 2013), New York/USA (Rich et al., 2012; M. Wang et al., 2016), Belgium (Pieters et al., 2015), and four scripted exposure studies, conducted in Spain (Kubesch et al., 2015), New Jersey/USA (Laumbach et al., 2014), Canada (Weichenthal et al., 2014) and China (Langrish et al., 2012) (Table A1d). Four of the studies investigated effects in two-pollutant models (Pieters et al., 2015; Rich et al., 2012; Weichenthal et al., 2014; J. Zhang et al., 2013) (Table A3d).

The cohort study by **Bind** et al. (2016) investigated effects of central-site measured PNC sized 7-3,000 nm on various cardiovascular endpoints in a sample of 1,112 veterans. Bind et al. observed increased levels of diastolic blood pressure (DBP) and systolic blood pressure (SBP) in in response to extended concentrations of PNC.

While associations for DBP was consistently positive, SBP was positive in the lower and medium quantiles of the outcome's distribution (e.g., 10th percentile estimate = 4.9 mmHg (1.4; 8.6)).

Chung et al. (2015) investigated PNC related effects on systolic, diastolic and pulse pressure (PP) in 220 participants residing near a highway. The authors observed significantly increased DBP of 2.4 mmHg per 10,000 particles/ml in response to central-site measured PNC during the averaged 24h prior to measurement. The association with SBP was non-significantly positive, with pulse pressure slightly inverse. **Gong** et al. (2014) investigated associations of single-site measured PNC and AccMPs with blood pressure in a panel of 125 healthy young Chinese adults. The authors observed inconsistent associations for SBP, DBP and HR in response to PNC sized 13 to 108 nm across 0 to 6-day lagged exposures in young adults. One significant positive association was observed for a 4-day lagged exposure to PNC associated with SBP. A panel-study with 125 young adults indicated associations between quasi-UFPs and SBP for 3- and 4-day lagged exposures, being significant at 4-day lagged exposure. Effect estimates for 0- to 3-day and 5- to 6-day lagged exposures were close to zero (J. Zhang et al., 2013). **Hoffmann** et al. (2012) found non-significantly increased SBP per IQR of central-site measured PNC total averaged over the 1–5 days before examination in a panel of diabetic persons. Effect estimates for DBP were close to zero and only slightly elevated for central mean pressure. **Rich** et al. (2012) studied effects of PNC sized 10 to 100 nm and AccMPs on different blood pressure indices in a panel of 76 individuals with previous myocardial infarction or unstable angina. The authors reported increases in SBP at almost all lags ranging from 1 to 5 days, being significant for a lag of 24-47h (ß-estimates: 0.89 mmHg (0.06; 1.72) per 2,680 particles UFP/ml and 0.94mmHg (0.02; 1.87) per 897 accumulation mode particles/ml). In the same panel, **Wang** et al. (2016) explored effects of UFPs and AccMPs in relation to AccMPs lagged for 5 hours up to 4 days. The authors found consistent increases in SBP, being most pronounced for 0-23h lagged exposures (ß-estimates: 1.48 (0.09; 2.86)). In relation to UFP PNC exposure, SBP decreased for 5h and 4-day lagged exposures and increased for 0-23h up to 72-95h, being strongest for 0-23h (1.38 (0.07; 2.68)) and 24-47h lags (1.60 (0.32; 2.89)). A study exploring UFP effects on SBP in schoolchildren observed most pronounced effects of increased BP estimated for the PNC size fraction 20-30 nm (**Pieters** et al., 2015). Whereas increases in PNC total were linked to elevated SBP measurements, PNC sized < 100 nm did not have an effect. DBP was not associated with PNC.

Four scripted exposure studies measured indices of autonomic function applying mobile measured PNC. **Kubesch** et al. (2015) found statistically significant increases in SBP and DBP in healthy adults 2 hours post exposure of PNC sized 10-1,000 nm. **Laumbach** et al. (2014) did not find any associations at any time point after a ride of healthy adults in a passenger vehicle. In a study with 53 middle-aged female cyclists, **Weichenthal** et al. (2014) observed borderline significant increases in DBP in relation to PNC sized 10 to 100 nm and slightly positive associations with DBP three hours post exposure. **Langrish** et al. (2012, study described in 4.3.1.2) found statistically significant differences in mean arterial blood pressure, but not in DBP, SBP in participants wearing a respiratory mask in comparison to participants not wearing a mask.

Only four studies examining associations between UFP and blood pressure indices applied two-pollutant models (Pieters et al., 2015; Rich et al., 2012; Weichenthal et al., 2014; J. Zhang et al., 2013). **Pieters** et al. (2015) found nearly unchanged adverse effects of PNC on SBP after adjusting for PM_{10}. **Zhang** et al. (2013) observed only slightly decreased SBP values upon adjustment for CO, O_3 and SO_2. The decrease was slightly more pronounced with NO_2 used as co-pollutant. **Weichenthal** found stronger effects for SBP, DBP upon adjustment for $PM_{2.5}$, NO_2 and O_3. **Rich** et al. adjusted single pollutant models of AccMP for $PM_{2.5}$. The authors observed reduced estimates for SBP upon adjustment for $PM_{2.5}$.

Summary of blood pressure indices

The majority of studies found adverse associations between blood pressure indices and exposure to UFP/quasi-UFP, indicating increases in BP. These results differed across different endpoints (SBP, DBP, PP), different size fractions and lag periods. Apart from one study with more than 1,000 participants, the studies consisted of smaller study populations. In addition, all study samples represented selected group, impeding a transfer to the general population. Apart from these limitations, the evidence from two-pollutant studies is too scarce to draw conclusions on independent UFP effects on blood pressure indices.

HRV indices

HRV indices have been assessed in a panel analysis within a cohort study located in Massachusetts/USA (Bind et al., 2016), 11 panel studies located in Spain (Cole-Hunter et al., 2016), Germany (Hampel et al., 2012; Peters et al., 2015), California/USA (Bartell et al., 2013), New York/USA (Rich et al., 2012; M. Wang et al., 2016), Georgia/ USA (Sarnat et al., 2014), China (Gong et al., 2014; Sun et al., 2015; J. Zhang et al., 2013) and Taiwan (Wu et al., 2012). Furthermore, four scripted exposure studies were conducted in New Jersey/USA (Laumbach et al., 2014), Canada (Shutt et al., 2017; Weichenthal et al., 2014) and China (Langrish et al., 2012). Finally, one ultra-short-term panel study (Hampel et al., 2014) was reviewed (Table A1d). Five studies examining HRV indices in response to UFPs applied two-pollutant models (Table A3d).

The evidence on HRV is mixed with studies showing adverse associations for at least single indices (Bind et al., 2016; Cole-Hunter et al., 2016; Gong et al., 2014; Hampel et al., 2012; Hampel et al., 2014; Shutt et al., 2017) and those with null or even protective associations (Bartell et al., 2012; Laumbach et al., 2014; Wu et al., 2012; Weichenthal et al., 2014; Zhang et al., 2013). Several studies on HRV were conducted in susceptible groups (i.e. diabetics or patients with coronary artery disease), showing mostly small, and in part significant adverse associations with various markers of HRV (Hampel et al., 2012; Peters et al., 2015; Sun et al., 2015; Rich et al.,2012; Wang et al., 2016; Sarnat et al., 2014; Langrish et al., 2012). Most associations of UFPs or quasi-UFPs with indicators of HRV were within a very short to short time-frame (5 minutes to 95 hours), with two studies showing changes within minutes of increased exposure (Peters et al., 2015; Hampel et al., 2014). One study investigated HRV indices in a medium-term time-frame (Bind et al., 2016).

Five studies examining associations between UFP and HRV indices applied two-pollutant models (Peters et al., 2015, Rich et al.,2012, Sun et al., 2015, Weichenthal et al., 2014, Zhang et al., 2013). **Peters** et al. (2015) observed unchanged effects for HR, the standard deviation of all NN beat interval (SDNN) and root mean square of the sucessive differences in ms.RMSSD upon adjustment for ambient $PM_{2.5}$. **Rich** et al. adjusted single pollutant models of AccMP for $PM_{2.5}$. The authors observed increased effect estimates for Tpeak– Tend (Tp/Te) and similar effect estimates for heart rate turbulence upon adjustment for $PM_{2.5}$. **Sun** et al. (2015) found similar UFP-related effect estimates for SDNN upon adjustment for NO_2 and O_3. In the study by **Weichenthal** et al., (2014) associations with SDNN decreased by more than half in size in two-pollutant models upon adjustment for $PM_{2.5}$, NO_2 and O_3. **Zhang** et al. (2013) adjusted his models for NO_2 and O_3. Whereas effect estimates for HR remained similar, effect estimates for high-frequency power turned negative upon adjustment for NO_2 and decreased upon adjustment for O_3.

Summary of HRV studies

A relatively large body of evidence is available for HRV indices, mostly observing effects on at least on one HRV outcome. This evidence base adds to prior evidence from the HEI report

(2013), which already included six studies with mixed results. Upon adjustment for co-pollutant, associations changed in both directions. Across studies, different time-windows and different co-pollutants were examined, so that no clear pattern can be observed.

Indices of arrythmia

Bartell (2013) assessed PNC (5-3,000 nm) related to arrhythmia and HRV (Table A1d). The authors observed non-significant associations, with a slight decrease in daily ventricular tachycardia (VT) counts and mostly positive effect estimates and inconsistent effect estimates for hourly VT absence/ presence across different lags and for daytime versus nighttime. One, three and five-day lagged PNC was inversely associated with daily ventricular tachycardia. Day- and nighttime measured hourly tachycardia for 1-hour to 5-day lagged PNC yielded non-significant inverse associations for 8- and 24h lagged PNC with hourly daytime ventricular tachycardia all hourly nighttime ventricular tachycardia with the exception of 8 hours. On the other hand, associations with $PM_{0.25}$ and daily VT counts yielded consistent and stronger effect estimates across 0-, 1- and 2-day lagged exposures. No adjustment for co-pollutants was conducted.

Summary of studies on arrhythmias

Considering the limited number of studies with only one study, the evidence base is still insufficient.

Indices of vascular function

Vascular function has been assessed within one cohort study (Amar J. Mehta et al., 2014), two cross-sectional studies (Karottki et al., 2014; Ljungman et al., 2014) three panel studies (Karottki et al., 2015; Zanobetti et al., 2014; X. Zhang, Staimer, Tjoa, et al., 2016) and one scripted exposure study (Weichenthal et al., 2014) (Table A1d). Two of these studies further adjusted for co-pollutants (Zhang, Staimer, Tjoa, et al., 2016; Weichenthal et al., 2014) (Table A3d).

Seven studies considered effects of UFP in relation to vascular function indices. One cohort study (Amar J. Mehta et al., 2014) with 370 elderly US veterans investigated associations between central-site measured PNC sized 7-3,000 nm and augmentation pressure and augmentations index. The authors observed consistent increases per IQR for both endpoints, being significant for mean average lag periods of 1, 3 and 14 days. Two cross-sectional studies conducted in Denmark (Karottki et al., 2014) and in Massachussets (Ljungman et al., 2014) explored central-site modelled PNC in relation to vascular endpoints. **Karottki** et al. (2014) observed a significant decrease in microvascular function in relation to 2-day lagged central-site measured PNC sized 10-280 nm in 49 adults. In a large sample of 2,072 adults being part of the Framingham Heart Study Offspring and Third Generation Cohorts, **Ljungman** et al. (2014) analyzed micro-vessel function measured by peripheral-arterial tonometry in relation to total PNC. The study showed significant increases in baseline pulse amplitude, but did not observe any consistent patterns for hyperemic response measured by PAT ratio across mean average exposures of 1 to 7 days.

Three panel studies investigated central-site measured UFP related effects on vascular function. In a panel of 48 adults in Denmark, **Karottki** et al. (2015) found significant decreases in MVF in relation to central-site measured PNC sized 20-280 nm. Incontrast, **Zanobetti** et al. (2014) did not observe associations between total PNC and brachial artery diameter in a panel of 64 adults with type 2 diabetes mellitus (T2DM). **Zhang** et al. (2016b) found no associations between measured $\mathbf{PM_{0.18}}$ in relation to reactive hyperemia index (RHI), but for $PM_{2.5}$.

Finally, one scripted exposure study (**Weichenthal** et al., 2014) applying mobile measurements of PNC sized 10 to 100 nm in 53 healthy female cyclists in Canada found significant decreases in RHI in response to UFP exposure.

Only two of the studies investigating vascular effects applied two-pollutant models. In the study by **Zhang** et al. (2016b), previously zero effect estimates for RHI turned in a positive direction upon adjustment for O_3. In the study by **Weichenthal** et al. (2014), effect estimates for RHI remained significantly inverse upon adjustment for $PM_{2.5}$ and NO_2.

Summary of studies on vascular function

The majority of the studies examining associations between UFP/quasi-UFP and vascular function indicate a possible association. However, a lack of consistency regarding the study design, specifically the outcome parameters, as well as missing co-pollutant models do not allow overall conclusions.

Biomarkers - Pulmonary inflammation

Nine panel studies located in Beijing (Gong et al., 2014; J. Zhang et al., 2013) Shanghai (Han et al., 2016), The Netherlands (Manney et al., 2012), Belgium, (Pieters et al., 2015), Boston/ USA (Peng et al., 2016) Atlanta/ USA (Sarnat et al., 2014), Atlanta/USA (X. Zhang, Staimer, Gillen, et al., 2016) and four scripted exposure studies located in New Jersey/ USA (Laumbach et al., 2014) , Georgia/USA (Mirabelli et al., 2015), Belgium (Bos et al., 2013) and The Netherlands (Strak et al., 2012) investigated UFP related effects on markers of pulmonary inflammation. Pulmonary indices were fractional exhaled nitric oxide (FeNO), nitric oxide in exhaled breath condensate pH (NO_x EBC), nitrite and/or nitrate, malondialdehyde (MDA) and IL-1ß in exhaled breath condensate (Table A1d). One of the studies adjusted for co-pollutants (Table A3d).

Nine studies investigated effects of UFPs and quasi-UFPs on fractional exhaled NO with time windows ranging from immediately after exposure up to 7 days after exposure. Five central-site measured panel studies (Gong et al., 2014; Han et al., 2016; Peng et al., 2016; X. Zhang, Staimer, Gillen, et al., 2016) indicate positive associations for shorter lag periods. A panel study using a microscale personal exposure model (Sarnat et al., 2014) also found significant associations. One panel study assessing $PM_{0.18}$[18], averaged across five days, observed significant positive effect estimates for exhaled NO in relation to $PM_{0.18}$ (X. Zhang, Staimer, Gillen, et al., 2016). Three scripted exposure studies using mobile measurements found significant increases immediately after exposure to total PNC (Bos et al., 2013; Mirabelli et al., 2015; Strak et al., 2013). However, all studies focusing on fractional exhaled NO used selected groups and PNC of different size fractions.

Five studies explored UFPs and quasi-UFPs in relation to MDA in exhaled breath condensate. The effects were less pronounced than for exhaled NO. Two panel studies (Gong et al., 2014; Sarnat et al., 2014) found inconsistent associations in response to PNC across lags, ranging from significant inverse to non-significant positive, or being non-significantly positive. A panel study assessing quasi-UFP $PM_{0.18}$ averaged across five days, found non-significant positive associations (X. Zhang, Staimer, Gillen, et al., 2016). A scripted exposure study by Mirabelli et al. (2015) observed slightly positive associations immediately after exposure.

[18] $PM_{0.18}$ is a standard UFP metric measured by MOUDI impactors. It describes particles with a cut-off diameter of 18 nm.

Few studies assessed nitrite (Gong et al., 2014; Laumbach et al., 2014), or nitrite and nitrate (Laumbach et al., 2014; Manney et al., 2012) in exhaled breath condensate. In summary, effect estimates for EBC nitrite + nitrate were mostly positive (except for (Manney et al., 2012)) and were strongest for shorter lag-periods, i.e. immediately after exposure (Laumbach et al., 2014) or in 0 to 2-day lagged exposures (Gong et al., 2014). Most effect estimates were inverse or null for EBC NO_x in Manney et al. (2012), suggesting a decrease in pulmonary inflammatory markers.

One study (Pieters et al., 2015) observed significant increases in the pulmonary inflammation marker IL-1ß in exhaled breath condensate for concurrent exposure to UFP PNC, being strongest for the smallest fraction of 20-30 nm.

Two-pollutant models were conducted in the study by Strak et al. (2012). Significant effect estimates remained unchanged after adjustment for $PM_{2.5}$, and increased after adjustment for NO_2, immediately after exposure. In the study by Zhang et al. (2013) effect estimates remained significant and attenuated only slightly upon adjustment for NO_2, O_3 and SO_2.

Summary of studies on pulmonary inflammation

The studies which have been investigated UFP-effects on pulmonary inflammations suggest positive associations between UFP and adverse changes in the pulmonary inflammation marker, in particular immediately after exposure. Nevertheless, the evidence base for pulmonary inflammation in response to UFP is still limited as the studies used different subgroups, exposure metrics, outcome measures and time frames. The two studies that conducted two-pollutant models observed overall robust effect estimates.

Biomarkers - Systemic inflammation

One cohort study located in Massachusetts/ USA (Bind et al., 2016), two cross-sectional studies located in Massachusetts/ USA (Fuller et al., 2015) and Denmark (Karottki et al., 2014), eleven panel studies conducted in New York/ USA (Croft et al., 2017; Rich et al., 2012; M. Wang et al., 2016), Finland (Huttunen et al., 2012), Denmark (Karottki et al., 2015), Germany (Rückerl et al., 2014; Rückerl et al., 2016), Georgia/USA (Sarnat et al., 2014), California/ USA (Wittkopp et al., 2013), China (Gong et al., 2014; J. Zhang et al., 2013), one Australian case-crossover study (Cole-Hunter et al., 2013) plus four scripted exposure studies conducted in Belgium and the Netherlands (Bos et al., 2013; Steenhof et al., 2014; Steenhof et al., 2013; Strak et al., 2013) investigated UFP-related effects on systemic inflammation markers.

The above named studies used various markers of systemic inflammation C-reactive protein (CRP), interleukin-6 (IL-6) and myeloperoxidase are the most commonly used markers of systemic inflammation, while fibrinogen was the most investigated measure for coagulation. Systemic inflammation and coagulation are considered as mediators for cardiovascular diseases.

The most common investigated inflammatory marker among our identified studies was hs-CRP. The largest study sample in which hs-CRP measures were realized was a repeated measures study which was embedded in the Normative Ageing Cohort study including 1,112 veterans (Bind et al., 2016). The quantile regression yielded significantly elevated hs-CRP values for the 70. to 90th percentiles (referring to the distribution of CRP in the cohort) per IQR of central-site measured total PNC averaged over 28 days. The lower quantiles yielded zero to non-significant positive effect estimates. The cross-sectional studies found mostly non-significant positive associations between central-site measured quasi-UFP PNC and hs-CRP with exposure time windows of averaged 28 days (Fuller et al., 2015) and a 2-day lag (Karottki et al., 2014). Fuller et al.

(2015) observed significantly decreased CRP estimates in relation to a second, near highway measurement site. However, this second measurement site had considerable missings in exposure data. Of the eleven mostly central-site measured PNC-related panel-studies, five studies (**Croft** et al., 2017; **Huttunen** et al., 2012; **Karottki** et al., 2015; **Rich** et al., 2012; **Sarnat** et al., 2014; **Rückerl** et al., 2014; **Wang** et al. (2016); Wittkopp et al. (2013)) did not observe strong associations. The effect estimates ranged mostly from slightly inverse to positive relations being significant for single lag periods. Rich et al. (2012), Wang et al. (2016) and Wittkopp et al. (2013) found the strongest associations for 1 and 2 day lag periods and Huttunen et al. (2012) for a lag period of 3 days. One scripted exposure study (**Strak** et al. (2013) found non-significant inverse relationships between total PNC and CRP 25 hours post exposure.

Of the studies investigating UFP-related effects on hs-CRP and described above, two adjusted their models for co-pollutants. In the study by **Rückerl** et al. (2014), associations lost significance after adjustment for $PM_{2.5}$. **Strak** et al. (2013) observed decreased effect estimates upon $PM_{2.5}$ which gained significance upon adjustment for NO_2. In contrary, adjustment for PM_{10} yielded less inverse associations.

The evidence base from the studies with regard to fibrinogen is inconsistent. Mostly positive associations were found in the studies by **Croft** et al. (2017) being strongest for shorter lag periods (0-11h, 0-47h), **Rich** et al. (2012) and **Wang** et al. (2016), most pronounced for lag hours 24-47. Inconsistent effect estimates were found in **Bind** et al. (2016), showing highest effect estimates for the 10., 20. and 90th percentile. **Gong** et al. (2014) and **Zhang** et al. (2013) found inconsistent effect estimates across 0 to 6-day lags, without a specific direction. In other studies, mostly inverse associations were observed (**Huttunen** et al. (2012) and **Strak** et al. (2013)). The inconsistent evidence base may originate from different subgroups being selected and different exposure metrics and time periods being assessed.

Four studies investigating UFP-related effects on fibrinogen in two-pollutant models. In the study by **Strak** et al. (2013), the negative effect estimates for fibrinogen even decreased upon adjustment for $PM_{2.5}$ and PM_{10}. Upon adjustment for NO_2, effect estimates became significantly inverse. Similar to Strak et al. (2013), **Zhang** et al. (2013) found decreased effect estimates upon adjustment for NO_2, SO_2, O_3 and CO. **Croft** et al. (2017) observed significantly increased effect estimates upon adjustment for $PM_{2.5}$, Delta-C and BC. **Rich** et al. (2012) found unchanged significant positive effect estimates for accumulation mode particles with a lag period of 24-47 hours upon adjustment for $PM_{2.5}$. As expected, estimates became insignificant upon adjustment for UFP and accumulation mode particles (vice versa).

Some panel and scripted exposure studies (Gong et al., 2014; Huttunen et al., 2012; Karottki et al., 2014; Rich et al., 2012; Steenhof et al., 2013; J. Zhang et al., 2013) measured blood cell counts as markers for systemic inflammation. The short-term studies indicate inconsistent results for the most frequent used white blood cell counts. In a group of young university workers, **Gong** et al. (2014) and **Zhang** et al. (2013) found positive association for shorter lag periods of 0 and 1-days, being significantly elevated at lag 0. However, 2 to 6-day lagged exposures to UFP PNC yielded inverse associations. **Huttunen** et al. (2012) oberserved non-significant effect estimates being lowest for 0-day lagged and most pronounced for 2-day lagged exposure to quasi-UFP PNC in a panel of elderly IHD patients. The study by **Rich** et al. (2012) suggests slightly positive associations between UFP PNC and white blood cell counts and generally (except from 4-day lag) slightly inverse associations with accumulation mode particles for 0 to 4-day lag periods. **Steenhof** et al. (2014) found significantly positive associations between quasi-UFP PNC and white blood cell count 2 hours post exposure for all and outdoor sites and significantly negative associations 18h post exposure. The same pattern was observed for neutrophiles. However, mono-

cyte, lympohocyte and eosinophile counts were consistently and significantly inverse associated. **Karottki** et al. (2014) found also (non-significant) associations for neutrophils and adverse associations for leukocyte, monocyte and lymphocyte counts. Eosinophiles were positively associated with 2-day lagged PNC sized 10-280 nm.

Associations with myeloperoxidase were explored in three studies (Croft et al., 2017; Huttunen et al., 2012; Rückerl et al., 2014). **Croft** et al. (2017) observed consistently inverse associations, being significant for a lag period of the previous 12 hours. **Huttunen** et al. (2012) found non-significant positive associations for 0 to 3-day lagged exposures. In the group of genetically susceptible persons, **Rückerl** et al. (2014) found significant positive associations for a 5-day averaged exposure to UFP PNC. However, effect estimates were non-significant inverse in the group of individuals suffering from T2DM or IGT.

Numerous studies investigated further inflammatory markers. **Bind** et al. (2016) found increasing effect estimates for Interleukin-6 along larger averaged lag periods, finally being significant at a mean average of 28 days. TNR-II was positively associated with PNC for lag periods from 7 to 28 days in comparison to lag days 0 to 3. Furthermore, the intercellular adhesion molecules-1 ICAM-1 was consistently and significantly positive associated with PNC whereas associations with the vascular cell adhesion molecule-1 VCAM-1 were less pronounced across the different quantiles. **Huttunen** et al. (2012) investigated the inflammation markers interleukin IL-8 and IL-12. He found strongest associations for 3-day lagged exposures compared to 0 to 2-day lagged exposures. 3-day lagged quasi-UFP PNC associations with IL-12 was significantly elevated. For urinary MDA and urinary 8-OHdG. **Wittkopp** et al. (2013) found non-significantly inverse associations for IL-6sR as well as positive associations for tumor necrosis factor TNFRII for 0 to 5 day-lagged exposures. In a scripted exposure approach, **Steenhof** et al. (2013) observed elevated levels of IL-6 in healthy participants 2h after exposure in a real-world scenario. **Gong** et al. (2012) found elevated effect estimates for 2 to 5-day lagged (urinary MDA) and 3- to 6-day lagged 8-OHdG.

Summary of associations with systemic biomarkers

Overall, the majority of the studies investigating UFP effects on systemic inflammation markers indicate inconsistent associations. Effects of UFP on indices for hs-CRP, fibrinogen, blood cell counts, myeloperoxidase varied, which may originate from different compositions of participants, assessed PNC fractions and exposure assessment types. In most studies, effects seem to be most pronounced for shorter lag periods. Only few multi-pollutant models do not allow statements on independent effects of UFPs/ quasi-UFPs.

Neurocognitive indices

Two scripted exposure studies conducted in Brussels (Bos et al., 2013; Bos et al., 2011) explored associations between mobile measured quasi-UFP exposures and serum brain-derieved neurotropic factor (BDNF) in a short-term (Bos et al., 2011) and medium-term time-frame of 12 weeks (Bos et al., 2013). In the later study, Bos and collegues additionally assessed cognitive tests. Whereas the short exposure of 20 minutes cycling along a busy road did result in any association with BDNF, the 12-week aerobic training program resulted in higher BDNF-levels differed significantly in participants exercising in a rural area versus participants exercising in an urban area. Due to the lack of co-pollutant adjustment, it is not possible to disentangle UFP associations from other pollutant effects.

Table 19: Summary table of conducted analyses in the 55 studies on subclinical outcomes

Outcome	Number of studies	Number of studies with single-pollutant--associations in expected direction	Number of studies with multi-pollutant associations in expected direction	Comments (i.e. studies with significant results in the non-expected direction)
Respiratory indices	11	4/11	3/3	Li et al. (2016) found significantly positive associations between UFP and FEV_1 & FVC
Blood pressure	13	9/13	2/4[19]	Two of the nine studies with associ. showed inconsistent results across lags
HRV	16	12/16	3/5	In Zhang et al. (2013), effect estimates decreased upon adj. for NO_2 and increased upon adj. for O_3
Arrhythmia	1	1/1	nc	Strong associations with $PM_{0.25}$, nearly protective associations between PN and hourly nighttime measured tachycardia
Vascular function	7	4/7	1/2	
Pulmonary inflammation	12	12/12	2/2	Most studies investigated effects on FeNO
Systemic inflammation (incl. fibrinogen)	18	7/18[20]	2/5	Significant inverse associations between fibrinogen & PNC upon adjustment for NO_2 (Strak et al., 2013)
Neurocognitive outcomes	2	1	nc	-

HRV: Heart rate variability.

4.3.2 Long-term effects

4.3.2.1 Mortality

One long-term study explored relations between all-cause, cardiovascular, ischemic heart disease and pulmonary mortality (Ostro et al., 2015) (Table A2a). The large cohort study included 101,884 female participants of the California Teachers Study and was realized from the beginning of 2001 to mid-2007. Ultrafine PNC sized 10 to 100 nm was applied using a chemical-transport model taking into account meteorological fields and emissions estimates for different sources to predict airborne particulate matter concentrations. Mortality outcomes were assessed via linkage to an administrative database. The authors adjusted their models extensively for individual covariates. Ostro et al. (2015) observed slightly increased hazard ratios for all-cause mortality (1.01 (95% CI 0.98; 1.05)), cardiovascular mortality (1.03 (95% CI 0.97; 1.08)), and no associations for pulmonary mortality (95% CI 1.01 (0.93; 1.10)). Among cardiovascular causes of death, hazard ratios for IHD mortality were significantly elevated (1.10 (95% CI 1.02; 1.18)). The study did not adjust for co-pollutants but for constituents of UFP, therefore prohibiting the evaluation of independent effects of UFPs. A further limitation of the study is the lack of

[19] One of the four studies did not show assoc. in single-pollutant models, either. A further study (Rich et al., 2012) did not show all results, therefore rated as non-associated here

[20] Most positive associations relate to fibrinogen

representativeness of the participants. Moreover, the spatial resolution of the exposure model was relatively large with 4x4 km^2, which prevents the assessment of small scale differences in exposure to UFPs/quasi-UFPs, which are characterized by a high spatial variability. Underestimation of the true association is therefore possible.

4.3.2.2 Morbidity

Four studies investigated associations between morbidity and long-term UFP exposure based on chemical transport or land regression models (Laurent et al., 2014; Laurent et al., 2016a, 2016b; Li et al., 2017). All four studies were carried out in North America with two cross-sectional, one nested case control and one case-cohort design. Three of these studies investigated birth outcomes, one cardiovascular and cerebrovascular morbidity. None of them conducted co-pollutant models.

Cardiovascular/ cerebrovascular/ metabolic morbidity

Three communities near Boston, representative for highway or urban background air pollution, were selected for the CAFEH study. Data from a subset of 435 participants which attended a field clinic was cross-sectionally analyzed by **Li** et al. (2017). Mobile monitoring and spatial-temporal regression models were used to estimate PNC (>4 nm) at each residential address (resolution of 20m). Subsequently, time activity information from all participants was used to assign individualized time activity adjusted annual PNC exposure. Self-reported prevalences of stroke or ischemic heart disease, hypertension and diabetes were non-significantly associated with ORs of 1.35 (95% CI 0.83; 2.22), 0.71 (95% CI 0.46; 1.1) and OR 1.14 (95% CI 0.81; 1.62). Due to lack of adjustment for co-pollutants, the independence of UFP associations cannot be evaluated.

Birth outcomes

Two studies explored associations between low birth weight (LBW) in term born infants (>37 gestational weeks) and submicron particle mass $PM_{0.1}$. **Laurent** et al. (2014) evaluated cross-sectional data of 960,945 singleton live births in Los Angeles between 2001 and 2008. Odds of having a term LBW infant was significantly increased with OR 1.03(95% CI 1.02; 1.03) per IQR of 0.4271 µg primary $PM_{0.1}/m^3$. In addition, the model was able to deliver broad source categories for $PM_{0.1}$, namely gasoline, diesel, shipping, high sulfur combustion sources, commercial meet cooking, wood burning and other sources. The source most strongly associated with term LBW was gasoline, followed by wood burning, meat cooking, diesel and high sulfur sources. In a very similar study setting with the same exposure assessment, **Laurent** et al. (2016b) were able to partly confirm their results with a case-cohort approach with over 70,000 term born LBW infants in California between 2001 and 2008. Even though primary $PM_{0.1}$ mass was not associated with increased odds for term LBW (OR 0.996 (95% CI 0.98; 1.01) per IQR 1.359 $\mu g/m^3$), when broken down by sources, odds for term LBW were significantly higher for on-road gasoline (OR 1.05 (95% CI 1.02; 1.09) per IQR), commercial meat cooking (OR 1.03 (95% CI 1.01; 1.06) per IQR) and on-road diesel (OR 1.03 (95% CI 1.0; 1.06) per IQR). In this second study, PNC (<100 nm) was also modeled with the line-source roadway dispersion model (CALINE4), however there was no association with term LBW (OR 1.001 (0.989; 1.014) per IQR 6,444 particles/cm^3). Using a nested matched case-control approach within the same cohort, **Laurent** et al. (2016a) studied the association between preterm birth (n= 442,314) and PM applying the same exposure measurements. ORs for preterm birth in association with IQR increases in average exposure during pregnancy were significantly elevated for primary $PM_{0.1}$ and its components organic carbon (OC), elemental carbon (EC) and secondary organic aerosols (SOA): 1.02 (95% CI 1.02; 1.03)

per 1.39 µg primary $PM_{0.1}/m^3$, 1.02 (95% CI 1.01; 1.03) per 0.99 µg OC in $PM_{0.1}/m^3$, 1.04 (95% CI 1.04; 1.05) per 0.13 µg EC in $PM_{0.1}/m^3$ and 1.13 (95% CI 1.12;1.143) per 0.061 µg SOA in $PM_{0.1}/m^3$. As for sources of primary $PM_{0.1}$, strongest and statistically significant associations per IQR increase in exposure were observed for on-road gasoline, followed by on-road diesel and commercial meat cooking. However, preterm birth was negative and significantly associated with submicron particle mass from wood burning. Furthermore, inverse effect estimates were observed for the association between preterm birth and PNC in all subjects with OR 0.99 (95% CI 0.99; 1.00) for a 6,480 particles/cm^3 increase in PNC. In a separate analysis, the authors explored the influence of geocoding accuracy with a subgroup of births, geocoded at the tax parcel level. In this subgroup, the effect estimate for an IQR increase in PNC was significantly elevated with OR 1.03 (95% CI 1.02; 1.04) for 6,770 particles/cm^3. Due to lack of adjustment for co-pollutants, the independence of UFP associations cannot be evaluated.

4.3.2.3 Emergency department visits

No studies found.

4.3.2.4 Subclinical outcomes

Three cohort studies located in Switzerland (Aguilera et al., 2016), Spain (Sunyer et al., 2015) and Germany (Viehmann et al., 2015) and two cross-sectional studies located in Boston/USA (Lane et al., 2015; Lane et al., 2016) investigated long-term subclinical health effects in response to quasi-UFP and UFP exposures (Table A2c). The study by Aguilera et al. (2016) additionally adjusted for co-pollutants (Table A4b).

Cardiovascular endpoints

Aguilera et al. (2016) explored associations between land-use regression modeled PNC 10-300 and LDSA and carotid intima-media thickness in 1,503 participants of the SAPALDIA cohort located in different areas in Switzerland. The cross-sectional analysis resulted in increased effect estimates for carotid intima-media thickness (2.06% (95% CI 0.03%; 4.10%) per increment of 10. to 90th percentile) and LDSA (2.32 (95% CI 0.23; 4.48)). Upon adjustment for lifestyle variables and T2DM prevalence, SBP, HDL cholesterol and different medications, effect estimates increased. Upon further adjustment for NO_2, the effect estimate in relation to PNC turned into a negative direction. However, the effect estimate for LDSA increased but lost its significance.

Inflammation markers

Three studies investigated long-term effects of quasi-UFP on inflammatory biomarkers within a time window of one year. **Viehmann** et al. (2015) analyzed repeated measures of hs-CRP, fibrinogen, white blood cell counts and platelets in relation to PNC sized 5 to 2,200 nm modeled by a chemical transport model with a spatial resolution of 1x1km^2. In their cohort with multiple measurements of 3,213 participants of the HNR cohort, Viehmann et al. (2015) observed consistently increased effect estimates for all endpoints being most pronounced for hs-CRP (3.8% (95% CI −0.6%; 8.4%) per IQR of 27,000 particles/ml) and fibrinogen (1.0% (95% CI 0.0%; 2.0%) per IQR of 27,000 particles/ml) and white blood cell counts (1.0 (95% CI −0.1; 2.1) per IQR of 27,000 particles/ml). Two cross-sectional studies by **Lane** (2015; 2016) investigated effects of a land-use regression model in combination with a time-activity pattern on hs-CRP, IL-6, TNFRII and fibrinogen within participants of the CAFEH cohort. In the first study with 140 par-

ticipants, **Lane** et al. (2015) used personal exposure model including the residential annual average + work + other + highway + air-condition at home at specific temperature in relation to hs-CRP and fibrinogen. His models yielded positive associations for hs-CRP (ß-estimate: 1.26 (95% CI −0.02; 2.75)) and (IL-6 0.65 (95% CI −0.26; 1.55)) in the fully adjusted models (increments unclear). The subsequent study with a larger cohort of 408 participants, **Lane** et al. (2016) observed elevated effect estimates for the markers of systemic inflammation hs-CRP 14.0% (95% CI -4.6%; 36.2%), IL-6 (8.9% (95% CI -2.6%; 21.8%) and TNFRII (5.1% (95% CI -0.4; 10.9) per 10,000 particles/ml in the fully adjusted models. For the coagulation marker fibrinogen, the authors reported inverse changes (-1.9 (95% CI -5.5; 1.6)). Due to lack of adjustment for co-pollutants, the independence of UFP associations cannot be evaluated.

Neurocognitive health endpoints

A Spanish cohort study in the framework of the BREATHE[21] project investigated quasi-UFP effects on the cognitive function of 2,715 children attending schools in low and high polluted areas (Sunyer et al., 2015). PNC sized 10-700 nm was assessed at school during two measurement campaigns complemented by exposures at the home address estimated by a land-use regression model. At baseline and within one year, cognitive function in children of high polluted areas developed less in schools of high polluted areas versus low polluted areas. Quasi-UFP exposure at the courtyard was related to inverse associations in working memory, superior working memory and positive associations in inattentiveness at baseline and after 12 months being significant for superior working memory and inattentiveness (e.g., difference of 3.9 (95% CI 0.31; 7.6) per increase of 6,110 particles/ml). The models were adjusted for maternal education and socioeconomic status, but not for co-pollutants, which prevents the investigation of independent associations of UFPs/quasi-UFPs.

Summary of long-term health effects

The above described study results of long-term studies on UFP health effects are summarized in table 20. A limited number of studies, varying outcome measures and exposure assessment methods as well the lack of or co-pollutant adjustments do not allow drawing final conclusions.

Table 20: Summary table of conducted analyses in the 10 long-term studies

Outcome type/ study	Outcome	Single pollutant associations	Multi-pollutant associations
Mortality/ Ostro et al. 2015	- all-cause - cardiovascular/ IHD - pulmonary	0 (+)/0 0	nc nc nc
Morbidity / Li et al. 2017 Laurent et al. 2014/2016b Laurent 2016a	- cardiometabolic - low birth weight - preterm birth	(+) +/(+) -/+	nc nc nc
Subclinical/ Aguilera et al. 2016	- carotid-intima-media thickness (PNC/LDSA)	+/+	-/(+)

[21] Brain Development and Air Pollution Ultrafine Particles in School Children

Viehmann et al. 2015	- hs-CRP/ fibrinogen/ WBC	(+)/+/(+)	nc
Lane et al. 2015	- hs-CRP/ IL-6	(+)/(+)	nc
Lane et al. 2016	- hs-CRP/ IL-6/ TNRFIII/ fibrinogen	(+)/(+)/(+)/(-)	nc
Sunyer et al. 2016	- working memory,	(+)	nc
	- superior working memory	+	
	- inattentiveness	+	

IHD: Ischemic heart disease, 0 indicates no association. (+) and (-) indicates primarily non-significant associations, + and - indicate significant associations. Nc: not conducted.

4.4 Summary of health effects

An overview on all included short-term and long-term studies reflects the inconsistency of the results (Table 21). More than half (n=49) of the studies on short-term effects (n=79) reported at least one significant effect in the single pollutant model, especially those studying mortality or subclinical outcomes. For less than half of the single-pollutant associations (21 of 49), the general pattern of the association was consistent regardless of the significance level. The associations in multi-pollutant studies (n=32) remained consistent in about half of the studies (n=7).

Associations between health outcomes and long-term exposure with ultrafines were more consistent in the single pollutant models even though there were considerably fewer studies. Nevertheless, long-term studies adjusting for other pollutants are still lacking with only one study, which did not show effects in the multi-pollutant model.

Table 21: Summary table of all included studies in single- and multi-pollutant associations

Outcome	Single pollutant effect	Consistency of general pattern	Multi-pollutant effect	Consistency of general pattern
Short-term	49/79*	21/49	18/32	7/18
Mortality	5/7	2/5	4/6	1/4
Morbidity	3/7	0/3	-	-
Hospital admission	4/10	2/4	0/5	-
Subclinical	37/55	17/37	14/21	6/14
Long-term	8/10	1/1	0/1	-
Mortality	1/1	1/1	-	-
Morbidity	3/4	-	-	-
Hospital admission	-	-	-	-
Subclinical	4/5	-	0/1	-

5 Discussion

5.1 Literature search

We conducted a systematic comprehensive search of relevant epidemiological studies on ultrafine and quasi-ultrafine particles for the period from 01.01.2011 until 11.05.2017. The different strategies of our search consisted of a MEDLINE search, using two alternative strategies, a search in the specialized data base LUDOK, and a hand search in review articles and reference lists of identified publications. Overall, the additional yield of the alternative MEDLINE search strategy, and of the complementary search strategies (LUDOK and hand search), and of the repeated search was substantial, with altogether 15 additional references added to the final analysis data base and an additional 13 articles identified per MEDLINE and hand search in February 2018. This relatively high yield reflects the lag in indexing newly published studies in large literature data bases as well as the fast development of an emerging scientific field. More specialized data bases such as the dedicated LUDOK literature data base are therefore very useful for targeted and timely research.

5.2 Evaluation of health relevance of ultrafine particles

Our evaluation of the health relevance of ultrafine particles is based on the above described epidemiologic studies and how they add to the available the evidence since the comprehensive review conducted by the HEI, published in 2013. Overall, the epidemiological evidence is quickly increasing and it can be expected, that the next few years will bring a substantial increase in relevant studies. Currently, we are still in the beginnings of health-related research of UFPs, which is in part due to the still developing methods (see sections below on exposure assessment).

The HEI concluded in its review that "the current database of experimental and epidemiologic studies does not support strong and consistent conclusions about the independent effects of UFPs on human health" (Health Effects Institute, 2013). Major reasons for this lack of evidence, specifically for epidemiologic studies, lie in the difficulty of assessing population-based exposure to UFPs for short-term as well as for long-term studies. Due to the specific properties of UFPs with a high temporal and spatial variability, common exposure assessment strategies, which have been developed for the more homogeneously distributed larger particle fractions, will lead to larger exposure misclassification when applied to UFPs. Nevertheless, HEI does not conclude that independent effects of UFPs can be ruled out, but rather recommends the exploration of alternative exposure metrics, spatial modeling techniques, and statistical methods.

In this review, we use similar design- and outcome-specific categories as in the HEI review to be able to integrate our findings with the prior evidence. Since independence of effects is the key question regarding the health relevance of UFPs, we specifically focus on studies with co-pollutant adjustment.

Inconcistency of results by endpoint

Previous evaluations have concluded, that the combined results for respiratory as well as for cardiovascular endpoints are still inconsistent (Health Effects Institute, 2013). When considering the newly acquired evidence during the years from 2011 to 2017, this picture has not changed substantially. Even though there is a growing number of specifically designed studies to investigate health effects of UFP/quasi-UFP, we cannot identify a consistent pattern of health

effects on either respiratory or cardiovascular disease across the different endpoints including mortality, morbidity, emergency department visits/hospital admissions or subclinical endpoints. For other outcomes such as mental disorders, neurocognitive function or birth outcomes, the evidence base is still too small to derive firm conclusions.

Even though results are not consistent across different outcomes types, the majority of the 11 studies investigating short-term effects on BP, the major risk factor for cardiovascular disease, indicate an association with increased blood pressure. Once again, evidence from the three co-pollutant-adjusted studies is mixed, which underscores the necessity of further studies with co-pollutant adjustments.

The lack of consistent findings can be explained by a number of factors. These include differences in exposure assessment (see below), endpoint assessment, study design and size, and different confounder control, specifically differences in the adjustment for co-pollutants (see below).

Long-term exposure and health effects

In contrast to the last prior comprehensive review by HEI (2013), ten studies have been published investigating long-term effects of UFPs on various health outcomes. While most of these studies found elevated point estimates for associations of UFPs with adverse health outcomes, only one study adjusted for co-pollutants, including NO_2. Adjustment with NO_2 led to a decrease in the effect estimate to an inverse association.

While the current evidence base does not support an independent effect of UFPs on health outcomes, this should by no means be mistaken for a proof of the absence of such an effect. As will be discussed below, current exposure assessment techniques are not well suited to describe and investigate long-term exposure to UFPs. More studies applying novel methods for individual-level exposure to UFPs are therefore urgently needed. Important applications are next to road traffic-related exposures also the emerging problem regarding exposure to UFPs in the vicinity of airports, which has only recently been described (Hudda, Simon, et al., 2016).

Exposure assessment

Overall, the number of studies including the assessment of exposure to and the investigation of health effects of UFPs is rapidly increasing. One important factor contributing to this rapid increase is the development of new instrumentation, which enables a less expensive assessment of UFP/quasi-UFP for example with condensation particle counters. However, research is still at the beginning and new exposure assessment methods need to be defined and employed in epidemiological studies.

Challenges of exposure assessment for UFPs include the high spatial and temporal variability of UFP/quasi-UFP, which necessitate different exposure assessment designs than the “classical” air pollutants like $PM_{2.5}$ and PM_{10} with a much more homogeneous spatial distribution. This high spatial variability is of concern not only for long-term health effects studies, which are based on long-term spatial differences in exposure, but also for short-term studies with a central-site measurement. These studies assume that the temporal changes from day to day are evenly distributed across the sometimes very large study areas; an assumption that might not hold true for UFPs. Given the possibility of a larger exposure estimation error for UFPs compared to other pollutants, a systematic bias towards the null in single-pollutant studies and in multi-pollutant studies is probable (Dionisio et al., 2014).

In the future, the development of enhanced spatiotemporal models can contribute to a more precise exposure assessment across larger areas. Current models such as the German EURAD model need to be adapted to incorporate specific sources, validation measurements and increase the spatial resolution.

A further challenge of UFP/quasi-UFP exposure assessment is the non-standardized equipment and the non-standardized use of size fractions in the studies. The commonly used measurement devices have different lower cutpoints for the particle size. Since the majority of particles are located in the nucleation mode (< 20 nm) of the particle size distribution, even small differences in the lower cutpoint between 1 and 20 nm can lead to substantial differences in particle number concentration. Futhermore, the reporting of the exposure assessment often does not include the exact size range of particles, which prevents direct comparisons of exposure between studies.

Independence of effects

Even though several studies across the investigated endpoints have observed positive associations of UFP/quasi-UFP with various health effects, the overall evidence for independent effects is still insufficient. We noticed, that specifically the newer studies conduct multi-pollutant models with a higher frequency than the older studies, which is a positive development (e.g., Aguilera et al., 2016; Croft et al., 2017; Lanzinger et al., 2016a; Samoli et al., 2016; Stafoggia et al., 2017). However, the type of adjustment still varies substantially between studies and there is no standard strategy for co-pollutant adjustment yet. At the moment, adjustment for NO_2 generally seems to exert a greater effect on the point estimate than other co-pollutants (e.g., Lanzinger et al., 2016a&b; Su et al., 2015; Samoli, Andersen et al., 2016, Zhang et al., 2013). One reason for this is the overlap in sources and spatial/temporal distribution of UFP/quasi-UFP and NO_2, which can lead to instability in the models and biased effect estimates in two-exposure models.

5.3 Transferability of results to the situation in Germany

The transferability of the above reported results to the situation in Germany will be judged according to the following criteria: Localizations of identified studies and level of exposure to ultrafine particles, level of exposure to airborne co-pollutants, baseline prevalence of investigated diseases and selection of study populations.

5.3.1 Exposure

The vast majority of the identified studies are located in North America (n=37, 43.5%) or Western Europe (n=27, 31.8%) and five studies (6%) located in more than one world region (Table 2). When examining the study sites of studies with multiple study centers, we can observe that the majority of study sites are located in Western and Southern Europe (n=44 of 101 study sites, 43.6%)(Table 3). The concentrations of ultrafine particles vary considerably in time and space and direct comparisons of single center measurements are subject to large variation depending on hour, day and season of measurement as well as exact placement of the measurement site (traffic, urban background, regional background site)(Birmili et al., 2016; UFIPOLNET, 2008). In the German Ultrafine Aerosol Network (GUAN), long-term measurements of ultrafine and fine particles have been conducted at 17 sites across Germany, including alpine sites (Zugspiptze), rural sites, urban background and roadside measurement sites (Birmili et al. 2016). Of note, the size of the measured particles ranges from 20 to 800 nm, thereby not encompassing the nucleation mode of particles and including the accumulation mode particles. Preliminary results of GUAN measurements indicate a range of hourly median concentrations of particle number (sized 20-800 nm) between 900/ml (Zugspitze) and 9,000/ml at the roadside in Leipzig. Hourly mean concentrations are higher with 1,120/ml at the Zugspitze and 10,500/ml in Leipzig. The 95 percentile of the distribution of hourly values reaches 22,400/ml in Leipzig-Mitte. All three road-

side measurement sites had P95 values above 19,900/ml, while the urban background sites ranged between 10,000 and 20,000/ml. GUAN also demonstrates the substantial variation in particle size distribution during the course of a week at six mainly urban sites.

The identified studies conducted in Western Europe typically have similar or higher mean total particle counts. A direct comparison is not possible with the available information, since instruments for measurements differ and have different lower cutpoints. 16 out of the 27 studies in Western Europe report the lower cutpoint of their measurement device as 10 nm or lower. Some devices go down as far as 3 nm as their lower cutpoint. Since the majority of particles is sized below 20 nm (nucleation mode) (HEI perspectives, 2013), small differences in the lower cutpoint leads to substantial differences in mean exposures. In addition, the upper cutpoint also varies considerably, with only few studies examining ultrafine particles in the more strict sense (<100 nm), but rather use the surrogate of total particle number concentration as the exposure of interest. This, however, presents a minor problem as total particle number is dominated by the size fraction below 100 nm (HEI perspectives, 2013).

For the benefit of this review, GUAN primarily demonstrates the large variability of exposures within Germany, but it is not well suited to compare absolute values with other studies, which used different measurement devices. The five studies from Germany included in the review are based on central-site or personal measurements (n=4) with lower cutpoints ranging between 3 and 10 nm. These studies yield mean exposures between 10.000/ml and 20.000/ml, which is comparable to other studies in this review. In comparison, the 13 studies located in the Western Pacific region or in South-East-Asia, in the metropolitan areas of China, South Korea or Taiwan, report measured mean particle number concentrations in similar or slightly higher ranges. The only German study based on modelled exposures applying the EURAD CTM yielded substantially higher mean exposures due to the modelling process, which included the complete nucleation mode and therefore also encompasses short-lived particles sized below 3 nm. We therefore conclude that the level of exposure in the identified studies, while very variable across time and space, is generally comparable to the German situation.

The development of population exposure to ultrafine and quasi-ultrafine particles in Germany in the coming years depends on several factors: (1) the formation and emission of these particles, (2) the spatial distribution of the population, and (3) the concentration of fine particles in ambient air.

According to a size-resolved pan-European anthropogenic particle number inventory, the most important sources of emissions are road traffic in urban areas and alongside highly trafficked roads (Health Effects Insitute, 2013) . Traffic-related emitters of primary UFP are direct injection engines in vehicles, which have increased in number during the last decade and will probably increase further (Köllner, 2016). On the other hand, vehicles with Diesel-powered engines, which also emit particles in the ultrafine and quasi-ultrafine size range, have been equipped with particle filters. This has reduced the emission of fine particles substantially (according to EURO5a less than 5 mg/km). For UFP, the EURO5b norm for the first time sets a limit at 6×10^{11} (European Union, 2007). Overall, with increasing traffic and a rising number of city dwellers expected in the future (Vallance et al., 2010), exposure to road traffic-related UFPs is likely to increase in Germany in the next decade.

A further source of mostly ultrafine particles is aircraft traffic. Several exposure studies have documented increased UFP exposure downwind of airports around the world (Hudda et al., 2014; Keuken et al., 2015; Masiol et al., 2017; Shirmohammadi et al., 2017; Stafoggia et al., 2016). The increased short-term exposure is correlated with aircraft movements over time and

reach concentrations up to 50,000 particles/ml (Keuken et al., 2015) 7 km downwind of the airport in Amsterdam and up to 75,000 particles/ml (Hudda et al., 2014) 8 km downwind in Los Angeles. The same studies show that long-term concentrations are elevated up to 3-fold 7 km downwind with more than 200,000 exposed inhabitants close to Schipohl airport, Amsterdam (Keuken et al., 2015) and up to 4-5-fold in Los Angeles, 8-10 km downwind (Hudda et al., 2014). Similar exposure studies are ongoiong in Germany and will yield first information about the exposure of residents close to German airports. Given the increase in air travel, the exposure due to airdraft emissions is likely to play an increasing role in the future.

Moreover, the concentration of fine particles in ambient air is a determinant of UFP in a way that UFP will collide and coagulate with larger particles. A high concentration of ambient fine particles will therefore support the clearance of UFP in ambient air. With the reduction of fine particles, UFP will likely stay longer airborne than in an environment with high PM concentrations.

5.3.2 Exposure to co-pollutants

The level of airborne co-pollutants are important, as most of these co-pollutants have own effects on the outcomes of interest. 78 of the 85 identified studies (92%) assessed the level of at least one other air pollutant; however, only 34 studies adjusted for at least one co-pollutants in their analysis (see section 4.3). Assessment of and adjustment for airborne co-pollutants is therefore not conducted in a comparable way across the identified studies.

Analysis of the multi-pollutant models revealed, that $PM_{2.5}$ and NO_2 are the co-pollutants which tend to influence the UFP/quasi-UFP estimate the most. Often, but not always, does the adjustment for NO_2 lead to an attenuation of the association of UFP/quasi-UFP with the health outcome (Leitte et al. 2012; Meng et al. 2012; Stafoggia et a.. 2017; Su et al. 2015; Iskandar et al. 2012; Lanzinger et al. 2016; Rosenthal et al. 2013; Gong et al. 2014; Janssen et al. 2015; Steenhof et al. 2013). Adjustment for PM_{10} and $PM_{2.5}$ also attenuates the UFP/quasi-UFP association in several studies, but in most cases less than the NO_2 adjustment.

The level of co-pollutants, and specifically $PM_{2.5}$ and NO_2, can be compared across Europe using the "Air quality in Europe — 2017 report" by the European Environmental Agency (European Environmental Agency, 2017). According to this report, Germany ranks top among the 28 member states regarding the annual mean of NO_2 at the included monitoring sites (European Environmental Agency, 2017; Fig 6.1). Similar to UFP/quasi-UFP, the annual mean at selected monitoring sites is not able to give a comprehensive overview of the exposures of the study populations in the included studies, as NO_2 concentrations are subject to a high variability across time and space. Of the 34 studies that adjusted for co-pollutants, 15 were conducted in Western Europe. Of those, three were conducted in Germany, Augsburg, and all other studies were conducted in mostly major cities in Switzerland, the Netherlands, Sweden, Denmark and Finland with comparable traffic exposures.

We therefore conclude that the findings of an at least partial overlap of effects between UFPs and NO_2, which we observe in the Western European studies included in this review (Iskandar et al. 2012; Janssen et al. 2015; Rosenthal et al. 2013; Stafoggia et al. 2017; Steenhof et al. 2013), hold true for Germany as well.

5.3.3 Disease prevalence

The majority of the studies identified in this review is located in Western/Southern Europe and North America. The cause-specific age-adjusted death rates for all non-communicable diseases

and for respiratory diseases for 2015 are similar for the WHO Region of the Americas (including South America, which is not included in this review) and the WHO European Region (World Health Organization, 2016b). On the other hand, the annual cause-specific age-adjusted death rates for cardiovascular diseases differ, with a substantially lower age-specific death rate in the Americas (211/10,000) compared to the European Region (344/10,000). This difference is primarily due to the combination of both Americas in this statistic. Compared to other European countries and the USA included in this review, Germany has a similar distribution of causes of premature deaths as the Netherlands with ischemic heart disease, lung cancer, Alzheimer disease, cerebrovascular disease and COPD ranking 1 to 5 in both countries. This ranking is very similar in the UK, Denmark, Sweden, Spain and the USA.

Moreover, the majority of studies investigate short-term sublinical outcomes (Table 8) and of those, cardiovascular, respiratory and biomarker outcomes present the focus of the included studies (Table 9). The outcome assessment of these studies is not subject to country-specific ICD-coding conventions. Unless baseline differences in physiological markers exist between the populations included in this review and the German population, which we have evidence for, transferability on results for Germany can be inferred.

5.3.4 Study population

Most studies included in this review are based on selected study populations (n=62, 72.9%) and only 10 (11.8%), respectively 13 (15.3%) studies were deemed representative or at least somewhat representative of the general population (Table 14). The studies deemed to be completely representative of the target population are the time-series studies, which are based on general populations of the city of study. One of these time-series studies (Diaz-Robles et al. 2014) targeted selected age-groups within the general population. Of the other studies, 13 (15%) studies include at least one random sample of the source population. Almost all identified articles describe the study population well. The 10 studies investigating long-term effects are mostly analyses based on existing cohorts of several hundreds to thousands of participants, exclusively located in Western Europe or North America. Of these, six studies target the adult population of either sex or limited to one sex (Ostro et al. 2015), and four studies target children (Laurent et al. 2014, 2016a and 2016b; Sunyer et al. 2015). Among the short-term studies, the study populations are mostly highly selected small groups of either healthy (younger) adults or participants with a respiratory or cardiovascular disorder such as asthma, COPD, coronary artery disease, etc..

5.3.5 Transferability – conclusions

Based on the above descriptions of exposure level, co-pollutant exposure, baseline disease prevalence and included study populations we conclude that the overall results of this review can be transferred with the appropriate caution to the German situation.

Important limitations are (1) the paucity of studies with co-pollutant adjustment, which is specifically important because of the high NO_2 exposures in Germany, and (2) the use of highly selected groups in short-term studies, as these often do not include specifically vulnerable populations such as patients with badly controlled disease, newborns and children.

5.4 Overall conclusions

The investigation of health effects in epidemiological studies is a rapidly increasing field of research and substantial developments have been made during the last seven years, tackling two of the most urgent open questions of research: First, several studies on long-term health effects of UFPs have been conducted and published. Second, specifically the more recent studies have undertaken efforts to control for co-pollutants to identify the independent effect of UFPs.

Despite the obvious development in the field, the overall conclusions have not changed substantially over the time period investigated in this study.

First, the evidence on health effects remains inconclusive or insufficient for most of the studied outcomes. Specifically, while a number of studies have investigated mortality and emergency department/hospital admission outcomes, the relatively few studies with co-pollutant adjustment reveal mixed and, up to now, inconclusive evidence. In terms of number of studies, most evidence is available from studies investigating subclinical outcomes. Within this group of studies, cardiovascular outcomes and outcomes of pulmonary and systemic inflammation show the most consistent patterns with associations generally pointing into the direction of the adverse health outcome. Nevertheless, the evidence for independence of effects remains limited here as well, as only few studies have adjusted for co-pollutants.

Second, exposure assessment in the population remains difficult, due to the specific characteristics of UFPs. Studies using central-site exposure assessment probably miss a large part of the variability. Studies using classical spatial modeling techniques need to incorporate the very high spatial and temporal variability. Null findings or reductions in UFP/quasi-UFP effect estimates upon co-pollutant adjustment can at least in part be explained by exposure misclassification and measurement error. Exposure assessment has to devote special attention to measurement techniques, size-fractions and localisations of monitor placement. Reporting needs to be standardized to make studies more easily comparable.

Third, the independence of UFPs cannot be evaluated at the moment, due to the low number of studies with adjustment and the above mentioned limitations to exposure assessment for UFPs. A positive development is the increase in studies paying attention to this issue.

Fourth, there is still an urgent need for long-term studies on health effects of UFPs. This will require the development of modeling techniques. Furthermore, specific high-exposure situations need to be identified and described in more detail to be able to assess long-term health effects. Specifically, while near road exposures have already been recognized as important factors, airport-related exposures, which have recently been shown to be substantially above background concentrations, have not been included in health effects studies yet.

In addition to these general conclusions, we conclude that the overall results of this review can be transferred with the appropriate caution to the German situation. Important limitations are (1) the paucity of studies with co-pollutant adjustment, which is specifically important because of the high NO_2 exposures in Germany, and (2) the use of highly selected groups in short-term studies.

6 List of Annexes

- Annex I: Literature search, used databases, selection process of the references, list of studies with UFP health effects published after the search period, indicators to describe and evaluate UFP-studies
- Annex II: Tables on short- and long-term health effects in the studies with co-pollutant effect estimates, quality aspects of the studies

7 References

Aguilera, I., Dratva, J., Caviezel, S., Burdet, L., de Groot, E., Ducret-Stich, R. E., . . . Probst-Hensch, N. (2016). Particulate Matter and Subclinical Atherosclerosis: Associations between Different Particle Sizes and Sources with Carotid Intima-Media Thickness in the SAPALDIA Study. *Environmental Health Perspectives, 124*(11), 1700-1706. doi:10.1289/ehp161

Baldauf, R. W., Devlin, R. B., Gehr, P., Giannelli, R., Hassett-Sipple, B., Jung, H., . . . Walker, K. (2016). Ultrafine Particle Metrics and Research Considerations: Review of the 2015 UFP Workshop. *Int J Environ Res Public Health, 13*(11). doi:10.3390/ijerph13111054

Bartell, S. M., Longhurst, J., Tjoa, T., Sioutas, C., & Delfino, R. J. (2013). Particulate air pollution, ambulatory heart rate variability, and cardiac arrhythmia in retirement community residents with coronary artery disease. *Environ Health Perspect, 121*(10), 1135-1141. doi:10.1289/ehp.1205914

Bind, M. A., Peters, A., Koutrakis, P., Coull, B., Vokonas, P., & Schwartz, J. (2016). Quantile Regression Analysis of the Distributional Effects of Air Pollution on Blood Pressure, Heart Rate Variability, Blood Lipids, and Biomarkers of Inflammation in Elderly American Men: The Normative Aging Study. *Environ Health Perspect, 124*(8), 1189-1198. doi:10.1289/ehp.1510044

Birmili, W., Weinhold, K., Rasch, F., Sonntag, A., Sun, J., Merkel, M., . . . Fiebig, M. (2016). Long-term observations of tropospheric particle number size distributions and equivalent black carbon mass concentrations in the German Ultrafine Aerosol Network (GUAN). *Earth System Science Data, 8*(2), 355-382. doi:10.5194/essd-8-355-2016

Bos, I., De Boever, P., Vanparijs, J., Pattyn, N., Panis, L. I., & Meeusen, R. (2013). Subclinical effects of aerobic training in urban environment. *Medicine and Science in Sports and Exercise, 45*(3), 439-447. doi:10.1249/MSS.0b013e31827767fc

Bos, I., Jacobs, L., Nawrot, T. S., de Geus, B., Torfs, R., Int Panis, L., . . . Meeusen, R. (2011). No exercise-induced increase in serum BDNF after cycling near a major traffic road. *Neuroscience Letters, 500*(2), 129-132. doi:10.1016/j.neulet.2011.06.019

Cassee, F. R., Heroux, M. E., Gerlofs-Nijland, M. E., & Kelly, F. J. (2013). Particulate matter beyond mass: recent health evidence on the role of fractions, chemical constituents and sources of emission. *Inhalation Toxicology, 25*(14), 802-812. doi:10.3109/08958378.2013.850127

Chung, M., Wang, D. D., Rizzo, A. M., Gachette, D., Delnord, M., Parambi, R., . . . Brugge, D. (2015). Association of PNC, BC, and PM2.5 Measured at a Central Monitoring Site with Blood Pressure in a Predominantly Near Highway Population. *International Journal of Environmental Research and Public Health, 12*(3), 2765-2780. doi:10.3390/ijerph120302765

Cohen, A. J., Brauer, M., Burnett, R., Anderson, H. R., Frostad, J., Estep, K., . . . Forouzanfar, M. H. (2017). 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. *The Lancet, 389*(10082), 1907-1918. doi:10.1016/s0140-6736(17)30505-6

Cole-Hunter, T., Jayaratne, R., Stewart, I., Hadaway, M., Morawska, L., & Solomon, C. (2013). Utility of an alternative bicycle commute route of lower proximity to motorised traffic in decreasing exposure to ultra-fine particles, respiratory symptoms and airway inflammation - a structured exposure experiment. *Environmental Health, 12*. doi:10.1186/1476-069x-12-29

Cole-Hunter, T., Weichenthal, S., Kubesch, N., Foraster, M., Carrasco-Turigas, G., Bouso, L., . . . Nieuwenhuijsen, M. (2016). Impact of traffic-related air pollution on acute changes in cardiac autonomic modulation during rest and physical activity: a cross-over study. *Journal of Exposure Science & Environmental Epidemiology, 26*(2), 133-140. doi:10.1038/jes.2015.66

Croft, D. P., Cameron, S. J., Morrell, C. N., Lowenstein, C. J., Ling, F., Zareba, W., . . . Rich, D. Q. (2017). Associations between ambient wood smoke and other particulate pollutants and biomarkers of systemic inflammation, coagulation and thrombosis in cardiac patients. *Environ Res, 154*, 352-361. doi:10.1016/j.envres.2017.01.027

Delfino, R. J., Wu, J., Tjoa, T., Gullesserian, S. K., Nickerson, B., & Gillen, D. L. (2014). Asthma morbidity and ambient air pollution: effect modification by residential traffic-related air pollution. *Epidemiology, 25*(1), 48-57. doi:10.1097/EDE.0000000000000016

Deutscher Wetterdienst. (2018). Particle Size-Distribution Retrieved from https://www.dwd.de/EN/research/observing_atmosphere/composition_atmosphere/aerosol/cont_nav/particle_size_distribution_node.html

Diaz-Robles, L. A., Fu, J. S., Vergara-Fernandez, A., Etcharren, P., Schiappacasse, L. N., Reed, G. D., & Silva, M. P. (2014). Health risks caused by short term exposure to ultrafine particles generated by residential wood combustion: A case study of Temuco,Chile. *Environment International, 66*, 174-181. doi:10.1016/j.envint.2014.01.017

Dionisio, K. L., Baxter, L. K., & Chang, H. H. (2014). An empirical assessment of exposure measurement error and effect attenuation in bipollutant epidemiologic models. *Environ Health Perspect, 122*(11), 1216-1224. doi:10.1289/ehp.1307772

European Environmental Agency. (2017). Air quality in Europe — 2017 report. *EEA Report No 13/2017*. doi:10.2800/358908

European Union. (2007). *Verordnung (EG) Nr. 715/2007 des europäischen Parlaments und des Rates vom 20. Juni 2007 über die Typgenehmigung von Kraftfahrzeugen hinsichtlich der Emissionen von leichten Personenkraftwagen und*

Nutzfahrzeugen (Euro 5 und Euro 6) und über den Zugang zu Reparatur- und Wartungsinformationen für Fahrzeuge. Amtsblatt der Europäischen Union Retrieved from http://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=uriserv:OJ.L_.2007.171.01.0001.01.DEU.

Evans, K. A., Halterman, J. S., Hopke, P. K., Fagnano, M., & Rich, D. Q. (2014). Increased ultrafine particles and carbon monoxide concentrations are associated with asthma exacerbation among urban children. *Environmental Research, 129*, 11-19. doi:10.1016/j.envres.2013.12.001

Fanning, E. W., Froines, J. R., Utell, M. J., Lippmann, M., Oberdorster, G., Frampton, M., . . . Larson, T. V. (2009). Particulate matter (PM) research centers (1999-2005) and the role of interdisciplinary center-based research. *Environ Health Perspect, 117*(2), 167-174. doi:10.1289/ehp.11543

Forouzanfar, M. H., Afshin, A., Alexander, L. T., Anderson, H. R., Bhutta, Z. A., Biryukov, S., . . . Murray, C. J. L. (2016). Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet, 388*(10053), 1659-1724. doi:10.1016/s0140-6736(16)31679-8

Fuller, C. H., Williams, P. L., Mittleman, M. A., Patton, A. P., Spengler, J. D., & Brugge, D. (2015). Response of biomarkers of inflammation and coagulation to short-term changes in central site, local, and predicted particle number concentrations. *Annals of Epidemiology, 25*(7), 505-511. doi:10.1016/j.annepidem.2015.02.003

Gardner, B., Ling, F., Hopke, P. K., Frampton, M. W., Utell, M. J., Zareba, W., . . . Rich, D. Q. (2014). Ambient fine particulate air pollution triggers ST-elevation myocardial infarction, but not non-ST elevation myocardial infarction: a case-crossover study. *Particle and Fibre Toxicology, 11*. doi:10.1186/1743-8977-11-1

Gong, J., Zhu, T., Kipen, H., Wang, G., Hu, M., Guo, Q., . . . Zhang, J. (2014). Comparisons of ultrafine and fine particles in their associations with biomarkers reflecting physiological pathways. *Environmental Science & Technology, 48*(9), 5264-5273. doi:10.1021/es5006016

Halonen, J. I., Lanki, T., Yli-Tuomi, T., Kulmala, M., Tiittanen, P., & Pekkanen, J. (2008). Urban air pollution, and asthma and COPD hospital emergency room visits. *Thorax, 63*(7), 635-641. doi:10.1136/thx.2007.091371

Hampel, R., Breitner, S., Schneider, A., Zareba, W., Kraus, U., Cyrys, J., . . . Cooperative Health Research in the Region of Augsburg Study, G. (2012). Acute air pollution effects on heart rate variability are modified by SNPs involved in cardiac rhythm in individuals with diabetes or impaired glucose tolerance. *Environ Res, 112*, 177-185. doi:10.1016/j.envres.2011.10.007

Hampel, R., Ruckerl, R., Yli-Tuomi, T., Breitner, S., Lanki, T., Kraus, U., . . . Schneider, A. (2014). Impact of personally measured pollutants on cardiac function. *International Journal of Hygiene and Environmental Health, 217*(4-5), 460-464. doi:10.1016/j.ijheh.2013.09.002

Han, Y. Q., Zhu, T., Guan, T. J., Zhu, Y., Liu, J., Ji, Y. F., . . . Huang, W. (2016). Association between size-segregated particles in ambient air and acute respiratory inflammation. *Science of the Total Environment, 565*, 412-419. doi:10.1016/j.scitotenv.2016.04.196

Health Effects Insitute. (2013). *Understanding the Health Effects of Ambient Ultrafine Particles. HEI Perspectives 3*: Health Effects Institute, Boston, MA.

Henschel, S., & Chen, G. (2013). Health risks of air pollution in Europe – HRAPIE project New emerging risks to health from air pollution – results from the survey of experts. *World Health Organization, Regional Office for Europe*.

Hoffmann, B., Luttmann-Gibson, H., Cohen, A., Zanobetti, A., de Souza, C., Foley, C., . . . Gold, D. R. (2012). Opposing effects of particle pollution, ozone, and ambient temperature on arterial blood pressure. *Environ Health Perspect, 120*(2), 241-246. doi:10.1289/ehp.1103647

Hudda, N., & Fruin, S. A. (2016). International Airport Impacts to Air Quality: Size and Related Properties of Large Increases in Ultrafine Particle Number Concentrations. *Environmental Science & Technology, 50*(7), 3362-3370. doi:10.1021/acs.est.5b05313

Hudda, N., Gould, T., Hartin, K., Larson, T. V., & Fruin, S. A. (2014). Emissions from an international airport increase particle number concentrations 4-fold at 10 km downwind. *Environmental Science & Technology, 48*(12), 6628-6635. doi:10.1021/es5001566

Hudda, N., Simon, M. C., Zamore, W., Brugge, D., & Durant, J. L. (2016). Aviation Emissions Impact Ambient Ultrafine Particle Concentrations in the Greater Boston Area. *Environmental Science & Technology, 50*(16), 8514-8521. doi:10.1021/acs.est.6b01815

Huttunen, K., Siponen, T., Salonen, I., Yli-Tuomi, T., Aurela, M., Dufva, H., . . . Lanki, T. (2012). Low-level exposure to ambient particulate matter is associated with systemic inflammation in ischemic heart disease patients. *Environmental Research, 116*, 44-51. doi:10.1016/j.envres.2012.04.004

Iskandar, A., Andersen, Z. J., Bonnelykke, K., Ellermann, T., Andersen, K. K., & Bisgaard, H. (2012). Coarse and fine particles but not ultrafine particles in urban air trigger hospital admission for asthma in children. *Thorax, 67*(3), 252-257. doi:10.1136/thoraxjnl-2011-200324

Janssen, N. A., Strak, M., Yang, A., Hellack, B., Kelly, F. J., Kuhlbusch, T. A., . . . Hoek, G. (2015). Associations between three specific a-cellular measures of the oxidative potential of particulate matter and markers of acute airway and nasal inflammation in healthy volunteers. *Occup Environ Med, 72*(1), 49-56. doi:10.1136/oemed-2014-102303

Jarjour, S., Jerrett, M., Westerdahl, D., de Nazelle, A., Hanning, C., Daly, L., . . . Balmes, J. (2013). Cyclist route choice, traffic-related air pollution, and lung function: a scripted exposure study. *Environmental Health: A Global Access Science Source, 12*, 14. doi:10.1186/1476-069x-12-14

Karakatsani, A., Analitis, A., Perifanou, D., Ayres, J. G., Harrison, R. M., Kotronarou, A., . . . Katsouyanni, K. (2012). Particulate matter air pollution and respiratory symptoms in individuals having either asthma or chronic obstructive pulmonary disease: a European multicentre panel study. *Environmental Health: A Global Access Science Source, 11*, 75. doi:10.1186/1476-069X-11-75

Karottki, D. G., Beko, G., Clausen, G., Madsen, A. M., Andersen, Z. J., Massling, A., . . . Loft, S. (2014). Cardiovascular and lung function in relation to outdoor and indoor exposure to fine and ultrafine particulate matter in middle-aged subjects. *Environment International, 73*, 372-381. doi:10.1016/j.envint.2014.08.019

Karottki, D. G., Spilak, M., Frederiksen, M., Jovanovic Andersen, Z., Madsen, A. M., Ketzel, M., . . . Loft, S. (2015). Indoor and outdoor exposure to ultrafine, fine and microbiologically derived particulate matter related to cardiovascular and respiratory effects in a panel of elderly urban citizens. *Int J Environ Res Public Health, 12*(2), 1667-1686. doi:10.3390/ijerph120201667

Kelly, F. J., & Fussell, J. C. (2012). Size, source and chemical composition as determinants of toxicity attributable to ambient particulate matter. *Atmospheric Environment, 60*, 504-526. doi:10.1016/j.atmosenv.2012.06.039

Keuken, M. P., Moerman, M., Zandveld, P., Henzing, J. S., & Hoek, G. (2015). Total and size-resolved particle number and black carbon concentrations in urban areas near Schiphol airport (the Netherlands). *Atmospheric Environment, 104*, 132-142. doi:10.1016/j.atmosenv.2015.01.015

Köllner, C. (2016, 05.12.2016). Kurze Geschichte der Benzindirekteinspritzung. *Springer Professional.* Retrieved from https://www.springerprofessional.de/einspritzung/ottomotor/kurze-geschichte-der-benzindirekteinspritzung/11104988

Kubesch, N., De Nazelle, A., Guerra, S., Westerdahl, D., Martinez, D., Bouso, L., . . . Nieuwenhuijsen, M. J. (2015). Arterial blood pressure responses to short-term exposure to low and high traffic-related air pollution with and without moderate physical activity. *Eur J Prev Cardiol, 22*(5), 548-557. doi:10.1177/2047487314555602

Lane, K. J., Levy, J. I., Scammell, M. K., Patton, A. P., Durant, J. L., Mwamburi, M., . . . Brugge, D. (2015). Effect of time-activity adjustment on exposure assessment for traffic-related ultrafine particles. *Journal of Exposure Science and Environmental Epidemiology, 25*(5), 506-516. doi:10.1038/jes.2015.11

Lane, K. J., Levy, J. I., Scammell, M. K., Peters, J. L., Patton, A. P., Reisner, E., . . . Brugge, D. (2016). Association of modeled long-term personal exposure to ultrafine particles with inflammatory and coagulation biomarkers. *Environ Int, 92-93*, 173-182. doi:10.1016/j.envint.2016.03.013

Langrish, J. P., Li, X., Wang, S., Lee, M. M., Barnes, G. D., Miller, M. R., . . . Jiang, L. (2012). Reducing personal exposure to particulate air pollution improves cardiovascular health in patients with coronary heart disease. *Environ Health Perspect, 120*(3), 367-372. doi:10.1289/ehp.1103898

Lanzinger, S., Schneider, A., Breitner, S., Stafoggia, M., Erzen, I., Dostal, M., . . . Peters, A. (2016a). Associations between ultrafine and fine particles and mortality in five central European cities - Results from the UFIREG study. *Environ Int, 88*, 44-52. doi:10.1016/j.envint.2015.12.006

Lanzinger, S., Schneider, A., Breitner, S., Stafoggia, M., Erzen, I., Dostal, M., . . . Peters, A. (2016b). Ultrafine and Fine Particles and Hospital Admissions in Central Europe. Results from the UFIREG Study. *American Journal of Respiratory and Critical Care Medicine, 194*(10), 1233-1241. doi:10.1164/rccm.201510-2042OC

Laumbach, R. J., Kipen, H. M., Ko, S., Kelly-McNeil, K., Cepeda, C., Pettit, A., . . . Gow, A. J. (2014). A controlled trial of acute effects of human exposure to traffic particles on pulmonary oxidative stress and heart rate variability. *Particle and Fibre Toxicology, 11*, 45. doi:10.1186/s12989-014-0045-5

Laurent, O., Hu, J., Li, L., Cockburn, M., Escobedo, L., Kleeman, M. J., & Wu, J. (2014). Sources and contents of air pollution affecting term low birth weight in Los Angeles County, California, 2001-2008. *Environ Res, 134*, 488-495. doi:10.1016/j.envres.2014.05.003

Laurent, O., Hu, J., Li, L., Kleeman, M. J., Bartell, S. M., Cockburn, M., . . . Wu, J. (2016a). Low birth weight and air pollution in California: Which sources and components drive the risk? *Environ Int, 92-93*, 471-477. doi:10.1016/j.envint.2016.04.034

Laurent, O., Hu, J., Li, L., Kleeman, M. J., Bartell, S. M., Cockburn, M., . . . Wu, J. (2016b). A Statewide Nested Case-Control Study of Preterm Birth and Air Pollution by Source and Composition: California, 2001-2008. *Environ Health Perspect, 124*(9), 1479-1486. doi:10.1289/ehp.1510133

Leitte, A. M., Schlink, U., Herbarth, O., Wiedensohler, A., Pan, X. C., Hu, M., . . . Franck, U. (2012). Associations between size-segregated particle number concentrations and respiratory mortality in Beijing, China. *International Journal of Environmental Health Research, 22*(2), 119-133. doi:10.1080/09603123.2011.605878

Li, N., Georas, S., Alexis, N., Fritz, P., Xia, T., Williams, M. A., . . . Nel, A. (2016). A work group report on ultrafine particles (American Academy of Allergy, Asthma & Immunology): Why ambient ultrafine and engineered nanoparticles should receive special attention for possible adverse health outcomes in human subjects. *Journal of Allergy and Clinical Immunology, 138*(2), 386-396. doi:10.1016/j.jaci.2016.02.023

Li, Y., Lane, K. J., Corlin, L., Patton, A. P., Durant, J. L., Thanikachalam, M., . . . Brugge, D. (2017). Association of Long-Term Near-Highway Exposure to Ultrafine Particles with Cardiovascular Diseases, Diabetes and Hypertension. *Int J Environ Res Public Health, 14*(5). doi:10.3390/ijerph14050461

Li, Y. R., Feng, L. T., Chen, B. Y., Kim, H., Yi, S. M., Guo, Y. L., & Wu, C. F. (2016). Association of urban particle numbers and sources with lung function among children with asthma or allergies. *Science of the Total Environment, 542*, 841-844. doi:10.1016/j.scitotenv.2015.10.098

Link, M. S., Luttmann-Gibson, H., Schwartz, J., Mittleman, M. A., Wessler, B., Gold, D. R., . . . Laden, F. (2013). Acute Exposure to Air Pollution Triggers Atrial Fibrillation. *Journal of the American College of Cardiology, 62*(9), 816-825. doi:10.1016/j.jacc.2013.05.043

Liu, L., Breitner, S., Schneider, A., Cyrys, J., Bruske, I., Franck, U., . . . Peters, A. (2013). Size-fractioned particulate air pollution and cardiovascular emergency room visits in Beijing, China. *Environ Res, 121*, 52-63. doi:10.1016/j.envres.2012.10.009

Ljungman, P. L., Wilker, E. H., Rice, M. B., Schwartz, J., Gold, D. R., Koutrakis, P., . . . Hamburg, N. M. (2014). Short-term exposure to air pollution and digital vascular function. *American Journal of Epidemiology, 180*(5), 482-489. doi:10.1093/aje/kwu161

Manney, S., Meddings, C. M., Harrison, R. M., Mansur, A. H., Karakatsani, A., Analitis, A., . . . Ayres, J. G. (2012). Association between exhaled breath condensate nitrate plus nitrite levels with ambient coarse particle exposure in subjects with airways disease. *Occupational and Environmental Medicine, 69*(9), 663-669. doi:10.1136/oemed-2011-100255

Masiol, M., Hopke, P. K., Felton, H. D., Frank, B. P., Rattigan, O. V., Wurth, M. J., & LaDuke, G. H. (2017). Analysis of major air pollutants and submicron particles in New York City and Long Island. *Atmospheric Environment, 148*, 203-214. doi:https://doi.org/10.1016/j.atmosenv.2016.10.043

Mehta, A. J., Kubzansky, L. D., Coull, B. A., Kloog, I., Koutrakis, P., Sparrow, D., . . . Schwartz, J. (2015). Associations between air pollution and perceived stress: the Veterans Administration Normative Aging Study. *Environmental Health, 14*. doi:10.1186/1476-069x-14-10

Mehta, A. J., Zanobetti, A., Koutrakis, P., Mittleman, M. A., Sparrow, D., Vokonas, P., & Schwartz, J. (2014). Associations Between Short-term Changes in Air Pollution and Correlates of Arterial Stiffness: The Veterans Affairs Normative Aging Study, 2007–2011. *American Journal of Epidemiology, 179*(2), 192-199. doi:10.1093/aje/kwt271

Meng, X., Ma, Y. J., Chen, R. J., Zhou, Z. J., Chen, B. H., & Kan, H. D. (2013). Size-Fractionated Particle Number Concentrations and Daily Mortality in a Chinese City. *Environmental Health Perspectives, 121*(10), 1174-1178. doi:10.1289/ehp.1206398

Mirabelli, M. C., Golan, R., Greenwald, R., Raysoni, A. U., Holguin, F., Kewada, P., . . . Sarnat, J. A. (2015). Modification of Traffic-related Respiratory Response by Asthma Control in a Population of Car Commuters. *Epidemiology, 26*(4), 546-555. doi:10.1097/EDE.0000000000000296

National Heart, L. a. B. I., , . (2014). Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies, US-Department of Health and Human Services, . Retrieved from http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp

Ostro, B., Hu, J., Goldberg, D., Reynolds, P., Hertz, A., Bernstein, L., & Kleeman, M. J. (2015). Associations of mortality with long-term exposures to fine and ultrafine particles, species and sources: results from the California Teachers Study Cohort. *Environ Health Perspect, 123*(6), 549-556. doi:10.1289/ehp.1408565

Park, H. Y., Gilbreath, S., & Barakatt, E. (2017). Respiratory outcomes of ultrafine particulate matter (UFPM) as a surrogate measure of near-roadway exposures among bicyclists. *Environmental Health, 16*. doi:10.1186/s12940-017-0212-x

Peng, C., Luttmann-Gibson, H., Zanobetti, A., Cohen, A., De Souza, C., Coull, B. A., . . . Gold, D. R. (2016). Air pollution influences on exhaled nitric oxide among people with type II diabetes. *Air Quality Atmosphere and Health, 9*(3), 265-273. doi:10.1007/s11869-015-0336-5

Peters, A., Hampel, R., Cyrys, J., Breitner, S., Geruschkat, U., Kraus, U., . . . Schneider, A. (2015). Elevated particle number concentrations induce immediate changes in heart rate variability: a panel study in individuals with impaired glucose metabolism or diabetes. *Particle and Fibre Toxicology, 12*. doi:10.1186/s12989-015-0083-7

Pieters, N., Koppen, G., Van Poppel, M., De Prins, S., Cox, B., Dons, E., . . . Nawrot, T. S. (2015). Blood Pressure and Same-Day Exposure to Air Pollution at School: Associations with Nano-Sized to Coarse PM in Children. *Environmental Health Perspectives, 123*(7), 737-742. doi:10.1289/ehp.1408121

Rich, D. Q., Zareba, W., Beckett, W., Hopke, P. K., Oakes, D., Frampton, M. W., . . . Utell, M. J. (2012). Are ambient ultrafine, accumulation mode, and fine particles associated with adverse cardiac responses in patients undergoing cardiac rehabilitation? *Environ Health Perspect, 120*(8), 1162-1169. doi:10.1289/ehp.1104262

Rosenthal, F. S., Kuisma, M., Lanki, T., Hussein, T., Boyd, J., Halonen, J. I., & Pekkanen, J. (2013). Association of ozone and particulate air pollution with out-of-hospital cardiac arrest in Helsinki, Finland: Evidence for two different etiologies. *Journal of Exposure Science and Environmental Epidemiology, 23*(3), 281-288. doi:10.1038/jes.2012.121

Rückerl, R., Hampel, R., Breitner, S., Cyrys, J., Kraus, U., Carter, J., . . . Schneider, A. (2014). Associations between ambient air pollution and blood markers of inflammation and coagulation/fibrinolysis in susceptible populations. *Environ Int, 70*, 32-49. doi:10.1016/j.envint.2014.05.013

Rückerl, R., Schneider, A., Breitner, S., Cyrys, J., & Peters, A. (2011). Health effects of particulate air pollution: A review of epidemiological evidence. *Inhalation Toxicology, 23*(10), 555-592. doi:10.3109/08958378.2011.593587

Rückerl, R., Schneider, A., Hampel, R., Breitner, S., Cyrys, J., Kraus, U., . . . Peters, A. (2016). Association of novel metrics of particulate matter with vascular markers of inflammation and coagulation in susceptible populations -results from a panel study. *Environ Res, 150*, 337-347. doi:10.1016/j.envres.2016.05.037

Samoli, E., Andersen, Z. J., Katsouyanni, K., Hennig, F., Kuhlbusch, T. A. J., Bellander, T., . . . Grp, U. H. S. (2016). Exposure to ultrafine particles and respiratory hospitalisations in five European cities. *European Respiratory Journal, 48*(3), 674-682. doi:10.1183/13993003.02108-2015

Samoli, E., Atkinson, R. W., Analitis, A., Fuller, G. W., Beddows, D., Green, D. C., . . . Kelly, F. J. (2016). Differential health effects of short-term exposure to source-specific particles in London, UK. *Environment International, 97*, 246-253. doi:10.1016/j.envint.2016.09.017

Sarnat, J. A., Golan, R., Greenwald, R., Raysoni, A. U., Kewada, P., Winquist, A., . . . Yip, F. (2014). Exposure to traffic pollution, acute inflammation and autonomic response in a panel of car commuters. *Environ Res, 133*, 66-76. doi:10.1016/j.envres.2014.05.004

Shirmohammadi, F., Sowlat, M. H., Hasheminassab, S., Saffari, A., Ban-Weiss, G., & Sioutas, C. (2017). Emission rates of particle number, mass and black carbon by the Los Angeles International Airport (LAX) and its impact on air quality in Los Angeles. *Atmospheric Environment, 151*, 82-93. doi:https://doi.org/10.1016/j.atmosenv.2016.12.005

Shutt, R. H., Kauri, L. M., Weichenthal, S., Kumarathasan, P., Vincent, R., Thomson, E. M., . . . Dales, R. (2017). Exposure to air pollution near a steel plant is associated with reduced heart rate variability: a randomised crossover study. *Environmental Health, 16*. doi:10.1186/s12940-016-0206-0

Song, S., Lee, K., Lee, Y. M., Lee, J. H., Lee, S. I., Yu, S. D., & Paek, D. (2011). Acute health effects of urban fine and ultrafine particles on children with atopic dermatitis. *Environ Res, 111*(3), 394-399. doi:10.1016/j.envres.2010.10.010

Song, S., Paek, D., Lee, K., Lee, Y. M., Lee, C., Park, C., & Yu, S. D. (2013). Effects of Ambient Fine Particles on Pulmonary Function in Children With Mild Atopic Dermatitis. *Archives of Environmental & Occupational Health, 68*(4), 228-234. doi:10.1080/19338244.2012.701247

Stafoggia, M., Cattani, G., Forastiere, F., Di Menno di Bucchianico, A., Gaeta, A., & Ancona, C. (2016). Particle number concentrations near the Rome-Ciampino city airport. *Atmospheric Environment, 147*, 264-273. doi:https://doi.org/10.1016/j.atmosenv.2016.09.062

Stafoggia, M., Schneider, A., Cyrys, J., Samoli, E., Andersen, Z. J., Bedada, G. B., . . . Grp, H. S. (2017). Association Between Short-term Exposure to Ultrafine Particles and Mortality in Eight European Urban Areas. *Epidemiology, 28*(2), 172-180. doi:10.1097/ede.0000000000000599

Steenhof, M., Janssen, N. A., Strak, M., Hoek, G., Gosens, I., Mudway, I. S., . . . Brunekreef, B. (2014). Air pollution exposure affects circulating white blood cell counts in healthy subjects: the role of particle composition, oxidative potential and gaseous pollutants - the RAPTES project. *Inhalation Toxicology, 26*(3), 141-165. doi:10.3109/08958378.2013.861884

Steenhof, M., Mudway, I. S., Gosens, I., Hoek, G., Godri, K. J., Kelly, F. J., . . . Janssen, N. A. (2013). Acute nasal pro-inflammatory response to air pollution depends on characteristics other than particle mass concentration or oxidative potential: the RAPTES project. *Occup Environ Med, 70*(5), 341-348. doi:10.1136/oemed-2012-100993

Stone, V., Miller, M. R., Clift, M. J., Elder, A., Mills, N. L., Moller, P., . . . Cassee, F. R. (2016). Nanomaterials vs Ambient Ultrafine Particles: an Opportunity to Exchange Toxicology Knowledge. *Environ Health Perspect*. doi:10.1289/ehp424

Strak, M., Hoek, G., Godri, K. J., Gosens, I., Mudway, I. S., van Oerle, R., . . . Janssen, N. A. (2013). Composition of PM affects acute vascular inflammatory and coagulative markers - the RAPTES project. *PLoS One, 8*(3), e58944. doi:10.1371/journal.pone.0058944

Strak, M., Janssen, N. A., Godri, K. J., Gosens, I., Mudway, I. S., Cassee, F. R., . . . Hoek, G. (2012). Respiratory health effects of airborne particulate matter: the role of particle size, composition, and oxidative potential-the RAPTES project. *Environ Health Perspect, 120*(8), 1183-1189. doi:10.1289/ehp.1104389

Su, C., Hampel, R., Franck, U., Wiedensohler, A., Cyrys, J., Pan, X. C., . . . Breitner, S. (2015). Assessing responses of cardiovascular mortality to particulate matter air pollution for pre-, during- and post-2008 Olympics periods. *Environmental Research, 142*, 112-122. doi:10.1016/j.envres.2015.06.025

Sun, Y., Song, X., Han, Y., Ji, Y., Gao, S., Shang, Y., . . . Huang, W. (2015). Size-fractioned ultrafine particles and black carbon associated with autonomic dysfunction in subjects with diabetes or impaired glucose tolerance in Shanghai, China. *Particle and Fibre Toxicology, 12*, 8. doi:10.1186/s12989-015-0084-6

Sunyer, J., Esnaola, M., Alvarez-Pedrerol, M., Forns, J., Rivas, I., Lopez-Vicente, M., . . . Querol, X. (2015). Association between Traffic-Related Air Pollution in Schools and Cognitive Development in Primary School Children: A Prospective Cohort Study. *PLoS Medicine, 12*(3). doi:10.1371/journal.pmed.1001792

Terzano, C., Di Stefano, F., Conti, V., Graziani, E., & Petroianni, A. (2010). Air pollution ultrafine particles: toxicity beyond the lung. *European Review for Medical and Pharmacological Sciences, 14*(10), 809-821.

Thurston, G. D., Kipen, H., Annesi-Maesano, I., Balmes, J., Brook, R. D., Cromar, K., . . . Brunekreef, B. (2017). A joint ERS/ATS policy statement: what constitutes an adverse health effect of air pollution? An analytical framework. *European Respiratory Journal, 49*(1). doi:10.1183/13993003.00419-2016

U.S. Environmental Protection Agency. (2009). Integrated Science Assessment for Particulate Matter. *Final Report. EPA/600/R-08/139F. National Center for Environmental Assessment-RTP Division, Office of Research and Development, Research Triangle Park, NC*.

UFIPOLNET. (2008). UFIPOLNET Technical Final Report. *Saxon State Agency for Environment and Geology, 01-12-2004 to 31-03-2008 LIFE04 ENV/D/000054*.

Vallance, S., & Perkins, H. (2010). Is another city possible? Towards an urbanised sustainability. *City, 14*(4), 448-456. doi:10.1080/13604813.2010.496217

Viehmann, A., Hertel, S., Fuks, K., Eisele, L., Moebus, S., Mohlenkamp, S., . . . Hoffmann, B. (2015). Long-term residential exposure to urban air pollution, and repeated measures of systemic blood markers of inflammation and coagulation. *Occup Environ Med, 72*(9), 656-663. doi:10.1136/oemed-2014-102800

Wang, M., Utell, M. J., Schneider, A., Zareba, W., Frampton, M. W., Oakes, D., . . . Rich, D. Q. (2016). Does total antioxidant capacity modify adverse cardiac responses associated with ambient ultrafine, accumulation mode, and fine particles in patients undergoing cardiac rehabilitation? *Environ Res, 149*, 15-22. doi:10.1016/j.envres.2016.04.031

Wang, Y., Eliot, M. N., Koutrakis, P., Gryparis, A., Schwartz, J. D., Coull, B. A., . . . Wellenius, G. A. (2014). Ambient Air Pollution and Depressive Symptoms in Older Adults: Results from the MOBILIZE Boston Study. *Environmental Health Perspectives, 122*(6), 553-558. doi:10.1289/ehp.1205909

Weichenthal, S., Hatzopoulou, M., & Goldberg, M. S. (2014). Exposure to traffic-related air pollution during physical activity and acute changes in blood pressure, autonomic and micro-vascular function in women: a cross-over study. *Particle and Fibre Toxicology, 11*. doi:10.1186/s12989-014-0070-4

Wichmann, J., Folke, F., Torp-Pedersen, C., Lippert, F., Ketzel, M., Ellermann, T., & Loft, S. (2013). Out-of-hospital cardiac arrests and outdoor air pollution exposure in Copenhagen, Denmark. *PLoS One, 8*(1), e53684. doi:10.1371/journal.pone.0053684

Wittkopp, S., Staimer, N., Tjoa, T., Gillen, D., Daher, N., Shafer, M., . . . Delfino, R. J. (2013). Mitochondrial genetic background modifies the relationship between traffic-related air pollution exposure and systemic biomarkers of inflammation. *PLoS One, 8*(5), e64444. doi:10.1371/journal.pone.0064444

Wolf, K., Schneider, A., Breitner, S., Meisinger, C., Heier, M., Cyrys, J., . . . Peters, A. (2015). Associations between short-term exposure to particulate matter and ultrafine particles and myocardial infarction in Augsburg, Germany. *Int J Hyg Environ Health, 218*(6), 535-542. doi:10.1016/j.ijheh.2015.05.002

World Health Organization. (2013). Review of evidence on health aspects of air pollution – REVIHAAP Project. Technical report.

World Health Organization. (2016a). *Ambient air pollution: a global assessment of exposure and burden of disease* (ISBN: 978 92 4 151135 3). Retrieved from http://apps.who.int/iris/bitstream/handle/10665/250141/9789241511353-eng.pdf?sequence=1

World Health Organization. (2016b). Global Health Estimates 2015: Deaths by Cause, Age, Sex, by Country and by Region, 2000-2015.

Wu, C. F., Li, Y. R., Kuo, I. C., Hsu, S. C., Lin, L. Y., & Su, T. C. (2012). Investigating the association of cardiovascular effects with personal exposure to particle components and sources. *Science of the Total Environment, 431*, 176-182. doi:10.1016/j.scitotenv.2012.05.015

Zanobetti, A., Luttmann-Gibson, H., Horton, E. S., Cohen, A., Coull, B. A., Hoffmann, B., . . . Gold, D. R. (2014). Brachial artery responses to ambient pollution, temperature, and humidity in people with type 2 diabetes: a repeated-measures study. *Environ Health Perspect, 122*(3), 242-248. doi:10.1289/ehp.1206136

Zhang, J., Zhu, T., Kipen, H., Wang, G., Huang, W., Rich, D., . . . Thomas, D. (2013). Cardiorespiratory biomarker responses in healthy young adults to drastic air quality changes surrounding the 2008 Beijing Olympics. *Research Report of the Health Effects Institute*(174), 5-174.

Zhang, X., Staimer, N., Gillen, D. L., Tjoa, T., Schauer, J. J., Shafer, M. M., . . . Delfino, R. J. (2016). Associations of oxidative stress and inflammatory biomarkers with chemically-characterized air pollutant exposures in an elderly cohort. *Environ Res, 150*, 306-319. doi:10.1016/j.envres.2016.06.019

Zhang, X., Staimer, N., Tjoa, T., Gillen, D. L., Schauer, J. J., Shafer, M. M., . . . Delfino, R. J. (2016). Associations between microvascular function and short-term exposure to traffic-related air pollution and particulate matter oxidative potential. *Environmental Health: A Global Access Science Source, 15*(1), 81. doi:10.1186/s12940-016-0157-5

8 Annex I

8.1 HEI search strategy

Our search strategy is based on the previous systematic search by the HEI (2013), which was conducted on 09.05.2011 in Web of Science and MEDLINE via Pubmed. The search was complemented extensively through hand researches in previous reviews (e.g. US PM ISA 2009).

- Search in the Web of Science on 09.05.2011
    - 966 references identified by the following search strategy:
        - Databases=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CCR-EXPANDED, IC Timespan=1900-2011
        - # 12 966 #11 AND #10
        - # 11 >100,000 #8 OR #7
        - # 10 2,996 #9 AND #6
        - # 9 >100,000 #5 OR #4 OR #3 OR #2 OR #1
        - # 8 >100,000 Topic=(epidemiology)
        - # 7 >100,000 Topic=(health)
        - # 6 40,369 Topic=(air pollution)
        - # 5 >100,000 Topic=(surface area)
        - # 4 17,045 Topic=(number count)
        - # 3 76,211 Topic=(number concentration)
        - # 2 1,398 Topic=(particle count)
        - # 1 14,632 Topic=(ultrafine)
        - [search was rerun on 9/19/2011 – search step #12 found 1475 refs, but only 2 additional pubs from summer 2011 looked relevant for epi section]
    - 779 unique refs added (records 1673-2638)
- Searched PubMed on 5/9/2011
    - 926 refs found using the same search structure as for Web of Science:
        - #12 Search #11 AND #10 926
        - #11 Search #8 OR #7 3086579
        - #10 Search #9 AND #6 1906
        - #9 Search #1 OR #2 OR #3 OR #4 OR #5 212253
        - #8 Search epidemiology 1366933
        - #7 Search health 2111539
        - #6 Search air pollution 44525
        - #5 Search surface area 86295
        - #4 Search number count 45896
        - #3 Search number concentration 84177
        - #2 Search particle count 1472
        - #1 Search ultrafine 1687
    - 695 unique refs added (records 2639-3564)

- Hand searches (including PM Integrated Science Assessment):

- 417 references (originating from Particulate Matter – Integrated Science Assessment) (PM ISA) 2009: 281)

Exclusion criteria

- Articles focused on nanotechnology and workplace engineered NP exposure
- Indoor allergen papers
- In vivo and in vitro and human controlled exposure articles
- Articles with no particle count or size measurements (e.g., studies of traffic using only distance to roadway measures)
- Excluded articles where smallest size fraction examined was PM1 (e.g., Slaughter 2005)

8.2 LUDOK search strategy

Aufnahmekriterien sind u. a.: Epidemiologische und experimentelle Originalarbeiten über die Auswirkungen der „klassischen“ Aussenluftschadstoffe auf Menschen, sowie von weiteren Schadstoffen, die via Luft auf die Allgemeinbevölkerung einwirken (d. h. keine alleinig arbeitsmedizinisch relevanten Stoffe), inkl. Metaanalysen und methodische Arbeiten zu diesem Zusammenhang.

Suchstrategien in LUDOK:

- Sprache: en, fr, de, it (für dieses Projekt wurden nur Artikel in Deutsch und Englisch genutzt)
- Zeitraum: seit Beginn der Lufthygieneforschung bis heute (ältester Artikel von 1929, ca. 20 Artikel aus der Zeit vor 1971)
- Handsuche in relevanten Fachzeitschriften und allgemein wichtigen Journals über wöchentliche Alerts (s. unten)
- Datenbanken:
    - PubMed: Dauerrecherche mit gleich bleibender, sehr breiter Formulierung (monatlich)
        - Suchtermini: “Air Pollutants/adverse effects” [Mesh] OR “Air Pollution/adverse effects” [Mesh] OR “Air Pollutants” [Pharmacological Action][22] OR “Environmental Exposure/adverse effects” [Mesh] OR “air pollutants” OR “air pollution” OR “air pollutant”
    - EMBASE: Auf eine Dauerrecherche wurde nach einem Probelauf von 2,5 Monaten verzichtet. Der Zusatzaufwand steht in keinem Verhältnis zum Ergebnis: Ein Teil der wichtigen Papers wird bereits über die PubMed-Suche gefunden. Die Handsuche wird auf die wichtigsten Zeitschriften, die via EMBASE erfasst werden, erweitert. Dies sollte die EMBASE-Suche ersetzen.
- Referenzlisten von Publikationen (Originalarbeiten und Reviews), Bibliographien
- Hinweise aus verschiedenen Quellen: Swiss TPH-intern, BAFU, WHO, Mitteilung anderer Forschungsteams.

Regelmäßige Handsuche in folgenden Zeitschriften

[22] Bringt keine zusätzlichen Treffer

Tabelle 21: Regelmäßige Handsuchen

Name	Art	Erscheinungskadenz	ISSN-Nr.
Air Quality Atmosphere and Health – Air Qual Atmos Health	Alert	vierteljährlich	1873-9318 1873-9326
American Journal of Epidemiology – Am J Epidemiology	Alert	PubMed 2/Monat	0002-9262 1476-6256
American Thoracic Society: e.g. American Journal of Respiratory and Critical Care Medicine- Am J Respir Crit Care Med	Alert Search query	wöchentlich	1073-449x 1535-4970
Asian Pacific Journal of tropical Biomedicine	Alert	monatlich	
Atmospheric Environment	Alert	monatlich	1352-2310
Environment International	Alert	monatlich	0160-4120 1873-6750
Environmental Health – Environ Health	Alert	Keine Angaben Wöchentlich?	1476-069x
Environmental Health Perspectives – Environ Health Perspect	Alert	Alle 3 Wochen 17/Jahr	0091-6765 1551-9924
Environmental Research – Environ Res	Alert	wöchentlich	0013-9351 1096-0953
Epidemiology – Epidemiol	Alert E-toc?	Alle 2 Monate	1531-5487 1044-3983
European Respiratory Journal – Eur Respir J	Alert	Monatlich	0903-1936 1399-3003
Inhalation Toxicology – Inhal Toxicol	Alert HTML	Alle 3-4 Wochen 14/Jahr	0895-8378 1091-7691
International Journal of Epidemiology – Int J Epidemiol	Alert	monatlich	
Journal of Air & Waste Management Association	Alert	monatlich	1096-2247
Journal of Environmental Protection	Alert	1/Monat	
Journal of Exposure Science and Environmental Epidemiology – J Expo Sci Environ Epidemiol	Alert	Alle 2 Monate	1559-0631 1559-064x
Lancet Respiratory Medicine	Alert	wöchentlich	
Occupational and Environmental Medicine – Occup Environ Med	Alert Etoc	monatlich	1351-0711 1470-7926
Science of the Total Environment – Sci Total Enviro	Alert	wöchentlich	0048-9697 1879-1026

Name	Art	Erscheinungska-denz	ISSN
Lancet	Alert	Weekly	0140-6736 1474-547x
Journal of the American Medical Associa-tion – JAMA	Alert	Weekly	0098-7484 1538-3598
British Medical Journal – BMJ	Alert	Weekly	0959-8138 1756-1833
New England Journal of Medicine – N Eng J Med	Alert	Weekly	0028-4793 1533-4406
Swiss Medical Weekly – Swiss Med Wkly	Alert	monatlich	1424-7860 1424-3997

Einschlusskriterien LUDOK

Es werden vor allem Originalarbeiten eingeschlossen, die für die Schweiz bzw. den europäischen Kontext oder das Verständnis von (weltweiten) Belastungs-Wirkungsbeziehungen relevant sind und sich mit Wirkungen von Schadstoffen befassen, welche in der Luftreinhalteverordnung reguliert werden bzw. für die eine Regulierung diskutiert werden. Die Literatur wird systematisch gesucht, allerdings werden nur die in diesem Kontext relevanten Studien in die Datenbank aufgenommen.

Bei Zeitreihenstudien ist man dazu übergegangen, nur noch Studien aufzunehmen, wenn sie neue Zielgrössen untersuchen oder wenn sie ein Multi-pollutant-Modell rechnen.

Tierstudien sind dann interessant, wenn der Expositionspfad inhalativ (keine Instillation, keine Aufnahme durch die Nahrung) erfolgt und die Expositionsdauer langfristig ist, also langfristige Folgen untersucht werden.

8.3 UKD search strategy in MEDLINE

Date: 11.05.2017

Languages: German, English;

Search period: 01.01.2011 – 11.05.2017

Since our search was based on the HEI-review, we searched from 01.01.2011. The search strategy was developed in collaboration with the project team and in accordance with the UBA.

	Suchwort	Field-Tag	Treffer
#1	"Particulate matter"	[All Fields]	9.159
#2	"Environmental exposure"	[All Fields]	16.540
#3	"Air Pollutants"	[All Fields]	24.235
#4	„Air Pollution"	[All Fields]	13.120
#5	"Air pollutant"	[All Fields]	761
#6	"Air Pollutants/adverse effects"	[Mesh]	4.018
#7	"Air Pollution/adverse effects"	[Mesh]	3.654
#8	"Environmental Exposure/adverse effects"	[Mesh]	11.065
#9	#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8		50.395
#10	"Surface area"	[All Fields]	22929
#11	"Ultrafine"	[All Fields]	1816
#12	"Ultrafine particle"	[All Fields]	174
#13	"Ultrafine particles"	[All Fields]	540
#14	"Nano particle"	[All Fields]	247
#15	"Nano particles"	[All Fields]	779
#16	Nanoparticle	[All Fields]	28.010
#17	Nanoparticles	[All Fields]	90.587
#18	$PM_{0.1}$	[All Fields]	24
#19	$PM_{0.25}$	[All Fields]	6
#20	PNC	[All Fields]	417
#21	"Particle number"	[All Fields]	813
#22	"Accumulation mode"	[All Fields]	95
#23	"Aitken mode"	[All Fields]	10
#24	Submicron*	[All Fields]	1.542
#25	#10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21 OR #22 OR #23 OR #24		127.172
#26	health	[All Fields]	1.294.298
#27	epidemiology	[All Fields]	633.487

#28	epidemiological	[All Fields]	285.652
#29	epidemiologic	[All Fields]	525.623
#30	#26 OR #27 OR #28 OR #29		1.737.258
#31	**#9 AND #25 AND #30**		**1.100**

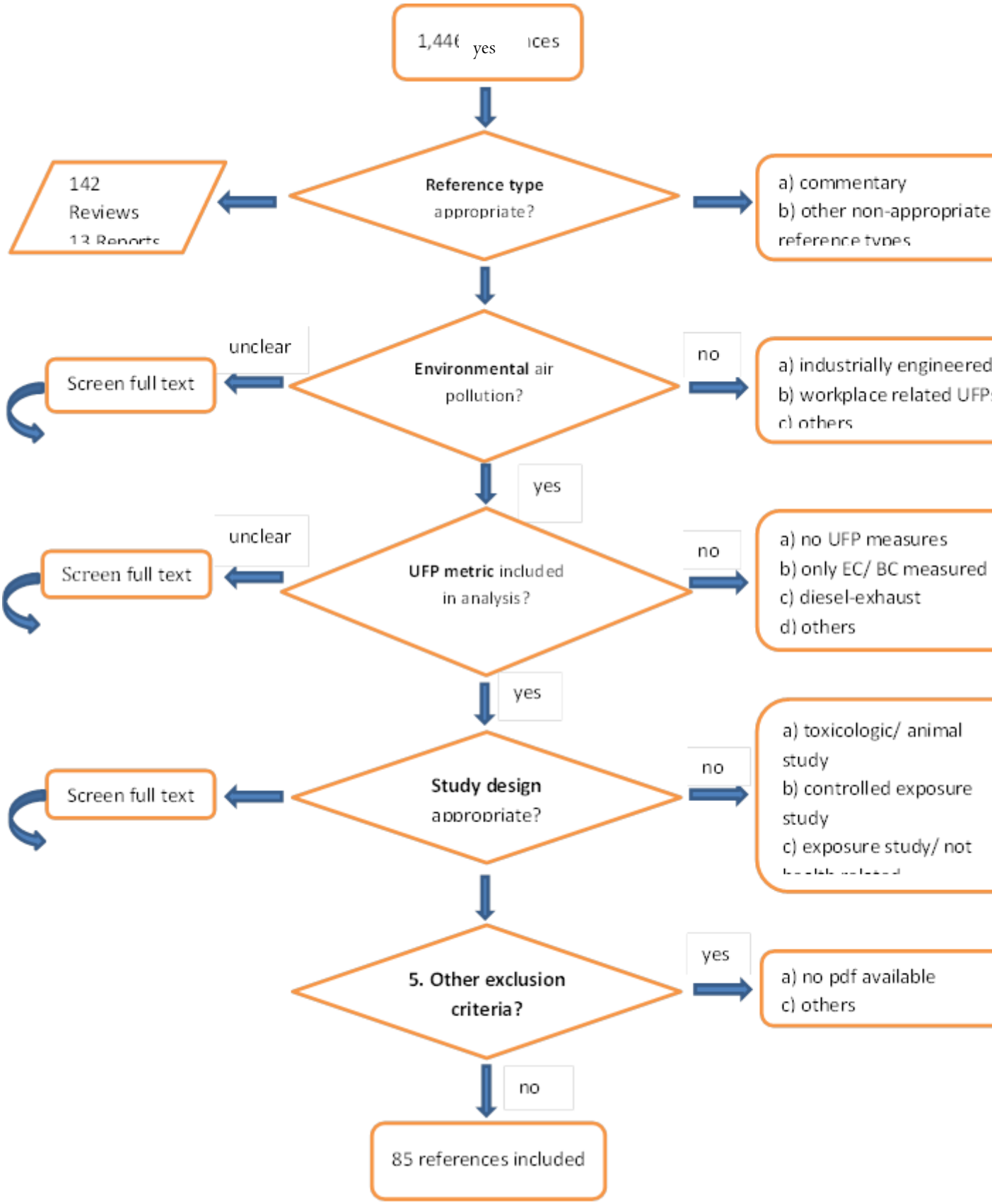
Additional search strategy using specific health outcomes (based on a template by the UBA)

	Suchwort	**Field-Tag**	**Treffer**
#1	"Particulate matter"	[All Fields]	9.159
#2	"Environmental exposure"	[All Fields]	16.540
#3	"Air Pollutants"	[All Fields]	24.235
#4	„Air Pollution"	[All Fields]	13.120
#5	"Air pollutant"	[All Fields]	761
#6	"Air Pollutants/adverse effects"	[Mesh]	4.018
#7	"Air Pollution/adverse effects"	[Mesh]	3.654
#8	"Environmental Exposure/adverse effects"	[Mesh]	11.065
#9	#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8		50.395
#10	"Surface area"	[All Fields]	22.929
#11	"Ultrafine"	[All Fields]	1816
#12	"Ultrafine particle"	[All Fields]	174
#13	"Ultrafine particles"	[All Fields]	540
#14	"Nano particle"	[All Fields]	247
#15	"Nano particles"	[All Fields]	779
#16	Nanoparticle	[All Fields]	28.010
#17	Nanoparticles	[All Fields]	90.587
#18	$PM_{0.1}$	[All Fields]	24
#19	$PM_{0.25}$	[All Fields]	6
#20	PNC	[All Fields]	417
#21	"Particle number"	[All Fields]	813
#22	"Accumulation mode"	[All Fields]	95
#23	"Aitken mode"	[All Fields]	10
#24	Submicron*	[All Fields]	1.542
#25	#10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21 OR #22 OR #23 OR #24		127.172
#26	Cardiovascular	[All Fields]	361.048
#27	vascular	[All Fields]	266.626

#28	cardiopulmonar*	[All Fields]	20.641
#29	ischaemic	[All Fields]	65.541
#30	ischemic	[All Fields]	90.058
#31	“myocardial infarction”	[All Fields]	49.221
#32	“heart attack”	[All Fields]	1.137
#33	“Cardiac infarction”	[All Fields]	84
#34	infarction	[All Fields]	63.263
#35	stroke	[All Fields]	99.768
#36	respirator*	[All Fields]	142.108
#37	pulmonar*	[All Fields]	145.534
#38	lung	[All Fields]	190.267
#39	asthma	[All Fields]	40.422
#40	copd	[All Fields]	25.258
#41	cancer	[All Fields]	918.403
#42	carcinoma	[All Fields]	209.814
#43	carcinogen*	[All Fields]	52.748
#44	malignan*	[All Fields]	141.485
#45	neoplas*	[All Fields]	570.667
#46	tumor	[All Fields]	825.207
#47	infectio*	[All Fields]	521.702
#48	disease	[All Fields]	1.146.966
#49	chronic inflammat*	[All Fields]	20.365
#50	systemic inflammat*	[All Fields]	12.404
#51	inflammat*	[All Fields]	326.167
#52	hospitaliz*	[All Fields]	72.874
#53	hospitalis*	[All Fields]	9.680
#54	“hospital admission”	[All Fields]	7.569
#55	emergency	[All Fields]	116.010
#56	mortality	[All Fields]	339.535
#57	death	[All Fields]	225.298
#58	depression	[All Fields]	111.407
#59	depressive	[All Fields]	47.920
#60	neurodegenerati*	[All Fields]	51.171
#61	alzheimer's	[All Fields]	50.714
#62	alzheimer*	[All Fields]	52.971
#63	parkinson's	[All Fields]	34.364

#64	parkinson*	[All Fields]	38.799
#65	dementia	[All Fields]	55.584
#66	diabetic	[All Fields]	64.131
#67	diabetes	[All Fields]	202.543
#68	metabolic	[All Fields]	175.842
#69	“low birth weight”	[All Fields]	9.703
#70	“low birthweight”	[All Fields]	1.476
#71	“preterm birth”	[All Fields]	6.737
#72	“premature birth”	[All Fields]	6.581
#73	“preterm delivery”	[All Fields]	3.086
#74	“premature delivery”	[All Fields]	438
#75	“premature infant”	[All Fields]	11.006
#76	“premature baby”	[All Fields]	76
#77	stillbirth	[All Fields]	3.267
#78	“dead birth”	[All Fields]	0
#79	stillborn	[All Fields]	572
#80	“immune system”	[All Fields]	37.177
#81	allergi*	[All Fields]	30.540
#82	#26 OR #27 OR #28 OR #29 ….#81		3.203.196
#83	**#9 AND #25 AND #82**		993

8.4 Flowchart on the selection of the studies



8.5 Repeated search in MEDLINE on 23.02.2018 in MEDLINE

Lucht S, Hennig F, Matthiessen C, Ohlwein S, Icks A, Moebus S, Jöckel K-H, Jakobs H, Hoffmann B. (in press). Air pollution and glucose metabolism: An analysis in non-diabetic participants of the Heinz Nixdorf Recall study. Accepted by Environ Health Perspect (in press, not yet indexed in MEDLINE).

Hennig F, Quass U, Hellack B, Küpper M, Kuhlbusch T, Stafoggia M, Hoffmann B. Ultrafine and Fine Particle Number and Surface Area Concentrations and Daily Cause-Specific Mortality in the Ruhr Area, Germany, 2009–2014. Environ Health Perspect. 2018; 126(2):1–10.; DOI:10.1289/EHP2054 (not yet indexed in MEDLINE).

Pilz V, Wolf K, Breitner S, Rückerl R, Koenig W, Rathmann W, Cyrys J, Peters A, Schneider A; KORA-Study group. C-reactive protein (CRP) and long-term air pollution with a focus on ultrafine particles. Int J Hyg Environ Health. 2018 Jan 31. pii: S1438-4639(17)30490-X. doi: 10.1016/j.ijheh.2018.01.016. [Epub ahead of print] PubMed PMID: 29428699.

Liu JY, Hsiao TC, Lee KY, Chuang HC, Cheng TJ, Chuang KJ. Association of ultrafine particles with cardiopulmonary health among adult subjects in the urban areas of northern Taiwan. Sci Total Environ. 2018 Jan 30;627:211-215. doi:10.1016/j.scitotenv.2018.01.218. [Epub ahead of print] PubMed PMID: 29426143.

Krauskopf J, Caiment F, van Veldhoven K, Chadeau-Hyam M, Sinharay R, Chung KF, Cullinan P, Collins P, Barratt B, Kelly FJ, Vermeulen R, Vineis P, de Kok TM, Kleinjans JC. The human circulating miRNome reflects multiple organ disease risks in association with short-term exposure to traffic-related air pollution. Environ Int. 2018 Jan 27;113:26-34. doi: 10.1016/j.envint.2018.01.014. [Epub ahead of print] PubMed PMID: 29421404.

Bai L, Chen H, Hatzopoulou M, Jerrett M, Kwong JC, Burnett RT, van Donkelaar A, Copes R, Martin RV, van Ryswyk K, Lu H, Kopp A, Weichenthal S. Exposure to Ambient Ultrafine Particles and Nitrogen Dioxide and Incident Hypertension and Diabetes. Epidemiology. 2018 Jan 9. doi: 10.1097/EDE.0000000000000798. [Epub ahead of print] PubMed PMID: 29319630.

Sinharay R, Gong J, Barratt B, Ohman-Strickland P, Ernst S, Kelly FJ, Zhang JJ, Collins P, Cullinan P, Chung KF. Respiratory and cardiovascular responses to walking down a traffic-polluted road compared with walking in a traffic-free area in participants aged 60 years and older with chronic lung or heart disease and age-matched healthy controls: a randomised, crossover study. Lancet. 2018 Jan 27;391(10118):339-349. doi: 10.1016/S0140-6736(17)32643-0. Epub 2017 Dec 5. Erratum in: Lancet. 2018 Jan 27;391(10118):308. PubMed PMID: 29221643; PubMed
Central PMCID: PMC5803182.

Forns J, Dadvand P, Esnaola M, Alvarez-Pedrerol M, López-Vicente M, Garcia-Esteban R, Cirach M, Basagaña X, Guxens M, Sunyer J. Longitudinal association between air pollution exposure at school and cognitive development in school children over a period of 3.5 years. Environ Res. 2017 Nov;159:416-421. doi: 10.1016/j.envres.2017.08.031. Epub 2017 Sep 1. PubMed PMID: 28858754.

Endes S, Schaffner E, Caviezel S, Dratva J, Stolz D, Schindler C, Künzli N, Schmidt-Trucksäss A, Probst-Hensch N. Is physical activity a modifier of the association between air pollution and arterial stiffness in older adults: The SAPALDIA cohort study. Int J Hyg Environ Health. 2017 Aug;220(6):1030-1038. doi: 10.1016/j.ijheh.2017.06.001. Epub 2017 Jun 13. PubMed PMID: 28629640.

Goldberg MS, Labrèche F, Weichenthal S, Lavigne E, Valois MF, Hatzopoulou M, Van Ryswyk K, Shekarrizfard M, Villeneuve PJ, Crouse D, Parent MÉ. The association between the incidence of postmenopausal breast cancer and
concentrations at street-level of nitrogen dioxide and ultrafine particles. Environ Res. 2017 Oct;158:7-15. doi: 10.1016/j.envres.2017.05.038. Epub 2017 Jun 5. PubMed PMID: 28595043.

Li Y, Lane KJ, Corlin L, Patton AP, Durant JL, Thanikachalam M, Woodin M, Wang M, Brugge D. Association of Long-Term Near-Highway Exposure to Ultrafine Particles with Cardiovascular Diseases, Diabetes and Hypertension. Int J Environ Res Public Health. 2017 Apr 26;14(5). pii: E461. doi: 10.3390/ijerph14050461. PubMed PMID: 28445425; PubMed Central PMCID: PMC5451912.

Bell G, Mora S, Greenland P, Tsai M, Gill E, Kaufman JD. Association of Air Pollution Exposures with High-Density Lipoprotein Cholesterol and Particle Number: The Multi-Ethnic Study of Atherosclerosis. Arterioscler Thromb Vasc Biol.
2017 May;37(5):976-982. doi: 10.1161/ATVBAHA.116.308193. Epub 2017 Apr 13. PubMed PMID: 28408373; PubMed Central PMCID: PMC5407952.

Weichenthal S, Lavigne E, Valois MF, Hatzopoulou M, Van Ryswyk K, Shekarrizfard M, Villeneuve PJ, Goldberg MS, Parent ME. Spatial variations in ambient ultrafine particle concentrations and the risk of incident prostate cancer: A case-control study. Environ Res. 2017 Jul;156:374-380. doi: 10.1016/j.envres.2017.03.035. Epub 2017 Apr 10. PubMed PMID: 28395241.

8.6 Indikators to describe and evaluate UFP studies

a) General Study Information

1. **Reference** [author et al. (year)]

   _____
   (free text)

   Example: Lane et al. (2016)

Source: Custom-made

2. **Link to PubMed** (Endnote reference below abstract)

   _____
   (free text)

Source: Custom-made

3. **Was the research question or objective in this paper clearly stated?** *Did the authors describe their goal in conducting this research? Is it easy to understand what they were looking to find? This issue is important for any scientific paper of any type. Higher quality scientific research explicitly defines a research question.*

   a) Yes
   b) No
   c) Not applicable

Source: Modified after QAT (for Observational Cohort and Cross-Sectional studies) **Question 1** –possible answer categories .

4. **What is the location of the study?** *[City, Country]*

_____

(free text)

Example: Copenhagen, Denmark

Source: Custom made

5. **Which world region is the country of the study assigned?**
   a) Africa
   b) North America
   c) South America
   d) Western Europe
   e) Eastern Europe
   f) South-East Asia
   g) Western Pacific
   h) Multiple Regions

Source: http://www.who.int/about/regions/en/

6. **What is the study name/ project abbreviation?** (e.g., ESCAPE)

_____

(free text)

   a) Not applicable
   b) Not reported/ reference given
   c) Not reported/ no reference given

-> *Use abbreviation + "study", e.g., ESCAPE study*

Source: Custom-made

7. **What is the cohort name?**

_____

(free text)

d) Not applicable
e) Not reported/ reference given
f) Not reported/ no reference given

-> *Use abbreviation + "cohort", e.g., SAPALDIA cohort*

Source: Custom-made

8. What was the study design?

a) Cohort	Long-term outcome
b) Case-control	
c) Case-crossover	
d) Cross-sectional	
e) Panel (cross-sectional)	Short-/Medium-Term
f) Panel (repeated measures)	Short-/Medium-Term
g) Scripted exposures	Particip. is assigned to prespecified expo, for example a specific bike route through a city
h) Time-series	
i) Other	

-> No free text answers allowed, if unclear state "Other".

Source: Custom-made

9. If other study design used, specify/Further details on study design, e.g., repeated measures (in cohort). Otherwise, leave free.

_____
(free text)

a) Not determinable
b) Not reported

Source: Custom-made

10. What was the time horizon of the study? (Filter question)

a) Short-term (hours to days)
b) Medium-term (weeks)
c) Long-term (months to years)
d) Combination of Short- and Long-term
e) Not reported

Source: Custom-made

11. Was it a multicenter-study?

a) Yes
b) No

Source: Custom-made

b) Specific aspects of study design

12. **Was the study population clearly specified and defined?** *Did the authors describe the group of people from which the study participants were selected or recruited, using demographics, location, and time period? If you were to conduct this study again, would you know who to recruit, from where, and from what time period? Is the cohort population free of the outcomes of interest at the time they were recruited? An example would be men over 40 years old with type 2 diabetes who began seeking medical care at Phoenix Good Samaritan Hospital between January 1, 1990 and December 31, 1994. In this example, the population is clearly described as: (1) who (men over 40 years old with type 2 diabetes); (2) where (Phoenix Good Samaritan Hospital); and (3) when (between January 1, 1990 and December 31, 1994). Another example is women ages 34 to 59 years of age in 1980 who were in the nursing profession and had no known coronary disease, stroke, cancer, hypercholesterolemia, or diabetes, and were recruited from the 11 most populous States, with contact information obtained from State nursing boards. In cohort studies, it is crucial that the population at baseline is free of the outcome of interest. For example, the nurses' population above would be an appropriate group in which to study incident coronary disease. You may need to look at prior papers on methods in order to make the assessment for this question. Those papers are usually in the reference list.*

a) Yes
b) Not specified/ reference given
c) Not specified/ no reference given
d) Not applicable

Source: Modified after **Question 2 of QAT** (Cohort and cross-sectional studies), modified answer categories.

13. **What was the sample size of the main study sample?**

_____
(free text)

a) Not reported

-> Write numbers without separation marks, e.g.: 1503

Source: Custom-made

14. What was **the main study population?** *(refers to the study group of the main analysis, e.g., male > 65 yrs)*

_____
(free text)

a) General population
b) Not reported/ reference given
c) Not reported/ no reference given

Examples:	Healthy Adults, > 40 yrs	Men with CAD, 35 - 70 yrs, Nonsmoking

-> After each characteristic, separate by a comma and press ALT + Enter and use a new line

15. What was the sample type of the study population? Convenience/ Random sample?

a) Convenience sample
b) Random sample
c) Random + Convenience sample
d) Other
e) Not reported/ reference given
f) Not reported/ no reference given
g) Not applicable

-> No freetext answers allowed, if unclear state "Other".

-> random sample: Zufallsstichprobe au seiner vorhandenen Gesamtpopulation (es muss also eine Liste mit allen potentiellen Teilnehmern vorliegen). Z. B. Kohortenstudie, bei der aus dem Einwohnermelderegister zufällig gezogen wurde.

-> convenience: Probanden werden gezielt angesprochen, z. B. Bewohner in der Nähe eines Monitors, Kinder in Schulklassen, Kranke im Krankenhaus, etc. , convenience ist auch z. B. die ACS-study (Nachbarn und Freunde der ACS-Mitlgieder)

-> Mischform: z. B. aus allen Schulen einer Stadt werden 3 zufällig ausgewählt, dann werden die Kinder um Teilnahme gebeten. Oder Subgruppe einer größeren Kohorte (random sample), die bei einer Zusatzstudie mitmachen.

Source: Custom-made

16. **What was the response rate of the study**? [e.g., 58%] *If fewer than 50% of eligible persons participated in the study, then there is concern that the study population does not adequately represent the target population. This increases the risk of bias.*

_____

(free text)

a) Not reported/ reference given

b) Not reported/ no reference given

-> Time-series and convenience sample: not applicable.

Source: **Modified after question 3 of QAT** (Cohort and cross-sectional studies), with modified answer categories (see QAT in appendix).

17. Was a sample size justification or power description provided?

Did the authors present their reasons for selecting or recruiting the number of people included or analyzed? Do they note or discuss the statistical power of the study? This question is about whether or not the study had enough participants to detect an association if one truly existed. A paragraph in the methods section of the article may explain the sample size needed to detect a hypothesized difference in outcomes. You may also find a discussion of power in the discussion section (such as the study had 85 percent power to detect a 20 percent increase in the rate of an outcome of interest, with a 2-sided alpha of 0.05). Sometimes estimates of variance and/or estimates of effect size are given, instead of sample size calculations. In any of these cases, the answer would be "yes." However, ***observational cohort studies*** *often do not report anything about power or sample sizes because the analyses are exploratory in nature. In this case, the answer would be "no." This is not a "fatal flaw." It just may indicate that attention was not paid to whether the study was sufficiently sized to answer a prespecified question–i.e., it may have been an exploratory, hypothesis-generating study.*

a) Yes
b) Not reported/ reference given
c) Not reported/ no reference given
d) Not applicable

-> A simple reference to design paper is not sufficient. Select yes only in case that authors refer to a sample size calculation for this analysis.

Source: Modified after question 3 of QAT (Cohort and cross-sectional studies), with modified answer categories (see QAT in appendix).

18. Were all the subjects selected or recruited from the same or similar populations?

Were the same underlying criteria used for all of the subjects involved? This issue is related to the description of the study population, above, and you may find the information for both of these questions in the same section of the paper. Most ***cohort studies*** *begin with the selection of the cohort; participants in this cohort are then measured or evaluated to determine their exposure status. However, some cohort studies may recruit or select exposed participants in a different time or place than unexposed participants, especially retrospective cohort studies–which is when data are obtained from the past (retrospectively), but the analysis examines exposures prior to outcomes. For example, one research question could be whether diabetic men with clinical depression are at higher risk for*

cardiovascular disease than those without clinical depression. So, diabetic men with depression might be selected from a mental health clinic, while diabetic men without depression might be selected from an internal medicine or endocrinology clinic. This study recruits groups from different clinic populations, so this example would get a "no."

a) Yes
b) No
c) Not reported/ reference given
d) Not reported/ no reference given
e) Not applicable

Source: Modified after question 4 (part 1) of QAT (Cohort and cross-sectional studies), with modified answer categories (see QAT in appendix).

19. **If case-control study**, how was **the selection of controls**?
    a) Community controls
    b) Hospital controls
    c) Other
    d) Not reported/ reference given
    e) Not reported/ no reference given
    f) Not applicable

Source: Modified after **question 3 of NOS/**Selection (Case-control studies.

20. **If case-control study, were controls selected or recruited from the same or similar population that gave rise to the cases?** *To determine whether cases and controls were recruited from the same population, one can ask hypothetically, "If a control was to develop the outcome of interest (the condition that was used to select cases), would that person have been eligible to become a case?" Case-control studies begin with the selection of the cases (those with the outcome of interest, e.g., lung cancer) and controls (those in whom the outcome is absent). Cases and controls are then evaluated and categorized by their exposure status. For the lung cancer example, cases and controls were recruited from hospitals in a given region. One may reasonably assume that controls in the catchment area for the hospitals, or those already in the hospitals for a different reason, would attend those hospitals if they became a case; therefore, the controls are drawn from the same population as the cases. If the controls were recruited or selected from a different region (e.g., a State other than Texas) or time period (e.g., 1991-2000), then the cases and controls were recruited from different populations, and the answer to this question would be "no." The following* ***example*** *further explores selection of controls. In a study, eligible cases were men and women, ages 18 to 39, who were diagnosed with atherosclerosis at hospitals in Perth, Australia, between July 1, 2000 and December 31, 2007. Appropriate controls for these cases might be sampled using voter registration information for men and women ages 18 to 39, living in Perth (population-based controls); they also could be sampled from patients without atherosclerosis at the same hospitals (hospital-based controls). As long as the controls are individuals who would have been eligible to be included in the study as cases (if they had been diagnosed with atherosclerosis), then the controls were selected appropriately from the same source population as cases. In a prospective case-control study, investigators may enroll individuals as cases at the time they are found to have the outcome of interest; the number of cases usually increases as time progresses. At this same time, they may recruit or select con-*

trols from the population without the outcome of interest. One way to identify or recruit cases is through a surveillance system. In turn, investigators can select controls from the population covered by that system. This is an example of population-based controls. Investigators also may identify and select cases from a ***cohort study population*** *and identify controls from outcome-free individuals in the same cohort study. This is known as a nested case-control study.*

a) Yes
b) No
c) Not reported/ reference given
d) Not reported/no reference given
e) Not applicable

Source: Modified after **question 4 (part 1) of QAT** (Case-control studies).

21. **Were all the subjects selected or recruited from the same time period**? *(...) However, some cohort studies may recruit or select exposed participants in a different time or place than unexposed participants, especially retrospective cohort studies–which is when data are obtained from the past (retrospectively), but the analysis examines exposures prior to outcomes.*

a) Yes
b) No
c) Not reported/ reference given
d) Not reported/no reference given
e) Not applicable

Source: Modified after **question 4 (part 2) of QAT** (Cohort and cross-sectional studies)

22. **Were inclusion and exclusion criteria for being in the study prespecified?** *Were the inclusion and exclusion criteria developed prior to recruitment or selection of the study population?*

a) Yes
b) No
c) Not reported/ reference given
b) Not reported/no reference given
c) Not applicable

Source: Modified after **question 4 (part 3) of QAT** (Cohort and cross-sectional studies).

23. Is the analyzed sample representative for the general population?

a) Yes, completely representative.
b) Yes, somewhat representative.
b) Not representative/ selected group
d) Not applicable

Source: Modified after **question 1 of NOS/**Selection (Cohort studies).

- ➔ Completely representative only for whole population studies (time series, register-based, possibly also administrative data,
- ➔ For example if random sample of a subgroup, then b) for example a representative sample of all children or of all adults above a certain age

24. If cohort study: Is lost to follow-up after baseline provided?

a) Yes
b) Not reported/ reference given
c) Not reported/no reference given
d) Not applicable

25. Are losses to follow-up likely to introduce bias?

Higher overall follow-up rates are always better than lower follow-up rates, even though higher rates are expected in shorter studies, whereas lower overall follow-up rates are often seen in studies of longer duration. Usually, an acceptable overall follow-up rate is considered 80 percent or more of participants whose exposures were measured at baseline. However, this is just a general guideline. For example, a 6-month cohort study examining the relationship between dietary sodium intake and BP level may have over 90 percent follow-up, but a 20-year cohort study examining effects of sodium intake on stroke may have only a 65 percent follow-up rate.

a) Yes
b) No
c) Cannot determine
d) Not applicable (e.g., if not reported)

Source: Modified after **question 13 of QAT** (Cohort and cross-sectional studies) and NOS/Outcome, Question 3 (Cohort studies)

26. What was the study period? [month/year]

(free text)

a) Not reported/ reference given
b) Not reported/no reference given
c) Not applicable

Example: ***03/2003-08/2004***

Source: Modified after HEI data extraction file (original: Study period, free text), answer categories inspired by QAT.

27. Was the timeframe sufficient so that one could reasonably expect to see an association between exposure and outcome if it existed?

Did the study allow enough time for a sufficient number of outcomes to occur or be observed, or enough time for an exposure to have a biological effect on an outcome? In the examples given above, if clinical depression has a biological effect on increasing risk for CVD, such an effect may take years. In the other example, if higher dietary sodium increases BP, a short timeframe may be sufficient to assess its association with BP, but a longer timeframe would be needed to examine its association with heart attacks.
The issue of timeframe is important to enable meaningful analysis of the relationships between exposures and outcomes to be conducted. This often requires at least several years, especially when looking at health outcomes, but it depends on the research question and outcomes being examined. ***Cross-sectional analyses*** *allow no time to see an effect, since the exposures and outcomes are assessed at the same time, so those would get a "no" response.*

a) Yes
b) No
a) Not reported/ reference given
b) Not reported/no reference given
c) Not applicable

Source: Modified after **question 7 of QAT** (Cohort and cross-sectional studies): Answer categories were modified.

27. **For the analyses in this paper, were the exposure(s) of interest measured prior to the outcome(s) being measured?** *This question is important because, in order to determine whether an exposure causes an outcome, the exposure must come before the outcome. For some prospective cohort studies, the investigator enrolls the cohort and then determines the exposure status of various members of the cohort (large epidemiological studies like Framingham used this approach). However, for other cohort studies, the cohort is selected based on its exposure status, as in the example above of depressed diabetic men (the exposure being depression). Other examples include a cohort identified by its exposure to fluoridated drinking water and then compared to a cohort living in an area without fluoridated water, or a cohort of military personnel exposed to combat in the Gulf War compared to a cohort of military personnel not deployed in a combat zone. With either of these types of cohort studies, the cohort is followed forward in time (i.e., prospectively) to assess the outcomes that occurred in the exposed members compared to nonexposed members of the cohort. Therefore, you begin the study in the present by looking at groups that were exposed (or not) to some biological or behavioral factor, intervention, etc., and then you follow them forward in time to examine outcomes. If a cohort study is conducted properly, the answer to this question should be "yes," since the exposure status of members of the cohort was determined at the beginning of the study before the outcomes occurred. For retrospective cohort studies, the same principal applies. The difference is that, rather than identifying a cohort in the present and following them forward in time, the investigators go back in time (i.e., retrospectively) and select a cohort based on their exposure status in the past and then follow them forward to assess the outcomes that occurred in the exposed and non-exposed cohort members. Because in retrospective cohort studies the exposure and outcomes may have already occurred (it depends on how long they follow the cohort), it is important to make sure that the exposure preceded the outcome. Sometimes cross-sectional studies are conducted (or cross-sectional analyses of cohort-study data), where the exposures and outcomes are measured during the same timeframe. As a result, cross-sectional analyses provide weaker evidence than regular cohort studies regarding a potential causal relationship between exposures and outcomes. For cross-sectional analyses, the answer to Question 6 should be "no."*

a) Yes

b) No

c) Not reported/ reference given

d) Not reported/ no reference given

d) Not applicable

Source: Modified after question 6 of QAT (Cohort and cross-sectional studies): Answer categories were modified.

c) Exposure assessment

28. Which type of exposure assessment technique was used (filter question)?

a) Model based

b) Measurements only

c) Other

-> No freetext answers allowed, if unclear state "Other".

Source: custom-made

29. If other exposure assessment technique was used, specify

(free text)

a) Not applicable

Source: custom-made

30. Which exposure assessment technique was used?

a) LUR

b) LUR: Spatio-temporal

c) CTM

d) Dispersion

e) Interpolation

f) Hybrid

g) Microscale personal exposure model

h) Measurement: satellite

i) Measurement: central site *(if only one measurement station was used)*

j) Measurement: residential

l) Measurement: mobile *(attached to car, bicycle, person)*

m) Other

-> No freetext answers allowed, if unclear state "Other".

Source: custom-made

31. If other exposure assessment technique was used, specify

(free text)

a) not applicable

32. What was the spatial resolution of the exposure? (E.g., 1x1km)

a) Mobile (for example personal or on bike or google cars)
b) Address-specific
c) Postal/ zip-code
d) City
e) 1x1 km^2
f) 5x5 km^2
g) 10x10 km^2
h) Other
i) No spatial resolution (for example only one monitor in one city)
j) Not reported/ reference given
k) Not reported/ no reference given

l) Not applicable

-> No freetext answers allowed, if unclear state "Other".

-> this only applies to the exposure assessment (model or measurements) and NOT to the assignment of exposure to the participants (separate question).

Source: custom-made

33. If other or unclear spatial resolution was used, specify

_____

(free text)

a) not applicable

34. **What was the temporal resolution of the exposure** measurement or modeling? Information on temp resol. of analysis in results section. Mehrfachnennung erlaubt [minute, hour, day, month, year, year-means].
If answer not included, specify as free text

_____

(free text)

a) Minute
b) Hour
c) Day
d) Month
e) Year
f) Year-means
g) Time-pattern
h) Not reported/ reference given

i) Not reported/ no reference given

j) Not applicable

Source: custom-made

35. To which level were the exposures allocated to participants?

If answer not included in list, specify as free text.

_____
(free text)

a) Mobile personal

b) Geocoded addresses (not corrected for mistakes in data base)

c) Microenvironments (incl. corrected addresses)

d) Zip code

e) City

f) County

g) Not reported/ reference given

h) Not reported/ no reference given

i) Not applicable

-> If exposure assessment was a central site measurement, select: "Not applicable"

Source: custom-made

36. Did the exposure assessment include a residential history?

a) Yes, complete or partial residential address history

b) No residential address history

c) Not reported / reference given

d) Not reported/ no reference given

e) Not applicable

*-> In case of **short-term studies**, select "not applicable"*

Source: custom-made

d) Assessment of UFP

37. Type of particle was assessed – UFP (ONLY below 100 nm) UFP in the most strict sense!

e) Yes

-> If UFP was not assessed, do not enter anything. The same procedure applies to the questions 39-42.

-> If the size fraction of UFP was not mentioned, select column 41) "Other" and specify as "not reported (42).

38. Type of particle was assessed – Quasi-UFP (PNC without cutpoint at 100 nm, for example total PNC or PNC 10-300 nm or PM0.25 or similar)

a) Yes

39. Type of particle was assessed – Surface Area

b) Yes

40. Type of particle was assessed - Other

c) Yes

41. If other type of particle was assessed, specify

_____

(free text)

b) Not reported

42. Particle metric - PNC?

-> If particle metric was not assessed, do not enter anything. The same pattern applies to the questions 44-51.

a) Yes

43. Particle metric – PM0.1?

b) Yes

44. Particle metric – PM0.25?

c) Yes

45. Particle metric – PM1.0?

d) Yes

46. Particle metric – Nucleation mode?

e) Yes

47. Particle metric – Aitken mode?

f) Yes

48. Particle metric – Accumulation mode?

g) Yes

49. Particle metric – Lung deposited surface area?

h) Yes

50. Particle metric – Other?

i) Yes

51. If other particle metric, specify

_____

(free text)

j) Not reported

52. Which size fractions were measured/modeled? Enter all fractions that were used in the analysis. Enter line change between each fraction (ALT + Enter)

x nm – y nm
(structured format)

a) Total
b) Not reported/ reference given
c) Not reported/ no reference given

d) Not applicable *(eg., LDSA)*

-> If no size fractions are mentioned, and a particle number counter was used, select "total"

Source: custom-made

53. **Which technical device was used to measure UFP?** (if various, give reference)

_____

(free text)

a) Various
b) Not reported/ reference given
c) Not reported/ no reference given
d) Not applicable

Source: custom-made

54. Was the measurement device/exposure model valid/reliable? (will be completed later)

55. Any mentioning of QA/QC measures described for the exposure assessment??

a) Yes
b) No
c) Not applicable

-> If a reference is given for QA/QC measures is given, select "Yes"

Source: custom-made

56. If QA/QC measures are referenced, specify

_____

(free text)

57. **Was the exposure assessment (independent variables) implemented consistently across all study participants?** Here is a final example that illustrates the point about why it is important to assess exposures consistently across all groups: If people with higher BP (exposed cohort) are seen by their providers more frequently than those without elevated BP (nonexposed group), it also increases the chances of detecting and documenting changes in health outcomes, including CVD-related events. Therefore, it may lead to the conclusion that higher BP leads to more CVD events. This may be true, but it could also be due to the fact that the subjects with higher BP were seen more often; thus, more CVD-related events were detected and documented simply because they had more encounters with the health care system. Thus, it could bias the results and lead to an erroneous conclusion.

a) Yes
b) No
c) Cannot determine
d) Not applicable

Source: Modified after **question 9** (partly) **of QAT**.

58. **Was the exposure assessment valid for the population?** Is the measurement/model appropriate to reflect the real exposure of the population?

a) Yes
b) No
c) Cannot determine
d) Not applicable

Source: Modified after **question 9 (partly) of QAT.**

59. **If cohort/panel/ crossover study, was the exposure assessed more than once over time?** Was the exposure for each person measured more than once during the course of the study period? Multiple measurements with the same result increase our

confidence that the exposure status was correctly classified. Also, multiple measurements enable investigators to look at changes in exposure over time, for example, people who ate high dietary sodium throughout the follow-up period, compared to those who started out high then reduced their intake, compared to those who ate low sodium throughout. Once again, this may not be applicable in all cases. In many older studies, exposure was measured only at baseline. However, multiple exposure measurements do result in a stronger study design.

a) Yes
b) No
c) Not reported/ reference given
d) Not reported/ no reference given
e) Not applicable

Source: Modified after QAT, Question 10 (answer categories).

f) Assessment of other exposures (air pollutants, noise, meteorologic data)

60. Were other air pollutants assessed?

f) Yes
g) No
h) Not applicable

Source: custom-made

Which technical device/exposure model was used to assess other air pollutants?

_____
(free text)

a) Not reported/ reference given

b) Not reported/ no reference given

c) Not applicable

-> If various, give reference

Source: custom-made

61. Was noise exposure assessed?

a) Yes, on residential level

b) Yes, on personal level

c) Yes, other

d) No

Source: custom-made

62. Was meteorological data measured/ modeled? (filter question)

a) Yes

b) No

c) Not applicable

Source: custom-made

63. Which meteorological data measured/ modeled? (MN)

If answer not included in list, specify as free text.

a) Temperature

b) Relative humidity

c) Barometric pressure

d) Precipitation

e) Season

f) Pollen counts
g) Other
h) Wind speed and direction
i) Not reported/ reference given
j) Not reported/ no reference given
k) Not applicable

Source: custom-made

64. How was meteorological data measured/ modeled? (MN)

a) Routine measurement
b) Study-specific measurement
c) Other
d) Not reported/ reference given
e) Not reported/ no reference given
f) Not applicable

Source: custom-made

65. Was neighborhood SES assessed? (filter question)

a) Yes
b) Not reported/ reference given
c) Not reported/ no reference given
d) Not applicable

Source: custom-made

66. How was neighborhood SES data measured/ modeled?

_____

(free text)

l) Not reported/ reference given

m) Not reported/ no reference given

n) Not applicable

Source: custom-made

67. What was the average submicron particle exposure of the study population (main analysis)?

_____

(free text)

a) Not reported/ reference given

b) Not reported/ no reference given

c) Not applicable

-> Specify if Mean or Median and add SD/IQR, if given.

-> Write numbers **without** any separation marks (“,” or “.”)

Example: Mean (SD): 15000 (4000)

Median (IQR): 13500 (3500)

g) Outcome assessment

68. Which outcome type was assessed? - Mortality

a) Yes

69. **Which outcome type was assessed? - Morbidity** (except emergency/admissions, etc.)

b) Yes

-> Code symptoms as morbidity

-> Except emergency/ hospital visits/admissions – see next question

70. Which outcome type was assessed? - Emergency/ hospital/ visits/ admissions

c) Yes

71. Which outcome type was assessed? - Subclinical

d) Yes

72. Which outcome type was assessed? - Other

e) Yes

Source: Custom-made

73. What was the main outcome of the study?

a) Total Mortality
b) Cardiovascular
c) Respiratory and atopy
d) inflammation
e) Oxidative stress
f) Neurocognitive
g) Other

-> No freetext answers allowed, if unclear state “Other”.

74. What was/were the specific outcome(s) of the study

_____
(free text)

d) Not reported/ reference given
e) Not reported/ no reference given
f) Not applicable

Source: custom-made

75. Were the outcome measures (dependent variables) clearly defined and implemented consistently across all study participants?

a) Yes
b) No
c) Not applicable

Source: Modified after **question 11 (partly) of QAT**.

76. How was the outcome assessed?

a) Standardized clinical examinations (e.g., in study center)
b) Official registry *(e.g., cancer registry)*
c) Administrative database *(e.g., insurance companies)*
d) Medical records *(e.g., hospital, general practitioner)*
e) Self-reported physician-diagnosed
f) Self-reported
g) Mobile device
h) Other

Source: custom-made

77. What was/were the ICD-codes of the outcome(s)?

_____

(free text)

a) Not reported
b) Not applicable

Source: custom-made

h) Statistical analysis

78. Which type of analysis was used?

a) Group comparison
b) Linear regression
c) Mixed linear regression
d) Logistic regression
e) Poisson regression
f) Cox-regression
g) Additive mixed model
h) Generalized estimated equation (GEE)
i) Other
j) Not reported/ reference given
k) Not reported/ no reference given
l) Not applicable

Source: custom-made

79. Which effect measure was estimated?

a) ß-estimates
b) %-change
c) OR
d) RR
e) HR
f) Other
g) No quantitative effect measures
h) Not applicable

Source: custom-made

80. Which unit of exposure was used?

a) Group comparison *(<=2)*
b) Categories *(>2)*
c) Fixed increment
d) IQR
e) Distribution based
f) Other
g) Not applicable

Source: custom-made

81.Absolute size of exposure unit?

_____
(free text)

c) Not reported

d) Not applicable

-> Write numbers without separation marks

Source: custom-made

82. Were the outcome assessors blinded to the exposure status resp. Case-control status of participants? Blinding means that outcome assessors did not know whether the participant was exposed or unexposed. It is also sometimes called "masking." The objective is to look for evidence in the article that the person(s) assessing the outcome(s) for the study (for example, examining medical records to determine the outcomes that occurred in the exposed and comparison groups) is masked to the exposure status of the participant. Sometimes the person measuring the exposure is the same person conducting the outcome assessment. In this case, the outcome assessor would most likely not be blinded to exposure status because they also took measurements of exposures. If so, make a note of that in the comments section.

a) Yes
b) No
c) Not reported/ reference given
d) Not reported/ no reference given
e) Not applicable

Source: Question 12 of QAT (Cohort and Cross-sectional, modified in question 11 (Case-Control))

83. Was the analysis adjusted for personal covariates? (e.g. demographic, lifestyle, medication)

a) Extensively
b) For main covariates
c) For some covariates
d) No
e) Not reported/ reference given
f) Not reported/ no reference given
g) Not applicable

Source: custom-made

84. Was the analysis adjusted for socioeconomic covariates?

a) Yes
b) No
c) Not reported/ reference given
d) Not reported/ no reference given
e) Not applicable

Source: custom-made

85. **Was the analysis adjusted for environmental covariates - NOISE?**

a) Yes

86. **Was the analysis adjusted for environmental covariates - METEOROLOGY**?

b) Yes

87. **Was the analysis adjusted for environmental covariates – Neighborhood SES?**

c) Yes

88. **Was the analysis adjusted for environmental covariates - Other?**

a) Yes

89. If adjusted for other environmental covariates, specify.

_____
(free text)

e) Not reported

f) Not applicable

Source: custom-made

90. Was the analysis adjusted for other air pollutants? / Were multi-pollutant-models conducted?

a) Yes

b) No

c) Not reported/ reference given

d) Not reported/ no reference given

e) Not applicable

Source: custom-made

91. For which co-pollutants were UFP-models adjusted?

_____

(free text)

g) Not reported

92. Covariate adjustment: List/ Specify

_____

(free text)

h) Not reported

-> Separate by commas, use capital letter for first entry, e.g.:

-> Age, dyslipidemia, prior MI, smoking, year, weekday, hour of the day, temperature, relative humidity

-> In case of different adjustment sets, separate by a), b), c) etc.

Source: custom-made

93. **Was confounder adjustment adequate?** Were key potential confounding variables measured and adjusted for, such as by statistical adjustment for baseline differences? Logistic regression or other regression methods are often used to account for the influence of variables not of interest. This is a key issue in cohort studies, because statistical analyses need to control for potential confounders, in contrast to an RCT, where the randomization process controls for potential confounders. All key factors that may be associated both with the exposure of interest and the outcome–that are not of interest to the research question–should be controlled for in the analyses. For example, in a study of the relationship between cardiorespiratory fitness and CVD events (heart attacks and strokes), the study should control for age, BP, blood cholesterol, and body weight, because all of these factors are associated both with low fitness and with CVD events. Well-done cohort studies control for multiple potential confounders.

a) Yes
b) Partly
c) No
d) Not applicable

Source: Modified after **question 14 of QAT** (Cohort and cross-sectional).

94. **If case-control and matching was used, did the investigators account for matching during study analysis?** Matching is a technique used to improve study efficiency and control for known confounders. For example, in the study of smoking and CVD events, an investigator might identify cases that have had a heart attack or stroke and then select controls of similar age, gender, and body weight to the cases. For case-control studies, it is important that if matching was performed during the selection or recruitment process, the variables used as matching criteria (e.g., age, gender, race) should be controlled for in the analysis.

a) Yes
b) No
c) Not reported/ reference given
d) Not reported/ no reference given
e) Not applicable

Source: Modified after **question 12 (part 2) of QAT** (Case-control studies)

95. For UFP-effect w/o co-pollutant adjustment: Was at least 1 estimate significantly elevated in the eypected adverse direction?

a) Yes

b) No

c) Not applicable

Source: custom-made

➔ If not clear, wich direction is expected and "adverse", generalize here to significantly changed

96. For UFP-effects w/o co-pollutant adjustment: Was a general pattern consistent with adverse association, regardless of significance?

a) Yes

b) No

c) Not applicable

Source: custom-made

97. For UFP-effect w/o co-pollutant adjustment: Were significant protective associations observed?

a) Yes

b) No

c) Not applicable

Source: custom-made

98. For UFP-effect with co-pollutant adjustment: Was at least 1 estimate significantly elevated in the expected adverse direction?

a) Yes

b) No

c) Not applicable

Source: custom-made

➔ If not clear, wich direction is expected and "adverse", generalize here to significantly changed

99. For UFP-effects with co-pollutant adjustment: Was a general pattern consistent with adverse association, regardless of significance?

a) Yes

b) No

c) Not applicable

Source: custom-made

100. For UFP-effect w/o co-pollutant adjustment: Were significant protective associations observed?

a) Yes

b) No

c) Not applicable

Source: custom-made

101. What was/were the size (incl. confidence intervals) of the UFP effect(s)? give estimate with most complete adjustment set. If estimate with and without copollutant is given, report both.

_____

(free text)

a) Not applicable

-> Use one line per estimate, write confidence intervals, separated by "-" in round brackets behind estimate).

-> In case of different outcomes/time lags, specify outcome/lag before estimates.

E.g.: 1-day: 1.03 (1.00-1.03)

2-day: 1.05 (1.02-1.07)

Source: custom-made

102. Was the model robust to the adjustment of other pollutant effects?

a) Yes
b) Mainly
c) Partly
d) No
e) Not applicable *(e.g., no adjustment for other pollutants)*

Source: custom-made

103. What was/were the effect size(s) of other pollutants?

_____
(free text)

a) Not reported/ reference given
b) Not reported/ no reference given
c) Not applicable

-> Format as UFP effect sizes.

-> Reference to table possible

Source: custom-made

104. Was the effect of other pollutants robust upon the inclusion of UFP?

a) Yes
b) Mainly
c) Partly
d) No
e) Not applicable *(e.g., no adjustment for UFP)*

Source: custom-made

105. Do sensitivity analyses support robustness of the associations? Does the main conclusion stays the same?

a) Yes

b) Partly

c) No

d) Not applicable *(e.g., no sensitivity analyses)*

Source: custom-made

106. Comments

_____

(free text)

107. Ersteingabe:

-> Name

108. Zweiteingabe

-> Name

109. Datum der Eingabe

-> z.B. 15.10.2017

8.7 Annexes

i) Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies

Criteria	Yes	No	Other (CD, NR, NA)*
1. Was the research question or objective in this paper clearly stated?			
2. Was the study population clearly specified and defined?			
3. Was the participation rate of eligible persons at least 50%?			
4. Were all the subjects selected or recruited from the same or similar populations (including the same time period)? Were inclusion and exclusion criteria for being in the study prespecified and applied uniformly to all participants?			
5. Was a sample size justification, power description, or variance and effect estimates provided?			
6. For the analyses in this paper, were the exposure(s) of interest measured prior to the outcome(s) being measured?			
7. Was the timeframe sufficient so that one could reasonably expect to see an association between exposure and outcome if it existed?			
8. For exposures that can vary in amount or level, did the study examine different levels of the exposure as related to the outcome (e.g., categories of exposure, or exposure measured as continuous variable)?			
9. Were the exposure measures (independent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?			
10. Was the exposure(s) assessed more than once over time?			

11. Were the outcome measures (dependent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?			
12. Were the outcome assessors blinded to the exposure status of participants?			
13. Was loss to follow-up after baseline 20% or less?			
14. Were key potential confounding variables measured and adjusted statistically for their impact on the relationship between exposure(s) and outcome(s)?			

Quality Rating (Good, Fair, or Poor) (see guidance)

Rater #1 initials:

Rater #2 initials:

Additional Comments (If POOR, please state why):

*CD, cannot determine; NA, not applicable; NR, not reported

Guidance for Assessing the Quality of Observational Cohort and Cross-Sectional Studies

The guidance document below is organized by question number from the tool for quality assessment of observational cohort and cross-sectional studies.

Question 1. Research question

Did the authors describe their goal in conducting this research? Is it easy to understand what they were looking to find? This issue is important for any scientific paper of any type. Higher quality scientific research explicitly defines a research question.

Questions 2 and 3. Study population

Did the authors describe the group of people from which the study participants were selected or recruited, using demographics, location, and time period? If you were to conduct

this study again, would you know who to recruit, from where, and from what time period? Is the cohort population free of the outcomes of interest at the time they were recruited?

An example would be men over 40 years old with type 2 diabetes who began seeking medical care at Phoenix Good Samaritan Hospital between January 1, 1990 and December 31, 1994. In this example, the population is clearly described as: (1) who (men over 40 years old with type 2 diabetes); (2) where (Phoenix Good Samaritan Hospital); and (3) when (between January 1, 1990 and December 31, 1994). Another example is women ages 34 to 59 years of age in 1980 who were in the nursing profession and had no known coronary disease, stroke, cancer, hypercholesterolemia, or diabetes, and were recruited from the 11 most populous States, with contact information obtained from State nursing boards.

In cohort studies, it is crucial that the population at baseline is free of the outcome of interest. For example, the nurses' population above would be an appropriate group in which to study incident coronary disease. This information is usually found either in descriptions of population recruitment, definitions of variables, or inclusion/exclusion criteria.

You may need to look at prior papers on methods in order to make the assessment for this question. Those papers are usually in the reference list.

If fewer than 50% of eligible persons participated in the study, then there is concern that the study population does not adequately represent the target population. This increases the risk of bias.

Question 4. Groups recruited from the same population and uniform eligibility criteria

Were the inclusion and exclusion criteria developed prior to recruitment or selection of the study population? Were the same underlying criteria used for all of the subjects involved? This issue is related to the description of the study population, above, and you may find the information for both of these questions in the same section of the paper.

Most cohort studies begin with the selection of the cohort; participants in this cohort are then measured or evaluated to determine their exposure status. However, some cohort studies may recruit or select exposed participants in a different time or place than unexposed participants, especially retrospective cohort studies–which is when data are obtained from the past (retrospectively), but the analysis examines exposures prior to outcomes. For example, one research question could be whether diabetic men with clinical

depression are at higher risk for cardiovascular disease than those without clinical depression. So, diabetic men with depression might be selected from a mental health clinic, while diabetic men without depression might be selected from an internal medicine or endocrinology clinic. This study recruits groups from different clinic populations, so this example would get a "no."

However, the women nurses described in the question above were selected based on the same inclusion/exclusion criteria, so that example would get a "yes."

Question 5. Sample size justification

Did the authors present their reasons for selecting or recruiting the number of people included or analyzed? Do they note or discuss the statistical power of the study? This question is about whether or not the study had enough participants to detect an association if one truly existed.

A paragraph in the methods section of the article may explain the sample size needed to detect a hypothesized difference in outcomes. You may also find a discussion of power in the discussion section (such as the study had 85 percent power to detect a 20 percent increase in the rate of an outcome of interest, with a 2-sided alpha of 0.05). Sometimes estimates of variance and/or estimates of effect size are given, instead of sample size calculations. In any of these cases, the answer would be "yes."

However, observational cohort studies often do not report anything about power or sample sizes because the analyses are exploratory in nature. In this case, the answer would be "no." This is not a "fatal flaw." It just may indicate that attention was not paid to whether the study was sufficiently sized to answer a prespecified question–i.e., it may have been an exploratory, hypothesis-generating study.

Question 6. Exposure assessed prior to outcome measurement

This question is important because, in order to determine whether an exposure causes an outcome, the exposure must come before the outcome.

For some prospective cohort studies, the investigator enrolls the cohort and then determines the exposure status of various members of the cohort (large epidemiological studies like Framingham used this approach). However, for other cohort studies, the cohort is selected based on its exposure status, as in the example above of depressed diabetic men (the

exposure being depression). Other examples include a cohort identified by its exposure to fluoridated drinking water and then compared to a cohort living in an area without fluoridated water, or a cohort of military personnel exposed to combat in the Gulf War compared to a cohort of military personnel not deployed in a combat zone.

With either of these types of cohort studies, the cohort is followed forward in time (i.e., prospectively) to assess the outcomes that occurred in the exposed members compared to nonexposed members of the cohort. Therefore, you begin the study in the present by looking at groups that were exposed (or not) to some biological or behavioral factor, intervention, etc., and then you follow them forward in time to examine outcomes. If a cohort study is conducted properly, the answer to this question should be "yes," since the exposure status of members of the cohort was determined at the beginning of the study before the outcomes occurred.

For retrospective cohort studies, the same principal applies. The difference is that, rather than identifying a cohort in the present and following them forward in time, the investigators go back in time (i.e., retrospectively) and select a cohort based on their exposure status in the past and then follow them forward to assess the outcomes that occurred in the exposed and nonexposed cohort members. Because in retrospective cohort studies the exposure and outcomes may have already occurred (it depends on how long they follow the cohort), it is important to make sure that the exposure preceded the outcome.

Sometimes cross-sectional studies are conducted (or cross-sectional analyses of cohort-study data), where the exposures and outcomes are measured during the same timeframe. As a result, cross-sectional analyses provide weaker evidence than regular cohort studies regarding a potential causal relationship between exposures and outcomes. For cross-sectional analyses, the answer to Question 6 should be "no."

Question 7. Sufficient timeframe to see an effect

Did the study allow enough time for a sufficient number of outcomes to occur or be observed, or enough time for an exposure to have a biological effect on an outcome? In the examples given above, if clinical depression has a biological effect on increasing risk for CVD, such an effect may take years. In the other example, if higher dietary sodium increases BP, a short timeframe may be sufficient to assess its association with BP, but a longer timeframe would be needed to examine its association with heart attacks.

The issue of timeframe is important to enable meaningful analysis of the relationships between exposures and outcomes to be conducted. This often requires at least several years, especially when looking at health outcomes, but it depends on the research question and outcomes being examined.

Cross-sectional analyses allow no time to see an effect, since the exposures and outcomes are assessed at the same time, so those would get a "no" response.

Question 8. Different levels of the exposure of interest

If the exposure can be defined as a range (examples: drug dosage, amount of physical activity, amount of sodium consumed), were multiple categories of that exposure assessed? (for example, for drugs: not on the medication, on a low dose, medium dose, high dose; for dietary sodium, higher than average U.S. consumption, lower than recommended consumption, between the two). Sometimes discrete categories of exposure are not used, but instead exposures are measured as continuous variables (for example, mg/day of dietary sodium or BP values).

In any case, studying different levels of exposure (where possible) enables investigators to assess trends or dose-response relationships between exposures and outcomes–e.g., the higher the exposure, the greater the rate of the health outcome. The presence of trends or dose-response relationships lends credibility to the hypothesis of causality between exposure and outcome.

For some exposures, however, this question may not be applicable (e.g., the exposure may be a dichotomous variable like living in a rural setting versus an urban setting, or vaccinated/not vaccinated with a one-time vaccine). If there are only two possible exposures (yes/no), then this question should be given an "NA," and it should not count negatively towards the quality rating.

Question 9. Exposure measures and assessment

Were the exposure measures defined in detail? Were the tools or methods used to measure exposure accurate and reliable–for example, have they been validated or are they objective? This issue is important as it influences confidence in the reported exposures. When exposures are measured with less accuracy or validity, it is harder to see an association between exposure and outcome even if one exists. Also as important is whether the expo-

sures were assessed in the same manner within groups and between groups; if not, bias may result.

For example, retrospective self-report of dietary salt intake is not as valid and reliable as prospectively using a standardized dietary log plus testing participants' urine for sodium content. Another example is measurement of BP, where there may be quite a difference between usual care, where clinicians measure BP however it is done in their practice setting (which can vary considerably), and use of trained BP assessors using standardized equipment (e.g., the same BP device which has been tested and calibrated) and a standardized protocol (e.g., patient is seated for 5 minutes with feet flat on the floor, BP is taken twice in each arm, and all four measurements are averaged). In each of these cases, the former would get a "no" and the latter a "yes."

Here is a final example that illustrates the point about why it is important to assess exposures consistently across all groups: If people with higher BP (exposed cohort) are seen by their providers more frequently than those without elevated BP (nonexposed group), it also increases the chances of detecting and documenting changes in health outcomes, including CVD-related events. Therefore, it may lead to the conclusion that higher BP leads to more CVD events. This may be true, but it could also be due to the fact that the subjects with higher BP were seen more often; thus, more CVD-related events were detected and documented simply because they had more encounters with the health care system. Thus, it could bias the results and lead to an erroneous conclusion.

Question 10. Repeated exposure assessment

Was the exposure for each person measured more than once during the course of the study period? Multiple measurements with the same result increase our confidence that the exposure status was correctly classified. Also, multiple measurements enable investigators to look at changes in exposure over time, for example, people who ate high dietary sodium throughout the followup period, compared to those who started out high then reduced their intake, compared to those who ate low sodium throughout. Once again, this may not be applicable in all cases. In many older studies, exposure was measured only at baseline. However, multiple exposure measurements do result in a stronger study design.

Question 11. Outcome measures

Were the outcomes defined in detail? Were the tools or methods for measuring outcomes accurate and reliable–for example, have they been validated or are they objective? This issue is important because it influences confidence in the validity of study results. Also important is whether the outcomes were assessed in the same manner within groups and between groups.

An example of an outcome measure that is objective, accurate, and reliable is death–the outcome measured with more accuracy than any other. But even with a measure as objective as death, there can be differences in the accuracy and reliability of how death was assessed by the investigators. Did they base it on an autopsy report, death certificate, death registry, or report from a family member? Another example is a study of whether dietary fat intake is related to blood cholesterol level (cholesterol level being the outcome), and the cholesterol level is measured from fasting blood samples that are all sent to the same laboratory. These examples would get a "yes." An example of a "no" would be self-report by subjects that they had a heart attack, or self-report of how much they weigh (if body weight is the outcome of interest).

Similar to the example in Question 9, results may be biased if one group (e.g., people with high BP) is seen more frequently than another group (people with normal BP) because more frequent encounters with the health care system increases the chances of outcomes being detected and documented.

Question 12. Blinding of outcome assessors

Blinding means that outcome assessors did not know whether the participant was exposed or unexposed. It is also sometimes called "masking." The objective is to look for evidence in the article that the person(s) assessing the outcome(s) for the study (for example, examining medical records to determine the outcomes that occurred in the exposed and comparison groups) is masked to the exposure status of the participant. Sometimes the person measuring the exposure is the same person conducting the outcome assessment. In this case, the outcome assessor would most likely not be blinded to exposure status because they also took measurements of exposures. If so, make a note of that in the comments section.

As you assess this criterion, think about whether it is likely that the person(s) doing the outcome assessment would know (or be able to figure out) the exposure status of the study participants. If the answer is no, then blinding is adequate. An example of adequate blind-

ing of the outcome assessors is to create a separate committee, whose members were not involved in the care of the patient and had no information about the study participants' exposure status. The committee would then be provided with copies of participants' medical records, which had been stripped of any potential exposure information or personally identifiable information. The committee would then review the records for prespecified outcomes according to the study protocol. If blinding was not possible, which is sometimes the case, mark "NA" and explain the potential for bias.

Question 13. Followup rate

Higher overall followup rates are always better than lower followup rates, even though higher rates are expected in shorter studies, whereas lower overall followup rates are often seen in studies of longer duration. Usually, an acceptable overall followup rate is considered 80 percent or more of participants whose exposures were measured at baseline. However, this is just a general guideline. For example, a 6-month cohort study examining the relationship between dietary sodium intake and BP level may have over 90 percent followup, but a 20-year cohort study examining effects of sodium intake on stroke may have only a 65 percent followup rate.

Question 14. Statistical analyses

Were key potential confounding variables measured and adjusted for, such as by statistical adjustment for baseline differences? Logistic regression or other regression methods are often used to account for the influence of variables not of interest.

This is a key issue in cohort studies, because statistical analyses need to control for potential confounders, in contrast to an RCT, where the randomization process controls for potential confounders. All key factors that may be associated both with the exposure of interest and the outcome–that are not of interest to the research question–should be controlled for in the analyses.

For example, in a study of the relationship between cardiorespiratory fitness and CVD events (heart attacks and strokes), the study should control for age, BP, blood cholesterol, and body weight, because all of these factors are associated both with low fitness and with CVD events. Well-done cohort studies control for multiple potential confounders.

Some general guidance for determining the overall quality rating of observational cohort and cross-sectional studies

The questions on the form are designed to help you focus on the key concepts for evaluating the internal validity of a study. They are not intended to create a list that you simply tally up to arrive at a summary judgment of quality.

Internal validity for cohort studies is the extent to which the results reported in the study can truly be attributed to the exposure being evaluated and not to flaws in the design or conduct of the study–in other words, the ability of the study to draw associative conclusions about the effects of the exposures being studied on outcomes. Any such flaws can increase the risk of bias.

Critical appraisal involves considering the risk of potential for selection bias, information bias, measurement bias, or confounding (the mixture of exposures that one cannot tease out from each other). Examples of confounding include co-interventions, differences at baseline in patient characteristics, and other issues throughout the questions above. High risk of bias translates to a rating of poor quality. Low risk of bias translates to a rating of good quality. (Thus, the greater the risk of bias, the lower the quality rating of the study.)

In addition, the more attention in the study design to issues that can help determine whether there is a causal relationship between the exposure and outcome, the higher quality the study. These include exposures occurring prior to outcomes, evaluation of a dose-response gradient, accuracy of measurement of both exposure and outcome, sufficient timeframe to see an effect, and appropriate control for confounding–all concepts reflected in the tool.

Generally, when you evaluate a study, you will not see a "fatal flaw," but you will find some risk of bias. By focusing on the concepts underlying the questions in the quality assessment tool, you should ask yourself about the potential for bias in the study you are critically appraising. For any box where you check "no" you should ask, "What is the potential risk of bias resulting from this flaw in study design or execution?" That is, does this factor cause you to doubt the results that are reported in the study or doubt the ability of the study to accurately assess an association between exposure and outcome?

The best approach is to think about the questions in the tool and how each one tells you something about the potential for bias in a study. The more you familiarize yourself with the key concepts, the more comfortable you will be with critical appraisal. Examples of studies rated good, fair, and poor are useful, but each study must be assessed on its own

based on the details that are reported and consideration of the concepts for minimizing bias.

Last Updated March 2014

j) Quality Assessment of Case-Control Studies

Criteria	Yes	No	Other (CD, NR, NA)*
1. Was the research question or objective in this paper clearly stated and appropriate?			
2. Was the study population clearly specified and defined?			
3. Did the authors include a sample size justification?			
4. Were controls selected or recruited from the same or similar population that gave rise to the cases (including the same timeframe)?			
5. Were the definitions, inclusion and exclusion criteria, algorithms or processes used to identify or select cases and controls valid, reliable, and implemented consistently across all study participants?			
6. Were the cases clearly defined and differentiated from controls?			
7. If less than 100 percent of eligible cases and/or controls were selected for the study, were the cases and/or controls randomly selected from those eligible?			
8. Was there use of concurrent controls?			
9. Were the investigators able to confirm that the exposure/risk occurred prior to the development of the condition or event that defined a participant as a case?			
10. Were the measures of exposure/risk clearly defined, valid, reliable, and implemented consistently (includingthe same time period) across all study participants?			
11. Were the assessors of exposure/risk blinded to the case or			

control status of participants?			
12. Were key potential confounding variables measured and adjusted statistically in the analyses? If matching was used, did the investigators account for matching during study analysis?			

Quality Rating (Good, Fair, or Poor) (see guidance)
Rater #1 initials:
Rater #2 initials:
Additional Comments (If POOR, please state why):

*CD, cannot determine; NA, not applicable; NR, not reported

Guidance for Assessing the Quality of Case-Control Studies

The guidance document below is organized by question number from the tool for quality assessment of case-control studies.

Question 1. Research question

Did the authors describe their goal in conducting this research? Is it easy to understand what they were looking to find? This issue is important for any scientific paper of any type. High quality scientific research explicitly defines a research question.

Question 2. Study population

Did the authors describe the group of individuals from which the cases and controls were selected or recruited, while using demographics, location, and time period? If the investigators conducted this study again, would they know exactly who to recruit, from where, and from what time period?

Investigators identify case-control study populations by location, time period, and inclusion criteria for cases (individuals with the disease, condition, or problem) and controls (individuals without the disease, condition, or problem). For example, the population for a study of lung cancer and chemical exposure would be all incident cases of lung cancer diagnosed in patients ages 35 to 79, from January 1, 2003 to December 31, 2008, living in Texas during that entire time period, as well as controls without lung cancer recruited from the same population during the same time period. The population is clearly described as: (1) who (men and women ages 35 to 79 with (cases) and without (controls) incident lung can-

cer); (2) where (living in Texas); and (3) when (between January 1, 2003 and December 31, 2008).

Other studies may use disease registries or data from cohort studies to identify cases. In these cases, the populations are individuals who live in the area covered by the disease registry or included in a cohort study (i.e., nested case-control or case-cohort). For example, a study of the relationship between vitamin D intake and myocardial infarction might use patients identified via the GRACE registry, a database of heart attack patients.

NHLBI staff encouraged reviewers to examine prior papers on methods (listed in the reference list) to make this assessment, if necessary.

Question 3. Target population and case representation

In order for a study to truly address the research question, the target population–the population from which the study population is drawn and to which study results are believed to apply–should be carefully defined. Some authors may compare characteristics of the study cases to characteristics of cases in the target population, either in text or in a table. When study cases are shown to be representative of cases in the appropriate target population, it increases the likelihood that the study was well-designed per the research question.

However, because these statistics are frequently difficult or impossible to measure, publications should not be penalized if case representation is not shown. For most papers, the response to question 3 will be "NR." Those subquestions are combined because the answer to the second subquestion–case representation–determines the response to this item. However, it cannot be determined without considering the response to the first subquestion. For example, if the answer to the first subquestion is "yes," and the second, "CD," then the response for item 3 is "CD."

Question 4. Sample size justification

Did the authors discuss their reasons for selecting or recruiting the number of individuals included? Did they discuss the statistical power of the study and provide a sample size calculation to ensure that the study is adequately powered to detect an association (if one exists)? This question does not refer to a description of the manner in which different groups were included or excluded using the inclusion/exclusion criteria (e.g., "Final study size was 1,378 participants after exclusion of 461 patients with missing data" is not considered a sample size justification for the purposes of this question).

An article's methods section usually contains information on sample size and the size needed to detect differences in exposures and on statistical power.

Question 5. Groups recruited from the same population

To determine whether cases and controls were recruited from the same population, one can ask hypothetically, "If a control was to develop the outcome of interest (the condition that was used to select cases), would that person have been eligible to become a case?" Case-control studies begin with the selection of the cases (those with the outcome of interest, e.g., lung cancer) and controls (those in whom the outcome is absent). Cases and controls are then evaluated and categorized by their exposure status. For the lung cancer example, cases and controls were recruited from hospitals in a given region. One may reasonably assume that controls in the catchment area for the hospitals, or those already in the hospitals for a different reason, would attend those hospitals if they became a case; therefore, the controls are drawn from the same population as the cases. If the controls were recruited or selected from a different region (e.g., a State other than Texas) or time period (e.g., 1991-2000), then the cases and controls were recruited from different populations, and the answer to this question would be "no."

The following example further explores selection of controls. In a study, eligible cases were men and women, ages 18 to 39, who were diagnosed with atherosclerosis at hospitals in Perth, Australia, between July 1, 2000 and December 31, 2007. Appropriate controls for these cases might be sampled using voter registration information for men and women ages 18 to 39, living in Perth (population-based controls); they also could be sampled from patients without atherosclerosis at the same hospitals (hospital-based controls). As long as the controls are individuals who would have been eligible to be included in the study as cases (if they had been diagnosed with atherosclerosis), then the controls were selected appropriately from the same source population as cases.

In a prospective case-control study, investigators may enroll individuals as cases at the time they are found to have the outcome of interest; the number of cases usually increases as time progresses. At this same time, they may recruit or select controls from the population without the outcome of interest. One way to identify or recruit cases is through a surveillance system. In turn, investigators can select controls from the population covered by that system. This is an example of population-based controls. Investigators also may identify and select cases from a cohort study population and identify controls from outcome-free individuals in the same cohort study. This is known as a nested case-control study.

Question 6. Inclusion and exclusion criteria prespecified and applied uniformly

Were the inclusion and exclusion criteria developed prior to recruitment or selection of the study population? Were the same underlying criteria used for all of the groups involved?

To answer this question, reviewers determined if the investigators developed I/E criteria prior to recruitment or selection of the study population and if they used the same underlying criteria for all groups. The investigators should have used the same selection criteria, except for study participants who had the disease or condition, which would be different for cases and controls by definition. Therefore, the investigators use the same age (or age range), gender, race, and other characteristics to select cases and controls. Information on this topic is usually found in a paper's section on the description of the study population.

Question 7. Case and control definitions

For this question, reviewers looked for descriptions of the validity of case and control definitions and processes or tools used to identify study participants as such. Was a specific description of "case" and "control" provided? Is there a discussion of the validity of the case and control definitions and the processes or tools used to identify study participants as such? They determined if the tools or methods were accurate, reliable, and objective. For example, cases might be identified as "adult patients admitted to a VA hospital from January 1, 2000 to December 31, 2009, with an ICD-9 discharge diagnosis code of acute myocardial infarction and at least one of the two confirmatory findings in their medical records: at least 2mm of ST elevation changes in two or more ECG leads and an elevated troponin level. Investigators might also use ICD-9 or CPT codes to identify patients. All cases should be identified using the same methods. Unless the distinction between cases and controls is accurate and reliable, investigators cannot use study results to draw valid conclusions.

Question 8. Random selection of study participants

If a case-control study did not use 100 percent of eligible cases and/or controls (e.g., not all disease-free participants were included as controls), did the authors indicate that random sampling was used to select controls? When it is possible to identify the source population fairly explicitly (e.g., in a nested case-control study, or in a registry-based study), then random sampling of controls is preferred. When investigators used consecutive sampling, which is frequently done for cases in prospective studies, then study participants are not considered randomly selected. In this case, the reviewers would answer "no" to Question 8. However, this would not be considered a fatal flaw.

If investigators included all eligible cases and controls as study participants, then reviewers marked "NA" in the tool. If 100 percent of cases were included (e.g., NA for cases) but only 50 percent of eligible controls, then the response would be "yes" if the controls were randomly selected, and "no" if they were not. If this cannot be determined, the appropriate response is "CD."

Question 9. Concurrent controls

A concurrent control is a control selected at the time another person became a case, usually on the same day. This means that one or more controls are recruited or selected from the population without the outcome of interest at the time a case is diagnosed. Investigators can use this method in both prospective case-control studies and retrospective case-control studies. For example, in a retrospective study of adenocarcinoma of the colon using data from hospital records, if hospital records indicate that Person A was diagnosed with adenocarcinoma of the colon on June 22, 2002, then investigators would select one or more controls from the population of patients without adenocarcinoma of the colon on that same day. This assumes they conducted the study retrospectively, using data from hospital records. The investigators could have also conducted this study using patient records from a cohort study, in which case it would be a nested case-control study.

Investigators can use concurrent controls in the presence or absence of matching and vice versa. A study that uses matching does not necessarily mean that concurrent controls were used.

Question 10. Exposure assessed prior to outcome measurement

Investigators first determine case or control status (based on presence or absence of outcome of interest), and then assess exposure history of the case or control; therefore, reviewers ascertained that the exposure preceded the outcome. For example, if the investigators used tissue samples to determine exposure, did they collect them from patients prior to their diagnosis? If hospital records were used, did investigators verify that the date a patient was exposed (e.g., received medication for atherosclerosis) occurred prior to the date they became a case (e.g., was diagnosed with type 2 diabetes)? For an association between an exposure and an outcome to be considered causal, the exposure must have occurred prior to the outcome.

Question 11. Exposure measures and assessment

Were the exposure measures defined in detail? Were the tools or methods used to measure exposure accurate and reliable–for example, have they been validated or are they objective? This is important, as it influences confidence in the reported exposures. Equally important is whether the exposures were assessed in the same manner within groups and between groups. This question pertains to bias resulting from exposure misclassification (i.e., exposure ascertainment).

For example, a retrospective self-report of dietary salt intake is not as valid and reliable as prospectively using a standardized dietary log plus testing participants' urine for sodium content because participants' retrospective recall of dietary salt intake may be inaccurate and result in misclassification of exposure status. Similarly, BP results from practices that use an established protocol for measuring BP would be considered more valid and reliable than results from practices that did not use standard protocols. A protocol may include using trained BP assessors, standardized equipment (e.g., the same BP device which has been tested and calibrated), and a standardized procedure (e.g., patient is seated for 5 minutes with feet flat on the floor, BP is taken twice in each arm, and all four measurements are averaged).

Question 12. Blinding of exposure assessors

Blinding or masking means that outcome assessors did not know whether participants were exposed or unexposed. To answer this question, reviewers examined articles for evidence that the outcome assessor(s) was masked to the exposure status of the research participants. An outcome assessor, for example, may examine medical records to determine the outcomes that occurred in the exposed and comparison groups. Sometimes the person measuring the exposure is the same person conducting the outcome assessment. In this case, the outcome assessor would most likely not be blinded to exposure status. A reviewer would note such a finding in the comments section of the assessment tool.

One way to ensure good blinding of exposure assessment is to have a separate committee, whose members have no information about the study participants' status as cases or controls, review research participants' records. To help answer the question above, reviewers determined if it was likely that the outcome assessor knew whether the study participant was a case or control. If it was unlikely, then the reviewers marked "no" to Question 12. Outcome assessors who used medical records to assess exposure should not have been directly involved in the study participants' care, since they probably would have known about their patients' conditions. If the medical records contained information on the patient's condition that identified him/her as a case (which is likely), that information would have had to be removed before the exposure assessors reviewed the records.

If blinding was not possible, which sometimes happens, the reviewers marked "NA" in the assessment tool and explained the potential for bias.

Question 13. Statistical analysis

Were key potential confounding variables measured and adjusted for, such as by statistical adjustment for baseline differences? Investigators often use logistic regression or other regression methods to account for the influence of variables not of interest.

This is a key issue in case-controlled studies; statistical analyses need to control for potential confounders, in contrast to RCTs in which the randomization process controls for potential confounders. In the analysis, investigators need to control for all key factors that may be associated with both the exposure of interest and the outcome and are not of interest to the research question.

A study of the relationship between smoking and CVD events illustrates this point. Such a study needs to control for age, gender, and body weight; all are associated with smoking and CVD events. Well-done case-control studies control for multiple potential confounders.

Matching is a technique used to improve study efficiency and control for known confounders. For example, in the study of smoking and CVD events, an investigator might identify cases that have had a heart attack or stroke and then select controls of similar age, gender, and body weight to the cases. For case-control studies, it is important that if matching was performed during the selection or recruitment process, the variables used as matching criteria (e.g., age, gender, race) should be controlled for in the analysis.

General Guidance for Determining the Overall Quality Rating of Case-Controlled Studies

NHLBI designed the questions in the assessment tool to help reviewers focus on the key concepts for evaluating a study's internal validity, not to use as a list from which to add up items to judge a study's quality.

Internal validity for case-control studies is the extent to which the associations between disease and exposure reported in the study can truly be attributed to the exposure being evaluated rather than to flaws in the design or conduct of the study. In other words, what is ability of the study to draw associative conclusions about the effects of the exposures on outcomes? Any such flaws can increase the risk of bias.

In critical appraising a study, the following factors need to be considered: risk of potential for selection bias, information bias, measurement bias, or confounding (the mixture of exposures that one cannot tease out from each other). Examples of confounding include co-interventions, differences at baseline in patient characteristics, and other issues addressed in the questions above. High risk of bias translates to a poor quality rating; low risk of bias

translates to a good quality rating. Again, the greater the risk of bias, the lower the quality rating of the study.

In addition, the more attention in the study design to issues that can help determine whether there is a causal relationship between the outcome and the exposure, the higher the quality of the study. These include exposures occurring prior to outcomes, evaluation of a dose-response gradient, accuracy of measurement of both exposure and outcome, sufficient timeframe to see an effect, and appropriate control for confounding–all concepts reflected in the tool.

If a study has a "fatal flaw," then risk of bias is significant; therefore, the study is deemed to be of poor quality. An example of a fatal flaw in case-control studies is a lack of a consistent standard process used to identify cases and controls.

Generally, when reviewers evaluated a study, they did not see a "fatal flaw," but instead found some risk of bias. By focusing on the concepts underlying the questions in the quality assessment tool, reviewers examined the potential for bias in the study. For any box checked "no," reviewers asked, "What is the potential risk of bias resulting from this flaw in study design or execution?" That is, did this factor lead to doubt about the results reported in the study or the ability of the study to accurately assess an association between exposure and outcome?

By examining questions in the assessment tool, reviewers were best able to assess the potential for bias in a study. Specific rules were not useful, as each study had specific nuances. In addition, being familiar with the key concepts helped reviewers assess the studies. Examples of studies rated good, fair, and poor were useful, yet each study had to be assessed on its own.

Last Updated March 2014

k) NEWCASTLE - OTTAWA QUALITY ASSESSMENT SCALE CASE CONTROL STUDIES

Note: A study can be awarded a maximum of one star for each numbered item within the Selection and Exposure categories. A maximum of two stars can be given for Comparability.

Selection

1) Is the case definition adequate?
   a) yes, with independent validation ☐
   b) yes, eg record linkage or based on self reports
   c) no description

2) Representativeness of the cases
   a) consecutive or obviously representative series of cases ☐
   b) potential for selection biases or not stated

3) Selection of Controls
   a) community controls ☐
   b) hospital controls
   c) no description

4) Definition of Controls
   a) no history of disease (endpoint) ☐
   b) no description of source

Comparability

1) Comparability of cases and controls on the basis of the design or analysis
   a) study controls for _____ (Select the most important factor.) ☐
   b) study controls for any additional factor ☐ (This criteria could be modified to indicate specific control for a second important factor.)

Exposure

1) Ascertainment of exposure
   a) secure record (eg surgical records) ☐
   b) structured interview where blind to case/control status ☐
   c) interview not blinded to case/control status
   d) written self report or medical record only
   e) no description

2) Same method of ascertainment for cases and controls
   a) yes ☐
   b) no

3) Non-Response rate
   a) same rate for both groups ☐
   b) non respondents described

c) rate different and no designation

l) NEWCASTLE - OTTAWA QUALITY ASSESSMENT SCALE COHORT STUDIES

Note: A study can be awarded a maximum of one star for each numbered item within the Selection and Outcome categories. A maximum of two stars can be given for Comparability

Selection

1) Representativeness of the exposed cohort
   a) truly representative of the average _____ (describe) in the community ⍰
   b) somewhat representative of the average _____ in the community ⍰
   c) selected group of users eg nurses, volunteers
   d) no description of the derivation of the cohort

2) Selection of the non exposed cohort
   a) drawn from the same community as the exposed cohort ⍰
   b) drawn from a different source
   c) no description of the derivation of the non exposed cohort

3) Ascertainment of exposure
   a) secure record (eg surgical records) ⍰
   b) structured interview ⍰
   c) written self report
   d) no description

4) Demonstration that outcome of interest was not present at start of study
   a) yes ⍰
   b) no

Comparability

1) Comparability of cohorts on the basis of the design or analysis
   a) study controls for _____ (select the most important factor) ⍰
   b) study controls for any additional factor ⍰ (This criteria could be modified to indicate specific control for a second important factor.)

Outcome

1) Assessment of outcome
   a) independent blind assessment ⍰
   b) record linkage ⍰
   c) self report

d) no description

2) Was follow-up long enough for outcomes to occur
a) yes (select an adequate follow up period for outcome of interest)
b) no

3) Adequacy of follow up of cohorts
a) complete follow up - all subjects accounted for
b) subjects lost to follow up unlikely to introduce bias - small number lost - > ____ % (select an adequate %) follow up, or description provided of those lost)
c) follow up rate < ____% (select an adequate %) and no description of those lost
d) no statement

CODING MANUAL FOR CASE-CONTROL STUDIES

SELECTION

1) Is the Case Definition Adequate?

a) Requires some independent validation (e.g. >1 person/record/time/process to extract information, or reference to primary record source such as x-rays or medical/hospital records) ☆
b) Record linkage (e.g. ICD codes in database) or self-report with no reference to primary record
c) No description

2) Representativeness of the Cases

a) All eligible cases with outcome of interest over a defined period of time, all cases in a defined catchment area, all cases in a defined hospital or clinic, group of hospitals, health maintenance organisation, or an appropriate sample of those cases (e.g. random sample) ☆
b) Not satisfying requirements in part (a), or not stated.

3) Selection of Controls

This item assesses whether the control series used in the study is derived from the same population as the cases and essentially would have been cases had the outcome been present.

a) Community controls (i.e. same community as cases and would be cases if had outcome) ☆
b) Hospital controls, within same community as cases (i.e. not another city) but derived from a hospitalised population
c) No description

4) Definition of Controls

a) If cases are first occurrence of outcome, then it must explicitly state that controls have no history of this outcome. If cases have new (not necessarily first) occurrence of outcome, then controls with previous occurrences of outcome of interest should not be excluded. ☆
b) No mention of history of outcome

COMPARABILITY

1) **Comparability of Cases and Controls on the Basis of the Design or Analysis**

A maximum of 2 stars can be allotted in this category

Either cases and controls must be matched in the design and/or confounders must be adjusted for in the analysis. Statements of no differences between groups or that differences were not statistically significant are not sufficient for establishing comparability. Note: If the odds ratio for the exposure of interest is adjusted for the confounders listed, then the groups will be considered to be comparable on each variable used in the adjustment. There may be multiple ratings for this item for different categories of exposure (e.g. ever vs. never, current vs. previous or never)
Age =☆ , Other controlled factors =☆

EXPOSURE

1) Ascertainment of Exposure

Allocation of stars as per rating sheet

2) Non-Response Rate

Allocation of stars as per rating sheet

m) References

US-Department of Health and Human Services, National Heart, Lung and Blood Institute. Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies (2014). URL: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp (Zugriff: 04.07.2016)

US-Department of Health and Human Services, National Heart, Lung and Blood Institute. Quality Assessment of Case-Control Studies (2014).
URL: https://www.nhlbi.nih.gov/health-pro/guidelines/in-develop/cardiovascular-risk-reduction/tools/cohort (Zugriff: 04.07.2016)

Wells et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. The Ottawa Hospital. Research Institute.
URL**:** https://www.nhlbi.nih.gov/health-pro/guidelines/in-develop/cardiovascular-risk-reduction/tools/case-control (Zugriff: 04.07.2016)

9 Annex II

Table A1a: Primary research articles presenting methods and results of UFP/ Quasi-UFP epidemiologic short-term studies, mortality

Reference	Country, City	Study period	Study Design	Sample Size, Main study population	Exposure Assessment	Size Fractions	Tech. Device	Covariate adjustment	Outcome Assess-ment	Outcome	Exposure time windows	Effect sizes (confidence intervals) per increment
Time-series												
Lanzinger et al. (2016) UFIREG study	4 Cities in Germany, Czech Republic, Slovenia, Ukraine,	01/2011-03/2014, city-specific times overlapping	Time-series	2,582,000 General population >1 year	Measurement: Central site	PNC 20-100 nm (UFP) PNC 20-800 nm	Differential or Scanning MPS	time-trend, DOW, public holidays, vacation periods, influenza periods, T, RH	Official registry	Natural mortality	Lag days ma 0-1, ma 2-5 ma 0-5	percent cahanges in RRs/ PNC20-100 per 2,750/ml ma 2-5: −1.2 (−4.0; 1.8) ma 0-1: 0.1 (−2.0; 2.4) RRs/ PNC20-800 per 3,675/ml ma 2-5: −1.2 (−4.1; 1.8) ma 0-1: −0.2 (−2.4; 2.1)
									Official registry	Cv mortality	ma 0-1 ma 2-5 ma 0-5	RRs/ PNC20-100 per 2,750/ml ma 0-1: −0.5 (−3.6; 2.8) ma 0-5: −0.2 (−5.5; 5.4) RRs/ PNC20-800 per 3,675/ml ma 0-1: −0.7 (−3.9; 2.5) ma 0-5: −0.1 (−5.8; 5.9)
									Official registry	Resp. mortality	ma 0-1 ma 2 -5 ma 0-5	RRs/ PNC20-100 per 2,750/ml ma 0-1: 3.7 (−5.8; 14.2) ma 0-5: 9.9 (-6.3; 28.8) RRs/ PNC20-800 per 3,675/ml ma 0-1: 1.5 (−8.0; 12.1) ma 0-5: 5.6 (−8.3; 21.7)
Leitte et al. (2012)	China, Beijing	03/2004-08/2005	Time-series	8,000,000 Beijing residents, for respiratory disease adults > 20 yrs	Measurement: Central site	PNC3–10 PNC10–30 PNC30–50 PNC50–100 PNC100–300 PNC300–1000 PNC3 –1 µm (NCtot) 3–100 nm	TDMPS and TSI	Seasonal pattern, T, DOW	Official registry	Resp. mortality	Lag days lag 0, lag 1 lag 2 ma 0-3 ma 0-4	Percentage change/ PNC300–1000 per 840/ml lag 1: 2.1 (-3.0; 7.5) lag 2: 0.7 (-3.8; 5.3) ma 0-4:.11.5 (3.0; 20.7) PNC 3-100 per 13,000/ml lag 1: -3.1 (-9.5; 3.9) ma 0-4: 3.9 (-7.3; 16.4) PNC total per 14,000/ml

Reference	Country, City	Study period	Study Design	Sample Size, Main study population	Exposure Assessment	Size Fractions	Tech. Device	Covariate adjustment	Outcome Assess-ment	Outcome	Exposure time windows	Effect sizes (confidence intervals) per increment
						(UFP)						lag 1: 0.3 (-7.5; 8.7) lag 2: 9.3 (1.3; 17.9)
Meng et al. (2013)	China, Chenyang	12/2006-11/2008	Time-series	NR/ total population General population	Measure-ment: Cen-tral site	PNC250–280, PNC280–300, PNC300–350, PNC350–400, PNC400–450, PNC450–500, PNC500–650, PNC650–1000	Ambient Dust Monitor 365 (GRIMM)	Calendar time, current day-mean T, RH, DOW	Adm. Database	Total mortality	ma 0-1	Percent change, All periods, ma 0-1 0.12 (-0.22; 0.45) per 63/ml PNC650-1000 2.41 (1.23; 3.58) per 2,600/ml PNC250–280, warm period, 2.11 (0.72; 3.49) per 193/ml PNC450–500 to 4.21 (2.43; 5.99) per 2,600/ml PNC250–280
									Adm. Database	cv mortal-ity	ma 0-1	All periods, range 0.37 (–0.10; 0.84) per 63/ml PNC650-1000 2.79 (1.09; 4.49) per 2,600/ml PNC250–280
									Adm. Database	resp, mortality	ma 0-1	All periods, range 0.42 (–0.59; 1.43) per 63/ml PNC650-1000 0.81 (–2.33; 3.96) per 1,510/ml PNC300–350
Samoli et al (2016a) Clearflo	UK, London	01/2011-12/2012	Time-Series	approxi-mately 9 million (>700/day)	Measure-ment: Cen-tral site	PNC> 6nm	CPC model 3022	trend, DOW, public holidays, T, RH	Medical records	Total non-accid. Mortality ICD-10 Chapters A–R	lag 1d	Percent changes per 5,180/ml −0.06 (−1.16; 1.06)

Reference	Country, City	Study period	Study Design	Sample Size, Main study population	Exposure Assessment	Size Fractions	Tech. Device	Covariate adjustment	Outcome Assess-ment	Outcome	Exposure time windows	Effect sizes (confidence intervals) per increment
									Medical records	cv mortal-ity	lag 1d	−2.04 (−3.94; −0.10)
									Medical records	resp, mortality	lag 2d	−1.86 (−4.50; 0.86)
Stafoggia et al. (2017) UF& HEALTH Study	8 Cities/ Areas in Finland, Sweden, Denmark, Germany, Italy, *Spain*, Greece	01/1999-12/2013	Time-series	12,000,000 General population	Measure-ment: Cen-tral site	Athens, Co-penhagen, Helsinki: 0-100 nm, Barcelone: 5-1,000 nm, Ruhr Area: 14-750 nm, Augs-burg: 7-3,000/ 10-2,000, Stockholm: 4-3,000/ 7-3,000	various (eTable 1; http://li nks.lww .com/ED E/B142	longterm and seasonal time trends, DOW, population dynamics due to summer vacation and holidays, influ-enza peaks, T	Official registry	Non-accid. ICD-10 codes: 1-799	lag 0-10 shown in table: lags 5-7, figure 1: all lags	Percent increases PNC per 10,000/ml lag 5: 0.32 (-0.08; 0.72) lag 6: 0.35 (-0.05; 0.75) lag 7: 0.37 (-0.03; 0.78) lags 0-4, 8-9, range: 0.00--0.35, lag 10 similar to lag 7
										Cv mortal-ity	lag 0-10	range: -0.58 (lag 0/ lag 9)to 0.45 (lag 7) no estimate significant.
										resp. mortality	lag 0-10	range: -0.6 (lag 1, lag 0 similar) to 0.65 (lag 6, lag 10 similar) significant protective estimate at lag 3 (estimate not visible in figure)
Su et al. (2015)	China, Beijing	05/2008-12/2008	Time-series	12,299,000 General population	Measure-ment: Cen-tral site	PNC3-10nm PNC10-30nm PNC30-50nm PNC50-100nm PNC3-100nm (UFP)	TDMPS	T, RH, DoW, public holidays, three specific periods, heat-ing period, season.	Official registry	Cv mortal-ity ICD10: I00–I99	Lag days lag 0, lag 1, ma 5	Percent increase per 1,758/ml PN 30-50: lag 0: 2.3 (-2.1; 6.8), lag 1: 6.0 (1.7; 10.6), ma 5: 7.4 (2.1; 12.9) per 8,328/ml PN3-100:

Reference	Country, City	Study period	Study Design	Sample Size, Main study population	Exposure Assessment	Size Fractions	Tech. Device	Covariate adjustment	Outcome Assess-ment	Outcome	Exposure time windows	Effect sizes (confidence intervals) per increment
												lag 0: 3.7 (-1.5; 9.1), lag 1: 5.7 (0.8; 10.7), ma 5: 8.8 (2.7; 15.2)
										IHD: ICD10: I20–I25,	lag 0, lag 1, ma 5	Percent increase Per 1,304/ml PN30-50: lag 0: 3.4 (-3.9; 9.2), lag 1:-4.0 (-4.0; 8.8), ma 5: 5.7 (-1.9; 14.0) per 8,328/ml PN3-100: lag 0: 2.7 (-4.7; 10.7), lag 1: -0.7 (-7.4; 6.5), ma 5: 4.4 (-4.2; 13.8),
										Cerebro-vascular: ICD10: I60–I69	lag 0, lag 1, ma 5	Percent increase Per 3,502/ml PN30-50: lag 0: 3.3 (-3.5; 10.7), lag 1: 10.3 (3.3; 17.8), ma 5: 7.5 (-0.8; 16.5) per 8,328/ml PN3-100 lag 0: 8.0 (0.4; 17.0) lag 2: 13.6 (5.7; 22.1) ma 5: 13.3 (3.4; 24.2)
Wolf et al. 2015	Germany, Augsburg	1999-2009	Time-series	15,417 General population	Measure-ment: Cen-tral site	PNC10-2000 nm	CPC, TDMPS	time trend, temperature, season, day of week	Official registry	MI and coronary deaths, fatal events	Lag days lag 0 lag 1 ma 5	percent change in RR per 6,800/ml (PNCm+f) lag 0: 1.3 (−2.0; 4.7) lag 1: 0.5 (−2.8; 4.0) ma5: -0.5 (−4.2; 3.3)

[a] IHD: Ischaemic heart disease, MI: Myocardial infarction, ICD: International Classification of disease, cv: cardiovascular, T: Temperature, RH: Relative humidity, DOW: Day of week, PNC: Particulate number concentration, ma: mean average, CPC: Condensation particle counter.

[b] TDMPS: Twin Differential Mobility Particle Sizer.

Table A1b: Primary research articles presenting methods and results of UFP/ quasi-UFP epidemiologic short-term studies, morbidity

Reference	Country, City	Study period	Study Design	Sample Size, Main study population	Exposure Assessment	Size Fractions	Techn. device	Covariate adjustment	Outcome	Outcome Assessment	Exposure time windows	Effect sizes (confidence intervals) per increment
Case-crossover												
Cole-Hunter et al. (2013)	**Australia, Brisbane**	**Not reported/ no reference given**	**Case-crossover**	**35, healthy cycling adults, Mean age: 39**	**Measurement: Mobile**	**PNC <100nm PD 1-300nm**	**Aerasense NanoTracer**	**NA**	**Nose irritation throat irritation Other symptoms, e.g. Cough peak flow rates Blood cell counts, e.g. Leukoc.**	**Self-reported**	**-**	**Mean ± SD high vs. low inbound exposure: Nose irritation 1.82 ± 0.33 versus 1.53 ± 0.23 Throat irritation 2.00 ± 0.40 vs.1.56 ± 0.24 Cough 1.62 ± 0.26 vs. 1.26 ± 0.16 Peak flow rates 1.28 ± 0.16 vs. 1.76 ± 0.31 Leucocytes 1.38 ± 0.43 vs. 1.37 ± 0.42**
Link et al. (2013)	USA, Boston (Massachussets)	09/2006 - 06/2010	Case-crossover	176, adults >18 yrs with prior implantation of dual (atrial + ventricular) chamber ICD	Measurement: Central site	Total	CPC	T, dew point	Events of atrial fibrillation	Other	ma 0-2h ma 0-6h ma 0-12h ma 0-24h ma 0-48h	ma 0-2h: 24% (-4%; 61%) per 10,900/ml ma 0-24: 12% (-19%; 56%) per 8,400/ml

Reference	Country, City	Study period	Study Design	Sample Size, Main study population	Exposure Assessment	Size Fractions	Techn. device	Covariate adjustment	Outcome	Outcome Assessment	Exposure time windows	Effect sizes (confidence intervals) per increment
Cohort												
Mehta et al. (2015) Veterans Affairs Normative Aging Study	USA, Boston (Massachussets)	1995-2007	Cohort	987, elderly men	Measurement: Central site	Total	CPC 3022A	Age, education, race, physical activity, seasonality, DoW, T, anti-depressant medic.	Perceived stress during previous week	Standardized-clinical examinations	ma 1 week	point increase per 15,997/ml PNC 3.2 (2.1; 4.3)
Wang et al. (2014) MOBILIZE Boston study	USA, Boston (Massachussets)	2005-2010	Cohort	1,314 baseline. and 732 follow-up, adults, ≥ 65 yrs, mean age: 78 yrs	Measurement: Central site	NR	NR	Age, sex, race/ethnicity, visit, dew point, T, barom. pressure, DOW, season, long-term temporal trends	CESD-R ≥ 16	Standardized Interview	ma 1, 2, 3, 5, 7, 14 days	OR per 6,630/ml PNC 1.04 (0.68; 1.57)
Panel (repeated measure)												
Karakatsani (2012)	The Netherlands, Amsterdam; Greece, Athens; UK, Birmingham; Finland, Helsinki	10/2002-03/2004	Panel (repeated measure)	136, adults ≥ 35 yrs, either asthmatic or COPD patient	Measurement: Central site	Total	CPC 3022A, TSI	Time, T, RH, DOW, medication use, individual differences in frequency of symptoms	Woken with breathing problems, Shortness of breath, Wheeze, Cough, Phlegm,	Self-reported	lag 0 lag 1 lag 2 ma0-6	ORs for total/asthmatic population per 10,000/ml Woken with breathing problems: lag 0: 0.97 (0.87; 1.0)/1.01 (0.84; 1.2) lag 1: 1.03 (0.952; 1.11)/ 1.05 (0.96; 1.14) lag 2: 0.96 (0.86; 1.06)/ 1.02 (0.94; 1.11) ma0-6: 0.910 (0.64; 1.30)/ 1.20 (0.95; 1.50)

Reference	Country, City	Study period	Study Design	Sample Size, Main study population	Exposure Assessment	Size Frac-tions	Techn. device	Covariate adjustment	Outcome	Outcome Assess-ment	Exposure time win-dows	Effect sizes (confidence intervals) per increment
									Limitation of vigorous activities, Limitation of moderate activities, limitation of walking			Shortness of breath: lag 0: 0.97 (0.9; 1.05)/0.98 (0.9; 1.06) lag 1: 0.91 (0.84; 0.98)/ 0.93 (0.82; 1.05) lag 2: 0.92 (0.86; 0.98)/0.95 (0.88; 10.3) ma 0-6: 0.91 (0.77; 1.07)/1.03 (0.86; 1.24) Wheezing: lag 0: 0.93 (0.79; 1.1)/0.96 (0.82; 1.17) lag 1: 0.95 (0.82; 1.10)/0.99 (0.82; 1.19) lag 2: 0.99(0.81; 1.15)/1.05 (0.84; 1.3) ma 0-6: 1.09 (0.64; 1.87)/ 1.41 (0.73; 2.71), Cough: lag 0: 0.98 (0.92; 1.05)/0.98 (0.91; 1.06) lag 1: 1.01 (0.94; 1.08)/ 0.97 (0.90; 1.05) lag 2: 0.97 (0.9; 1.05)/0.92 (0.81; 1.04) ma 0-6: 0.89 (0.71; 1.12)/0.82(0.62; 1.1)
Scripted Exposure												
Langrish et al. (2012)	China, Beijing	03/2009-05/2009	Scripted Exposure	98, non-smoking adults, mean age: 62 yrs, history of CAD	Measure-ment: Mo-bile	Total	CPC 3007	NA	Symptoms	Standard-ized-clinical examina-tions, Self-reported	2 hour walk, 24 hour study period	Group comparison: Mask use vs. no mask:
Time-series												

Reference	Country, City	Study period	Study Design	Sample Size, Main study population	Exposure Assessment	Size Fractions	Techn. device	Covariate adjustment	Outcome	Outcome Assessment	Exposure time windows	Effect sizes (confidence intervals) per increment
Wolf et al. 2015	Germany, Augsburg	1999-2009	Time-series	15,417, general population	Measurement: Central site	PNC10-2000 nm	CPC, TDMPS	time trend, T, season, DOW	MI and coronary deaths	Official registry	lag 0 lag 1 ma 0-5	RR for total events per 6'800/ml lag 1: 1.5% (-0.8; 3.7) lag 2: 0.4& (-1.9; 2.8) ma 0-5: 0.8% (-1.7; 3.4) **Nonfatal** events: lag 1: 1.6% (-1.5; 4.8) lag 2: 0.3% (-2.9; 3.6) ma 0-5: 2% (-1.5; 5.8) Incident events: lag 1: 0.7% (-2.1; 3.5) lag 2: -0.1% (-2.9; 2.8) ma 0-5: -0.2% (-3.3; 2.9) Recurrent events: lag 1: 4.1% (-0.9; 9) lag 2: 3.8% (-1.1; 8.9) ma 0-5: 6% (0.6; 11.7)

[a] CAD: Coronary artery disease, CESD-R: Revised Center for Epidemiological Studies Depression Scale, COPD: Chronic obstructive pulmonary disease, DOW: Days of week, MA: Mean average, MI: Myocardial infarction, NA: Not available, NR: No reference, OR: Odds ratio, PNC: Particulate number concentration, T: Temperature.

[b] CPC: Condensation particle counter, TDMPS: Twin Differential Mobility Particle Sizer.

c MOBILIZE: Maintenace of Balance, Independent Living, Intellect and Zest in the Elderly of BOSTON.

Table A1c: Primary research articles presenting methods and results of UFP/quasi-UFP epidemiologic short-term studies, emergency/hospital admissions

Reference	Country, City	Study period	Study Design	Sample Size, Main study population	Exposure Assessment	Size Fractions	Tech. device	Covariate adjustment	Outcome	Outcome Assessment	Exposure time windows	UFP effect sizes (conficence intervals)
Case-crossover												
Evans et al. (2014)	USA, Rochester (NY)	08/2006-06/2009	Case-crossover	74, asthmatic children, 3-10 yrs,	Measurement: Central site	PNC <100nm, AccMP: 100-500nm	SMPS	T, RH	Number of paediatric asthma visits	Medical records	Lag **days** ma 1 ma 2 ma 3 ma 4 ma 5 ma 6 ma 7	**PNC/ORs** ma 1: 0.89 (0.64; 1.24)per 3,007/ml ma 4: 1.27 (0.9; 1.79) per 2,088/ml **AccMP/ORs** ma 1: 0.73 (0.50; 1.08) per 874/ml ma 4: 1.00 (0.71; 1.4) per 638/ml
Gardner et al. (2014)	USA, Rochester (NY)	01/2007-12/2010	Case-crossover	338 STEMI 339 NSTEMI events	Measurement: Central site	PNC <100nm, AccMP: 100-500nm	SMPS	T, RH	Cardiac Catheterizations due to acute coronary symptom, STEMI	Medical records	Lag **hours** ma 0 ma 0-2 ma 0-11 ma 0-23 ma 0-47 ma 0-71 ma 0-95	**PNC/ORs** lag 0: 1.03 (0.95; 1.12) per 4,245/ml lag 0-23: 1.06 (0.89; 1.26) per 3,284/ml **AccMP/ORs** Lag 0: 1.07 (0.91; 1.27) per 860/ml, Lag 0-23: 1.12 (0.92; 1.38) per 775/ml
									Cardiac Catheterizations due to acute coronary symptom, NSTEMI	Medical records	Lag **hours** lag 0 lag 0-2 lag 0-11 lag 0-23 lag 0-47 lag 0-71 lag 0-95	**PNC/ORs** lag 0: 0.99 (0.90; 1.10) per 4245/ml lag 0-23: 1.01 (0.86; 1.18), per 3284/ml **AccMP/ORs** Lag 0: 0.97 (0.82; 1.15) per 860/ml Lag 0-23: 0.97 (0.81; 1.17) per 775/ml

Reference	Country, City	Study period	Study Design	Sample Size, Main study population	Exposure Assessment	Size Fractions	Tech. device	Covariate adjustment	Outcome	Outcome Assessment	Exposure time windows	UFP effect sizes (conficence intervals)
Iskandar et al. (2012)	Denmark, Copenhagen	05/2001-12/2008	Case-crossover	8,226, children aged 0-18 years admitted in 8 specific hospitals	Measurement: Central site	PNC 10-700nm	DMPS	Dew point, wind speed, global radiation	Hospital admission due to Asthma	Official registry	ma 5 (lag 0-4)	**ORs** per 3,812.86/ml overall: 1.06 (0.98; 1.14) 0-1 year-olds: 1.08 (0.97; 1.22) 2-5 year-olds: 1.07 (0.96; 1.20) 6-18 year-olds: 1.02 (0.91; 1.15)
Rosenthal et al. (2013)	Finland, Helsinki	1998-2006	Case-crossover	2,134 (all cardiac), MI: 629, other: 1505, patients with out-of-hospital cardiac arrest, mean age: 68 yrs	Measurement: Central site	PNC <100nm AccMP: 100-320 nm	DMPS	T, RH	Out-of hospital card. arrest, all cardiac causes	Adm. database	lag 0h lag 1h lag 2h lag 3h ma 07h Lag 0d Lag 1d Lag 2d Lag 3d ma 03d	ORs/ PNC per 10,624/ml ma 3d: 0.92 (0.78; 1.09) - lag 3d: 1.03 (0.93; 1.15) ORs/ ACCMP per 1,007/ml lag 2d: 0.96 (0.89; 1.03) lag 0d: 1.04 (0.97; 1.12)
									Out-of hospital card. arrest, MI	Adm. database		ORs/ PNC per 10,624/ml: lag 3d: 0.97 (0.80; 1.05) lag 0d: 1.27 (1.05; 1.54) ORs/ ACCMP per 1,007/ml lag 2d: 0.96 (0.8; 1.10) lag 0d: 1.19 (1.04; 1.35)

Reference	Country, City	Study period	Study Design	Sample Size, Main study population	Exposure Assessment	Size Fractions	Tech. device	Covariate adjustment	Outcome	Outcome Assessment	Exposure time windows	UFP effect sizes (conficence intervals)
									Out-of hospital card. arrest, Other cardiac	Adm. database		ORs/ UFP per 10,624/ml: range: 0. 86; 1.07 ACCMP: Other cardiac, range: 0.95; 1.04
Wichmann et al. (2013)	Denmark, Copenhagen	01/2000-12/2010	Case-crossover	4,657 Patients with OHCA, mostly older than 75 yrs	Measurement: Central site	PNC: 10-700nm, PAC: 10-700nm, PVC: 10-700nm	DMPS, custom built	Public holidays, T, RH	Out-of hospital cardiac arrest	Adm. database	Lag days lags 0-5, ma 2,4,6,	Estimated by figures in supplement, per cent excess risk: PNC: range: -3 to +3 per 3,828 PAC, range: -4.5 to + 2.5 per 155.00 $\mu m^2/m^3$ PVC, range: -4 - +2 per 7.14 m^3/m^3 increase per hourly AP levels: PNC, range: -4- +1 per 4,856/ml PAC, range: -5 - -3 per 174.71 $\mu m^2/m^3$ PVC, range: -6 - -4 per 7.77 $\mu m^3/m^3$
Time-series												
Delfino et al. (2014)	USA, California	2000-2008	Time-series	7,492 children 0-18 with a primary diagnosis of asthma, (11,177)	Dispersion	Not reported/ reference given	NR	T, RH	hospital admission/emergency dep. visits with a primary diagnosis of asthma	Medical records	Lag days 0-7 ma 1,3,5,7	PNC analyzed only as a mediator per cool: 1266 particles/m3 warm: 1041 particles/m3
Diaz-Robles et al. (2014)	Chile, Temuco	08/2009-06/2009	Time-series	2001: 255594 2011: 309354 (68 visits	Measurement: Central site	PM < 100nm	MOUDI, 100-NR model	T, RH, wind speed, Thermo-hygro-	Outpatient visits for respiratory illness	Medical records	Lag **days** 0-5	RRs per 4.73 $\mu g/m^3$ lag 1: 0.99 (0.96; 1.01) lag 4: 1.07 (1.04; 1.10)

Reference	Country, City	Study period	Study Design	Sample Size, Main study population	Exposure Assessment	Size Fractions	Tech. device	Covariate adjustment	Outcome	Outcome Assessment	Exposure time windows	UFP effect sizes (conficence intervals)
				/ day), general population				metric index, Steadman index				
Lanzinger et al. (2016b)	Germany, Augsburg and Dresden; Czech Republic, Prague; Slovenia, Ljubljana; Ukraine, Chernivtsi	01/2011-03/2014, city-specific times overlapping	Time-series	2,582,000, general population	Measurement: Central site	PNC 20-100 nm (UFP) PNC 20-800 nm (PNC)	custom made Differential or Scanning MPS	Time-trend, DOW, public holidays, vacation periods, influenza periods, T, RH	Cv. hospital adm.	Adm. database	Lag days: ma 0-1 ma 2-5 ma 0-5	Percent changes in RRs/ UFP per 2,750/ml ma 0-1: -0.6 (-2.4; 1.1) ma 2-5: 0.3 (-1.7; 2.4) RRs/ PNC per 3675/m ma 0-1-0.6 (-2.3; 1.3) ma 2-5: 0.8 (-1.3; 2.9)
									Resp. hospital adm.	Adm. database	Lag days: Lag 1-5 ma 0-1 ma 2-5 ma 0-5	RRs/ UFP per 2750/ml ma 0-1: 1.5 (-3.4; 6.7) ma 0-5: 3.4 (-3.2; 7.3) RRs/ PNC per 3675/m ma 0-1: 1.9 (-3.2; 7.3) ma 0-5: 4.3 (-0.9; 9.8)
									Diabetes hospital adm.	Adm. database	Lag days: Lag 1-5 ma 0-1 ma 2-5 ma 0-5	RRs/ UFP per 2750/ml ma 0-1: 0.4 (-4.7; 5.7) ma 0-5: 2.9 (-4.5; 10.9) RRs/ PNC per 3675/m ma 0-1: 0.6 (-4.7; 6.3) ma 0-5: 3.9 (-3.7; 12.1)
Samoli et al (2016a) Clearflo	UK, London	01/2011-12/2012	Time-Series	appr. 9 million (>700/day)	Measurement: Central site	PNC>0.6nm	CPC model 3022	trend, DOW, public holidays, T, RH	cv hospital admissions	Medical records	lag 1d	Percent changes per 5,180/ml 15-64y: 0.81 (−0.78; 2.42) 65+: −0.07 (−1.27; 1.15)

Reference	Country, City	Study period	Study Design	Sample Size, Main study population	Exposure Assessment	Size Fractions	Tech. device	Covariate adjustment	Outcome	Outcome Assessment	Exposure time windows	UFP effect sizes (conficence intervals)
									resp. hospital admissions	Medical records	lag 2d	0-14y: 1.86 (−0.28; 4.05) 15-64y: −1.14 (−2.66; 0.41) 65+: −1.09 (−2.42; 0.27)
Samoli et al. (2016b) UF Health	Denmark, Copenhagen; Finland, Helsinki; Italy, Rome, Sweden, Stockholm, Spain, Barcelona	2001-2011	Time-Series	appr.. 9 million General population	Measurement: depending on site, mostly single site	B: 5-1000nm, C: 6-700nm, H: 10-100nm, R: 7-3000, S: 7-3000/ 4-3000	B: WPCP, C: DMPS, H: ?, R: CPC, S: CPC	T, influenza periods	resp. hospitalizations	Medical records	lag days: 0-10	Percentage change per 10,000/ml lag 0: -0.44 (-1.73; 0.87) lag 1: -0.58 (-1.93; 0.79) lag 2: -0.22 (-0.92; 0.38) lag 5: 0.43 (-0.58; 1.45) lag 7: -0.37 (-1.39; 0.66)
Liu et al. (2013)	China, Beijing	03/2004-12/2006	Time-series	15,380,000 , general population	Measurement: Central site	only PNC: PNC3-10nm PNC10-30nm PNC30-50nm PNC50-100nm PNC & mass: 100-300nm 300-1000nm 3-100nm 3-1000nm	TDMPS (TSI model 3221)	T, RH, Public holidays, season	total cv emergency room visits	Medical records	Lag days ma 0-1 ma 0-10	Percentage changes, PNC 3-100 ma 0-10: 7.2 (1.1; 13.7) per 9,040/ml ma 0-1: 1.1 (-3.0; 5.3) per 10,340/ml PNC 3-1000 ma 0-10: 5.8 (-0.5; 12.4) per 10,310/ml ma 0-1: 2.2 (-2.2; 6.8) per 11,990/ml PM 3-1000 ma 0-10: -0.3 (-3.2; 2.6) per 40.7 µm/m³ ma 0-1: 1.4 (-1.4; 4.3) per 68.5 µm/m³

[a] AccMP: Accumulation mode particles, cv: Cardiovascular, MA: Mean average, NSTEMI: non ST-elevation myocardial infarction, OHCA: Out-of-hospital cardiac arrests, OR: Odds ratio, PAC: Particle area concentrations, PM: Particulate matter, PNC: Particulate number concentration, PVC: Particle volume concentration, RH: Relative humidity, RR: Relative risk, SMPS: Scanning Mobility Particle Sizer, STEMI: ST-elevation myocardial infarction, T: Temperature.

[b] DMPS: Differential Mobility Particle Sizer, MOUDI: Micro-Orifice-Uniform-Deposit Impactor, TDMPS: Twin Differential Mobility Particle Sizer.

Table A1d: Primary research articles presenting methods and results of UFP/ quasi-UFP epidemiologic short-term Studies, Subclinical Outcomes

Reference	Country, City	Study period	Study Design	Sample Size, Main study population	Exposure Assessment	Size Fractions	Tech. device	Covariate adjustment	Outcome	Outcome Assessment	Exposure time windows	UFP effect sizes (conficence intervals)
Cohort studies												
Bind et al. (2016) Normative Ageing Study	USA, Boston (Massachussets)	1995-2013	Cohort	1,112, Men (veterans), mean age: 69 yrs	Measurement: Central site	PNC7-3000	CPC, Model 3022A	T, RH, season, age, diabetes, BMI, smoking, pack-years, current use of statin, current use of AHM	SBP, DBP,	Standardized-clinical examinations		SBP difference per 13,845/ml PNC: 10th percentile: 4.9 (1.4; 8.6), 90th percentile: 1.2 (-1.7; 5.1) DBP difference per 13,845/ml PNC: 10th percentile: 3.6 (1.8; 5.6) 90th percentile: 2.9 (1.7; 4.8)
									HR SDNN LF:HF Corrected QT HDL LDL CRP (further outcomes-see article)			difference per 13,845/ml PNC: HR 10th percentile: -1.2 (-5; 2), 50th percentil): -0.2 (-5.5; 2.5) 90th percentile: 6.8 (-3; 17) SDNN 10th percentile: 0.0 (-0.003; 0.003) 90th percentile: -0.03 (-0.07; 0.01) LF:HF 10th percentile: 0.03(-0.065; 0.12) 90th percentile: 0.08 (0.01; 0.13) Corrected QT 10th percentile: -10 (0; -19) 90th percentile: -0.3 (-8; 7) HDL 10th percentile: -0.3 (-1.8; 8) 90th percentile: 1.7 (-1.8; 3.4) LDL 10th percentile: 3.0 (-1.8; 7) 90th percentile: 8.5 (4; 14) CRP 60th percentile (levels > 2 mg/L: 0.4 mg/L (0.1;

Reference	Country, City	Study period	Study Design	Sample Size, Main study population	Exposure Assessment	Size Fractions	Tech. device	Covariate adjustment	Outcome	Outcome Assessment	Exposure time windows	UFP effect sizes (conficence intervals)
												0.7) interval.
Mehta et al. (2014) Veterans Affairs Normative Aging Study	USA, Boston (Massachussets)	07/2007-08/2008	Cohort	370, elderly men	Measurement: Central site	7- to 3000nm	CPC 3022A	Age, BMI, HDL, education, race, alcohol, smoking status/ dose, diabetes, seasonality, DOW, T, RH	Arterial stiffness (AI, AP)	Standardized-clinical examinations	ma 04 hours, ma 01 ma 03 ma 07 ma 14 days	AI/ percentage changes ma 04h: 0.6 (-0.3; 1.7) per IQR (NR) ma 01d: 1.7 (0.4; 2.9) per 8,740/ml ma 03d: 2.2 (0.9; 3.5) per 7,874/ml ma 14d: 2.7 (1.3; 4.2) per IQR (NR) AP/ mmHg ma 04h: 0.2 (-0.5 ; 1.1) per IQR (NR) ma 01d: 0.8 (0.0; 1.7) per 8,740/ml ma 03d: 1.2 (0.2; 2.0) per 7,874/ml ma 14d: 1.6 (0.6; 2.7) per IQR (NR)
Cross-sectional												
Fuller et al. (2015) CAFEH	USA, Somerville / Boston, Massachusetts	clinical examinations: 08/2009-09/2010 UFP measure-ure-ments:11/ 2009-12/2010	Cross-sectional	142 (250 samples)aged > 40 yrs	Central site, spatiotemp. model	NR	SPH site: butanol-based CPC (Model 3022A near-highway site: water-based CPC (Model 3781)	Age, education, BMI, smoking, HTM, income, DoW.	IL-6 (pg/mL),	Standardized-clinical examinations	Lag days, lag 1, lag 2, lag 3, ma 3 ma 7 ma 14 ma 21	Effect estimates were highest for a 28-day moving average, with 91.5% (9.4%, 235%) increase in IL-6, Per 5,000 particles/cm3
								Age, BMI, employment, income, DoW, T	hs-CRP (mg/L)		Lag days, lag 1, lag 2, lag 3,	Effect estimates were highest for a 28-day moving average, with a 74% (−6.6%; 223.0%) increase in hs-CRP Analyses using PNC concentrations at the MAC (near motorway central site) did not identify

Reference	Country, City	Study period	Study Design	Sample Size, Main study population	Exposure Assessment	Size Fractions	Tech. device	Covariate adjustment	Outcome	Outcome Assessment	Exposure time windows	UFP effect sizes (conficence intervals)
											ma 3 ma 7 ma 14 ma 21	strong trends in effect estimates with the biomarkers. There was, however, a statistically significant 12.3% (−17.8%; −6.4%) decrease in hsCRP.
								Age, race, education, BMI, CHF, employment, T,	TNF-RII (pg/mL)		Lag days, lag 1, lag 2, lag 3, ma 3 ma 7 ma 14 ma 21	There were statistically significant associations for a 14-day MA with TNF-RII
								Age, education, BMI, smoking, CHF, DoW	Fibrinogen (mg/dL)		Lag days, lag 1, lag 2, lag 3, ma 3 ma 7 ma 14 ma 21	Effect estimates were highest for a 28-day moving average, with 58.7 pg/mL (−12.8%; 130.2%) increase in fibrinogen with each 5000 unit increase in the 28-day MA of PNC. MAC did not identify strong trends in effect estimates with the biomarkers.
Karottki et al. (2014)	Denmark, Copenhagen	10/2011-02/2012	Cross-sectional	outdoor: 49, indoor: 75, non-smoking adults, 41-68 yrs	Measurement: Central site, Measurement: Residential	outdoor: PNC10-280 nm, indoor: PNC10-300 nm	CPC, DMPS	Age, sex, BMI, vasoactive drugs	MVF, lung function, inflammatory markers	Standardized-clinical examinations	Lag 2d	Outdoor effect of PNC per 1,001/ml, percent changes: MVF: -8.9 (-15.9; -1.4), HBA1c: −1.5 (−8.1; 5.5) hsCRP: 46.5 (−10.5; 139.9) FEV1/FVC: 2.2 (−0.8; 5.3)
Ljungman et al. (2014) Framing-	USA, Boston (Massachussets	3rd generation cohort: 2003-	Cross-sectional	2,072, mean age: 56 yrs,	Measurement: Central site	Total	CPC, Model 3022A	Age, sex, cohort, diabetes, BMI,	Peripheral arterial tonometry ratio,	Other	ma 1 ma 2 ma 3 ma 5	PAT ratio: No consistent pattern of association was evident between averaging periods of particle number hyperemic response Baseline pulse amplitude

Reference	Country, City	Study period	Study Design	Sample Size, Main study popula-tion	Exposure Assessment	Size Frac-tions	Tech. device	Covariate adjust-ment	Outcome	Outcome Assess-ment	Expo-sure time win-dows	UFP effect sizes (conficence intervals)
ham heart study	)	2005, Offspring Cohort: 2005-2008						triglycer-ide level, ratio of total choles-terol to HDL, SBP, income, educa-tion, smoking, DoW, season, time trend, T, RH, T × RH, use of statin /AHM	baseline pulse ampli-tude		ma 7	ma 1: 12.5 (5; 21) per 15,000/ml ma 5: 18.2 (8.9; 28.2) per 15,000/ml ma 7: 18.4 (8.9; 28.7) per 15,000/ml
Panel (cross-sectional)												

Reference	Country, City	Study period	Study Design	Sample Size, Main study population	Exposure Assessment	Size Fractions	Tech. device	Covariate adjustment	Outcome	Outcome Assessment	Exposure time windows	UFP effect sizes (conficence intervals)
Croft et al. (2017)	USA, Rochester (NY)	11/2011-12/2013	Panel (cross-sectional)	135, adults ≥ 18 yrs, with acute coronary syndrome	Measurement: Central site	PNC10-100 nm (UFP) PNC100-500 nm (AccMP)	3071 Electrostatic Classifier with a 3010 CPC	Age, dyslipidemia, prior MI, smoking , year, weekday, hour of the day, T, RH	CRP, Fibrinogen, MPO, D-dimer	Standardized-clinical examinations	Lag hours:1 h 12 h 24 h 48 h 72 h 96 h	AccMP, percent changes 1-24h lags, most distict estimate Fibr,12h: 2.40 (1.30; 3.50) per 452/ml CRP, 1h: 3.17 (−0.75; 7.09) per 395/ml MPO: 12h: −2.80 (−4.68; −0.92) per 452/ml d-dimer, 12h: 0.23 (−3.25; 3.71) per 452/ml 72 and 96 h lags less distinct. UFP, percent changes CRP: 1h: 1.25 (−0.63; 3.12) per 2202/ml 12h: 3.11 (−1.40; 7.62) per 2477/ml 48-76h lags inconsistent. Fibrinogen, 12h: 2.33 (1.07; 3.59) per 2477/ml MPO, 12h: −3.28 (−5.32; −1.23) per 2477/ml
Panel (repeated measure)												
Bartell et al. (2013)	USA, Los Angeles	2005-2007	Panel (repeated measure)	50, Retirement community residents, ≥ 65 yrs, history of CAD, non-smoking, w/o exposure to ETS	Measurement: Residential	PNC5 - 3000 nm PM0.25: 0-250 nm	CPC model 3785	Daily average actigraph-derived physical activity and heart rate, T, DOW, seasonal study phase, community group	HRV, Arrhythmia	Standardized-clinical examinations	Lags PNC 1h 8h 24h 25-48h 2-day 3-day 5-day Lags PM0.25: 0 (24 hr), 1 (25–48 hr), 2d	ventricular tachycardia, per 6,351/ml PNC:: RRs, daily 24h: 0.70 (0.41; 1.20), 3d:: 0.42 (0.09; 1.94), 5d: 0.20 (0.02; 1.67), ORs, hourly daytime: 1h1.06 (0.86; 1.30), 8h 0.90 (0.64; 1.26), 3d: 1.16 (0.41; 3.26), 5d: 2.43 (0.55; 10.7), hourly nighttime ORs: 1h: 0.77 (0.59; 1.01), 8h: 1.09 (0.70; 1.70), 3d: 0.70 (0.26; 1.92), 5d: 0.88 (0.10; 7.89). per 7.0 microg/m³ PM0.25, Daily RRs 0d: 1.04 (0.67; 1.60),

Reference	Country, City	Study period	Study Design	Sample Size, Main study population	Exposure Assessment	Size Fractions	Tech. device	Covariate adjustment	Outcome	Outcome Assessment	Exposure time windows	UFP effect sizes (conficence intervals)
												1d: 1.20 (0.97; 1.47), 2d: 1.29 (0.73; 2.29)
Chung M. et al. (2015) CAFEH	USA, City of Somerville, the Dorchester, South Boston, MA	first visit: 08/2009-4/20111 second visit: 02/2010-06/2011	Panel (repeated measure)	220, resident near highway	Measurement: Central site	PNC < 100nm	CPC (Model 3022a)	Age, gender, race, income, education level, smoking, obesity, AHM, sampling method, distance to highway	SBP, DBP, PP	Standardized-clinical examinations	Daily average (24 h prior to clinic date)	ß-estimates per 10,000/ml PNC SBP: 2.19, Robust SE: 1.82, P.0.23 DBP: 2.40, Robust SE: 1.11, P: 0.03 PP: -0.16, Robust SE: 1.34, P0.91
Cole-Hunter et al. (2016)	Spain, Barcelona	02/2011-11/2011	Panel (cross-sectional)	28, healthy cycling adults	Microscale personal exposure model	PNC <100nm	CPC Model 3007, TSI,	BMI, ambient temperature, noise, linear and quadratic terms for HR	HRV (SDNN, rMSSD, LF, HF, LF:HF)	Standardized-clinical examinations	2 hours	Percentage changes per 10,000/ml SDNN(ms) low traffic site -4.9 (-7.1; -2.7), high traffic site: -0.52 (-0.96; -0.08), similary for RMSSD and LF and HF. Positive estimates for LF:HF e.g. at low traffic site: 1.0 (-3.1; 5.2), high traffic site: 0.17 (−0.66; 1.0)
Frampton et al. (2012)	USA, Rochester (NY)	Not reported/ no reference given	Panel (repeated measure)	19 never smokers,30–60 yrs, with T2D	Measurement: Central site	PNC 10-100nm	SMPS, version 3071	Temperature, relative humidity, order of meas-	platelet expression of CD62P and CD40L,	Standardized-clinical examinations	lag days (1–5), lag days 1–5 combined	ß-estimates per 2,482/ml Platelet CD62P ↓ D2,4, 1–5, Platelet-Leukocyte Conjugates↓ D1,2, 1–5, Platelet CD40L ↓ D1,4, 1–5, Soluble CD40L ↑ D1 only figures and summarizing table

Reference	Country, City	Study period	Study Design	Sample Size, Main study population	Exposure Assessment	Size Fractions	Tech. device	Covariate adjustment	Outcome	Outcome Assessment	Exposure time windows	UFP effect sizes (conficence intervals)
								urement, age-group, and sex	platelet-leukocyte conjugates, circul. MP, CD40L			Number of platelet-leukocyte conjugates decreased by −80 (−123 to −37, p=0.001) on the first lag day (20–44 hours prior to the blood draw) and by −85 (−139 to −31, p=0.005) on combined lag days 1 to 5, However, levels of soluble CD40L increased 104 (3 to 205, p=0.04) pg/ml on lag day 1, a finding consistent with prior platelet activation
Gong et al. (2014)	China, Beijing	06/2008-10/2008	Panel (repeated measure)	125, non-smoking 22–27 yrs, working on campus of hospital and most (92%) residing in dormitories of nearby University	Measurement: Central site	PNC13-108.2nm AccMP: 108.3-764.7nm	TDMPS, CPC	T, RH, sex, DOW	HR BP, vWF, CD40 ligand, P-selectin, pulm.inflammation, OS, FeNO, malondialdehyde, nitrite, OS (urinary malondialdehyde and 8-hydroxy-2'-deoxy-guanosine, plasma fibr., WBC)	Standardized-clinical examinations	lags 0-6d	Percent changes per IQR (not reported) FeNO, lag 0: 25.34% (12.96%; 39.09%) EBC pH value, lag 1: 1.54% (0.79%; 2.28%) EBC nitrite, lag 6: 25.64% (16.12%; 35.94%) WBC, lag 0: 4.1% (1.2%; 7%) urinary MDA, lag 3 10.89% (0.56%; 22.3%) 8-OHdG, lag 5: 42.8% (18.2%; 72.6%) EBC MDA and Plasma fibrinogen showed no significant association

Reference	Country, City	Study period	Study Design	Sample Size, Main study population	Exposure Assessment	Size Fractions	Tech. device	Covariate adjustment	Outcome	Outcome Assessment	Exposure time windows	UFP effect sizes (conficence intervals)
Hampel et al. (2012)	Germany, Augsburg	03/2007-12/2008	Panel (repeated measure)	61, with Diabetes or IGT, non-smoking, w/o cardiac disease	Measurement: Central site	PNC 10-100nm	TDMPS system consisting of two DMA.	Long-term time trend, time of day, DoW, T, RH, bar. pressure	HR, SDNN, rMSSD	Standardized-clinical examinations	1h	Percent changes per 7,157 /ml UFP were only related with lagged decreases in SDNN showing the strongest associations -1.9% [-3.4; -0.4%].
Hampel et al. (2014) KORA	Germany, Augsburg	04/2008-11/2008	Other	5, non-smoking, w/o history of angina pectoris, heart attack or stroke.	Measurement: Mobile	PNC 200->1000	PTRAK, Model 8525	For each outcome separately. T, RH barom. press.	HR, SDNN, RMSSD, HF, LF	Standardized-clinical examinations	Lag minutes concurrent 0-4 5-9 10-14 15-19 20-24 25-29	No association with HR, SDNN and LF. Elevated PNC levels led to delayed reductions in RMSSD and HF. The strongest effects were observed with lags of 15–19 min, 20–24 min, and 25–29 min for RMSSD and with lags of 10–14 min, 15–19 min, 20–24 min for HF. Percent changes per 9,581/ml RMSSD, 0-4 min: -2.2 (-4.16; -0.19) 25-29 min: -4.51(-6.38; -2.61) HF, 25-29 min: 2.26 [−4.26; −0.23] 15-19 min:-3.89 [−6.08; −1.65]
Han et al. (2016)	China, Shanghai	04/2010-09/2011	Panel (repeated measure)	55, elderly retired adults, 50-70 yrs, with T2DM or IGT	Measurement: Central site	PNC5.6–100.0 nm (UFP) PNC5.6–20.5 nm (AitMP) PNC100.0–560.0 nm (AccMP) PNC5.6–10.0 nm PNC10.0–20.5 nm	FMPS, TSI	T, RH, DoW, age, Sensitive adjustment: gender, condition of obese, diabetes, hypertension and use of	FeNO	Standardized-clinical examinations	ma up to 24 h	Percent changes, ma 08: 9.25 (2.87; 16.03) per 8,523/ml UFP: 1.44 (−3.21; 6.31) per 3,709/ml PNCnuc 11.68 (4.90; 18.89) per 5.673/ml PNCait: 8.49 (1.71; 15.72) per 2,279/ml PNCacc

Reference	Country, City	Study period	Study Design	Sample Size, Main study population	Exposure Assessment	Size Fractions	Tech. device	Covariate adjustment	Outcome	Outcome Assessment	Exposure time windows	UFP effect sizes (conficence intervals)
						PNC20.5–48.7 nm PNC48.7–100.0 nm PNC100.0–205.4 nm PNC205.4–560 nm.		medication				
Hoffmann et al. (2012)	USA, Boston (Massachussets)	09/2006-07/2010	Panel (repeated measure)	70, non-smoking adults, 40-85 yrs, with T2DM	Measurement: Central site	Total	CPC 3022A TSI	Age, sex, BMI, HbA1c, season, T, years of diabetes, glucose, AHM	Blood pressure	Standardized-clinical examinations	Ma 1-5 days	Percentage changes in SBP: ma 2d: 1.6 (–0.6; 3.9) per 7,300/ml ma 5d: 1.1 (–1.6; 4.0) per 6,600/ml
Huttunen et al. (2012)	Finland, Kotka	11/2005-05/2006	Panel (repeated measure)	52, non-smoking adults, >50 yrs, IHD patients	Measurement: Central site	PNC>20nm	CPC 3007	Time-trend, T	interleukin (IL)-1b, IL-6, IL-8, IL-12, IFN, CRP, fibrinogen, myeloperoxidase and WBC	Standardized-clinical examinations	Lag days: lag 0 lag 1 lag 2 lag 3	Percent changes per 4,841/ml Interleukin 12, lag 0: 2.73 (8.15; 3.01), lag 1: 2.06 (3.53; 7.98), lag 3: 6.41 (0.28; 12.90) Interleukin 8, lag 1: 3.35 (-5.10; 12.55) CRP, lag 1: 4.33 (-4.84; 14.38) Myeloperoxidase, lag 1: 1.29 (-1.83; 4.50) Fibrinogen, lag 1: -0.12 (-1.77; 1.5) WBC, lag 1: 0.17 (-1.44; 1.78)
Karottki et al (2015)	Denmark, Copenhagen	11/2010-05/2011	Panel (repeated measure)	48, non-smoking adults, middle	Measurement: Central site	indoor: PNC10-300 nm, outdoor: PNC10-280	CPC, DMPS	Age, sex, BMI, vasoactive drugs, T,	MVF, lung function, inflammatory	Standardized-clinical examinations	lag 48h	Percent changes of outdoor PNC per 3,000/ml: MVF: -3.4 (-6.6; -0.05), CRP: 3.4 (−6.2; 13.9) FEV1/FVC: −4.0 (−8.1; 0.5) *further outcomes view table 4 in original

Reference	Country, City	Study period	Study Design	Sample Size, Main study population	Exposure Assessment	Size Fractions	Tech. device	Covariate adjustment	Outcome	Outcome Assessment	Exposure time windows	UFP effect sizes (conficence intervals)
				aged (mean age: 68)		nm		season, air filtration	markers			article
Li et al. (2016) CAFEH	Taiwan, Xinzhuang distrcit, New Taipei	02/2008-06/2008	Panel (repeated measure)	59, school children with asthma and/or allergic rhinitis	Measurement: Central site	UFP: 10-100 nm; AccMP: 100-2500 nm; TP: 10-2500 nm	SMPS (TSI); optical aerosol spectrometer (PMS)	Ozone	Spirometric indices	Standardized-clinical examinations	Lag 1d	ß-estimates per 5,646.4/ml UFP: 0.2-0.25, significant for FEF 50% and FEF 75%, Adverse estimates only for Factor 5, secondary aerosol contributors No significant associations of FVC with AccMP
Manney et al. (2012) RUPIOH study	The Netherlands, Amsterdam; Greece, Athens; UK, Birmingham; Finland, Helsinki	10/2002-03/2004	Panel (repeated measure)	133, adults, ≥ 35 yrs, asthmatic or COPD patient	Measurement: Central site, Measurement: Residential	PNC >7nm	CPC 3022A, TSI	City, T, season, trend	levels of nitrite plus nitrate (NOx) in exhaled breath condensate (EBC)	Standardized-clinical examinations	Lag days Lag 0, Lag 1, Lag 2	Percent change per 10,000/ml PNC central site / residential outdoor lag 0: -4.3 (-17.7; 11.1 / 2.9% (-8.6; 15.7) lag 1: -5.1 (- 17.9; 9.8) / -4.3% (-16.6; 9.8) lag 2:-14.0 (-26.6;- 0.8) / -6.1% (- 17.7; 7.1)
Peng et al. (2016)	USA, Boston (Massachussets)	08/2006-07/2010	Panel (repeated measure)	70, non-smoking adults, 40-85 yrs, with T2DM	Measurement: Central site	Total	NR/ no reference given	Subject, T, water vapor pressure, season, scrubbed room NO	NO in exhaled breath (FeNO)	Standardized-clinical examinations	lag 6h lag 24h lag 2d lag 5d lag 7d	Percent changes per 8,270/ml Lag 6h: 9.86 (3.59; 16.52), in general slightly decreasing estimates with greater lags to app. 9.00 (-1; 20) at lag 7d.

Reference	Country, City	Study period	Study Design	Sample Size, Main study population	Exposure Assessment	Size Fractions	Tech. device	Covariate adjustment	Outcome	Outcome Assessment	Exposure time windows	UFP effect sizes (conficence intervals)
Peters et al. (2015)	Germany, Augsburg	03/2007 - 12/2008	Panel (repeated measure)	64, non-smoking adults, mean age: 66, 32 with confirmed T2DM and 32 with IGT	Measurement: Mobile, Measurement: Personal	Personal: PNC10-1000 nm, Central: PNC10-100nm, PNC100-800nm	personal: CPC 3007, central: TDMPS	Trend, meteorology, time of day, further adjustment for noise	HR and measures of HRV incl. SDNN	Standardized-clinical examinations	Concurrent, 0-4 min, 5-9 min, 10-14 min	Percent changes per 16,000/ml personal PNC measurements: SDNN, concurrent −0.56 (-1.02; −0.09), lag 0-4 min: 0.36 (−0.11; 0.83) HR, concurrent: -0.06 (−0.18; 0.07) lag 0-4 min: 0.23 (0.11; 0.36) lag 5-0 min: 0.16 (0.04; 0.28) RMSSD: estimates close to 0
Pieters et al. (2015) HEAPS study	Belgium, Antwert	05/2011-12/2011	Panel (repeated measure)	130, children 6-12 yrs, attending two primary schools, not exposed to ETS	Measurement: Central site	20–30 nm, 30–50 nm, 50–70 nm, 70–100 nm, 100–200 nm, > 200 nm, total	SMPS; model 3080	Sex, age, height, weight, parental education, neighborhood SES, fish consumption, HR school, DoW, season, wind speed, T, RH, season x T	BP, IL–1β	Standardized-clinical examinations	Lag 0	SBP, ß-estimates (mmHg): PN20-30nm: 6.35 (1.56; 11.47) per 860/cm3 30–50 nm: 1.18 (0.05; 2.31), per 712/ml, 50–70 nm, 0.92 (–0.05; 1.89) per 540/ml, 70–100 nm: 0.86 (0.05; 1.68) per 358/ml, Total UFP: 2.92 (0.30; 5.61) per 1,666/ml IL-1ß: 20-30nm: 24.20 (4.83; 47.16) 30-50 nm: 4.27 (–0.56; 9.35) 50-70 nm: 3.79 (–0.30; 8.05) 70-100 nm: 3.28 (0.33; 6.31) 100-200nm: 1.40 (0.13; 2.68) >200nm: 1.98 (–0.48; 4.49) Total UFP: 2.92 (0.30; 5.61)

Reference	Country, City	Study period	Study Design	Sample Size, Main study population	Exposure Assessment	Size Fractions	Tech. device	Covariate adjustment	Outcome	Outcome Assessment	Exposure time windows	UFP effect sizes (conficence intervals)
Rich et al. (2012)	USA, Rochester (NY)	06/2006-11/2009	Panel (repeated measure)	76, with previous MI or unstable angina	Measurement: Central site	PN10-100 nm (UFP) PN100-500 nm (ACCMP)	Wide range particle spectrometer (model 1000XP)	Visit number, calendar time since the beginning of the study for each participant, month of year, hour of day.	Preexercise resting period: MeanNN, SDNN, rMSSD, QTc, TpTe; whole session: MeanNN, SDNN, rMSSD, HRT, DC; preexercise meas. CRP, fibrinogen, WBC, BP	Standardized-clinical examinations	lag hours 0–5, 0-23, 24–47, 48–71, 72–95, 96–119	ß-estimates, preexercise resting period: MeanNN or SDNN: no clear pattern, rMSSD: ACCMP (similar but less distinct pattern for UFP) : 0-5 h:-3.65 ms (–6.39; –0.91) per 897/ml 0-23h: -4.33 msec (–7.27; –1.38) per 897/ml QTc duration: no pattern, TpTe (msec): 0-23h: 0.78 msec (0.02; 1.53) per 897/ml 24-47h: 1.05 msec (0.28; 1.82) per 897/ml SBP: increase for UFP per 2,680/ml & ACCMP per 897/ml at almost all lags, of which, the largest were significant 0.89 mmHg (95% CI: 0.06, 1.72) and 0.94 mmHg (95% CI: 0.02; 1.87) increases associated with IQR increases in UFP lagged 24–47 hr
Rückerl et al. (2014)	Germany, Augsburg	03/2007-12/2008	Panel (repeated measure)	274, T2DM: 83, IGT: 104, genet. susc.: 87, non-smoking adults, mean age: 62	Measurement: Central site	3–10 nm, 10–30 nm, 30–50 nm, 50–100 nm	TDMPS	T, RH, Pressure, weekday	CRP, interleukin (IL)-6, soluble CD40 ligand (sCD40L), fibrinogen, myeloperoxidase (MPO),	Standardized-clinical examinations	Lag days 0-4, ma 5d	percent change in the panel of T2DM or IGT CRP, NC3-100nm lag 3: 11.7 (3.0; 21.1) per 5722/ml ma 5: 12.2 (2.1; 23.3) per 4,279/ml: NC3–10 nm, ma5: 5.8 (0.7; 11.1) per 390/ml NC30–50 nm: lag 3: 10.9 (2.2; 20.4) per 1,748/ml MPO NC3-100nm: ma 5: 5.8 (0.7; 11.1) per 4,279/ml:

Reference	Country, City	Study period	Study Design	Sample Size, Main study population	Exposure Assessment	Size Fractions	Tech. device	Covariate adjustment	Outcome	Outcome Assessment	Exposure time windows	UFP effect sizes (conficence intervals)
				yrs					and plasminogen activator inhibitor-1 (PAI-1)			NC30–50 nm, ma 5:6.0 (0.9; 11.4) per 1251/ml NC50–100 nm: ma 5: 5.8 (1.6; 10.1) per 1546/ml sCD40L, NC3–10 nm lag 0h: 7 (1.1; 13.2) per 481/ml Results for PAI-1, IL6, Fibrinogen see table D1.1
Rückerl et al. (2016) KORA	Germany, Augsburg	03/2007-12/2009	Panel (repeated measure)	274, T2DM: 83, IGT: 104, genet. susc.: 88, non-smoking adults, mean age: 62 yrs,	Measurement: Central site	PNC3–10, PNC10–30, PNC30–50, PNC50–100, LC(EAD), LC10-800, LC3–10, LC10–30, LC30–50, LC50–100, SC(DCPS), SC10–800, SC3–10, SC10–30, SC30–50, SC50–100	LC(EAD): electric aerosol detector (EAD, model 3070 A), Active surface of the particles, SC(DCPS): Diffusion Charging Particle Sensor (DCPS) (model LQ1)	T, RH, Pressure, weekday	CRP, interleukin (IL)-6, fibrinogen, myeloperoxidase (MPO)	Standardized-clinical examinations	Lag days 0-4, ma 5d	Percent change CRP PNC10–30 nm, lag 3: 13.1 [3.3; 23.8] PNC50–100 nm lag 3: 9.6 [1.8; 18.9] per 0.3 mm/cm³ LC(EAD), lag 1: 6.6 (0.1; 13.6) ma 5: 8.7 (0.3; 17.8) per 0.00 mm/cm³ LC3-10nm, ma 5: 11.7 (2.5; 21.7) per 22.3 mm2/cm³ SC(DCPS) ma 5: 29.8 [15.9; 45.3] per 168.9 mm²/cm³ SC10-800, ma 5: 9.2 (0.8; 18.3) per 0.06 SC3-10nm, ma 5: 9.6 (1.9; 18.0) per 5.7 SC30-50nm, ma 5, 3.2 (-3.9; 10.9) per 24.7 SC50-100nm, ma 5, 4.2 (-2.5; 11.4), similar pictures with significant estimates for MPO, IL-6 less significant, for fibrinogen in some lags significant. In general, estimates for genetically susceptible higher and more often significant.
Sarnat et al. (2014)	USA, Atlanta	2009/12-2011/04	Panel (repeated measure)	42, 21 asthmatics & 21 healthy	Microscale personal exposure model	Not reported/ no reference given	CPC model 3007	noise, cortisol level	HRV (HR, SDNN, rMSSD), CRP, eNO,	Standardized-clinical examinations		At measurement time points within 3 h after the commute, we observed mild to pronounced elevations relative to baseline in exhaled nitric oxide, C-reactive-protein, and exhaled malondialdehyde, indicative of pul-

Reference	Country, City	Study period	Study Design	Sample Size, Main study population	Exposure Assessment	Size Fractions	Tech. device	Covariate adjustment	Outcome	Outcome Assessment	Exposure time windows	UFP effect sizes (conficence intervals)
				non-asthmatics					FEV1, FVC, MDA			monary and systemic inflammation and oxidative stress initiation, as well as decreases relative to baseline levels in the time-domain heart-rate variability parameters, SDNN and rMSSD, indicative of autonomic dysfunction. FEV1 levels were slightly elevated relative to baseline levels among asthmatic subjects at the 1 h and 2 h post-commute time points, the frequency-domain heart-rate variability parameter or other systemic biomarkers ofvascular injury. Water soluble organic carbonwas associated with changes in eNO at all post-commute time-points (po0.0001)
Song et al. (2013a)	South Korea, Inchon City	03/2009-06/2009	Panel (repeated measure)	84, 41 with eczema and 43 healthy children, 8-12 yrs, without ETS at home	Measurement: Central site	PM1 PNC11-101 (UFP) PNC111-930 (AccMP)	SMPS+C comprising a DMA and a CPC (UFP & ACCMP), multi-channel (31 different sizes, 0.25–32 µm) aerosol spectrometer (PM1)	Age, gender, height, day of the week, linear time trend, and meteorological variables such as temperature and humidity (lag 1)	Peak expiratory flow rates (PEFR)	Standardized-clinical examinations	Lag 1d, ma 1-3d	PEFR changes, PM1, children with AD/ without AD ma 1:−2.71 (−4.81; −0.61) /-0.26 (−2.14; 1.61) per 34.1 µg/m³ ma 3: −2.42 (−4.18; −0.65) / −0.36 (−1.91; 1.18) per 19.4 µg/m³ PNC0.1–1, children with AD /without AD ma 1: −1.90 (−4.56; 0.76)/ 0.88 (−1.46; 3.21) per 7,100/ml ma 3: −1.27 (−5.35; 2.80)/ −2.01 (−5.48; 1.46) per 5,370/ml PNC0.01–0.1, children with AD/ without AD ma 1: −1.17 (−3.81; 1.47)/ 1.65 (−0.66; 3.95) per 28,140/ml ma 3: 1.91 (−1.66; 5.48)/ −2.00 (−5.05; 1.06) per 17,680/ml
Song et al. (2013b)	South Korea, Inchon City	04/2009-06-2009	Panel (repeated measure)	84, 41 with eczema and 43	Measurement: Central site	PM1 PNC11.1–101 nm (UFP),	SMPS+C comprising a DMA and CPC	Age, BMI, passive smoking, tempera-	Urinary 8-OHdG levels	Standardized-clinical examina-	Lag days 1-3	Percent changes PM1, Children with eczema/ without eczema lag 1: 4.51 (−1.83; 11.26)/ 0.91 (−5.36; 7.58) per 31.81 µg/m³

Reference	Country, City	Study period	Study Design	Sample Size, Main study population	Exposure Assessment	Size Fractions	Tech. device	Covariate adjustment	Outcome	Outcome Assessment	Exposure time windows	UFP effect sizes (conficence intervals)
				healthy children, 8-12 yrs, without ETS at home		PNC111-454	(UFP & ACCMP), amulti-channel (31 different sizes, 0.25–32 µm) aerosol spectrometer (PM1)	ture on the previous day and time trend (sampling date)		tions		lag 2: −4.48 (−9.50; 0.79)/ −0.06 (−5.49; 5.67) per 31.21 µg/m³ lag 3: −3.58 (−9.78; 3.06)/ 3.73 (−2.91; 10.81) per 31.46 µg/m³ PNC0.1-0.5, Children with eczema/ without eczema lag 1: 5.96 (0.15; 12.10)/ −0.92 (−7.02; 5.58) per 5.49/ml lag 2: 4.11 (−2.68; 11.38)/ 8.14 (1.13; 15.63) per 5.32/ml lag 3: 1.38 (−8.23; 12.00)/ 11.32 (0.58; 23.20) pper 5.51/ml PNC0.01-0.1, children with eczema/ without eczema lag 1: 5.65 (1.31; 10.18)/ 1.99 (−2.93; 7.16) per 32.30/ml lag 2: 6.62 (0.12; 13.54)/ 13.37 (4.74; 22.71) per 32.29/ml lag 3: 2.77 (−2.24; 8.02) 5.87 (−3.71; 16.41) per 32.30/ml
Sun et al. (2015)	China, Shanghai	04/2010-10/2010	Panel (repeated measure)	53, Elderly retired adults, 50-70 yrs, with T2DM or IGT	Measurement: Cen ral site	PNC5-560 nm	Fast Mobility Particle Sizer Spectrometer (FMPS Model 3091)	Age, gender, BMI, visit, DOW, T, RH	HRV (SDNN, rMSSD, LF, HF)	Standardized-clinical examinations	ma 1h, 4h, 12h, 18h, 24h	Percent change in SDNN, ma 4h: PNC5-560: -7.9 (-9.7; -6.1) PNC 10-20: -7 (-8.9; -5.1) PNC 20-50: -6.6 (-8.1; -5) PNC 50-100: -5.4 (-7.3; -3.4) PNC 100-200: -3 (-4.6; -1.3) PNC 200-560: -0.45 (-2.43; 1.56). Other lag hours less positive, with positive estimates at ma 18h and ma 24h. Similar association patterns are observed for other HRV measures, including the root mean square of successive differences between adjacent normal cycles (rMSSD), low frequency (LF) (0.04; 0.15 Hz) and high frequency (HF)

Reference	Country, City	Study period	Study Design	Sample Size, Main study population	Exposure Assessment	Size Fractions	Tech. device	Covariate adjustment	Outcome	Outcome Assessment	Exposure time windows	UFP effect sizes (conficence intervals)
												(0.15; 0.4 Hz), whereas the magnitude of reduction for frequency-domain measure LF and HF were greater IQRs not reported
Wang et al. (2016)	USA, Rochester (NY)	06/2006-11/2009	Panel (repeated measure)	76, postinfarction non-smokers patients with MI or unstable angina	Measurement: Central site	PNC10-100 nm (UFP) PNC100-500 nm (ACCMP)	Wide range particle spectrometer (model 1000XP)	T, calendar time since the beginning of the study, indicator variables for visit number, month of year, and hour of day.	CRP, fibrinogen, SBP, and T-wave complexity, SDNN, rMSSD	Standardized-clinical examinations	lag hours 0-5, 0-23, 24-47, 48-71, 72-95, 96-119. 0-23h 24-47 0-23h:	ß-estimates per IQR (0.87 log particles/ml (6-hour mean) and 0.81 log particles/ml (24-hour) mean log UFP & 1.21 log particles/ml (6-hour mean) and 0.99 log particles/ml (24-hour mean) log ACCMP SBP: lag 0-23h: 1.38 (0.07; 2.68) lag 24-47h: 1.60 (0.32; 2.89) ACCMP and per 0.99 log particles/ml (24-hour mean) lag 0-23h: 1.48 (0.09; 2.86) lag 24-47: 0.61 (-0.89; 2.11) CRP: values close to zero, e.g.: 0-23h: UFP, 0.039 (-0.024; 0.102), ACCMP: 0.051 (-0.017; 0.119), Fibrinogen: 0-23h:UFP: 0.04 (-0.03; 0.11), ACCMP: 0.06 (-0.02; 0.13), 24-47h, UFP: 0.07 (0.00; 0.14), ACCMP: 0.10 (0.02; 0.18), rMSSD, 0-23h, UFP: -3.71 (-7.18; -0.25), ACCMP: -1.95 (-5.64; 1.74), 72-95h, UFP: -7.48 (-10.77; -4.20), ACCMP: -3.54 (-7.02; -0.06), SDNN, 0-23h, UFP: -1.14 (-4.00; 1.71), ACCMP: -1.05 (-4.10; 2.01), Log T wave complexity,

Reference	Country, City	Study period	Study Design	Sample Size, Main study population	Exposure Assessment	Size Fractions	Tech. device	Covariate adjustment	Outcome	Outcome Assessment	Exposure time windows	UFP effect sizes (conficence intervals)
												0-23h, UFP: -0.042 (-0.102; 0.017), ACCMP: -0.059 (-0.123; 0.005)
Wittkopp et al. (2013)	USA, Los Angeles	Not reported/ reference given	Panel (repeated measure)	38, non-smoking adults > 65 yrs with coronary artery disease	Measurement:Retirement communities	PM >250nm AccMP; 250-2500nm	Teflon Filters	Respiratory, urinary tract or other infections during week of biomarker measurements	CRP, TNF-alpha, soluble TNF-alpha receptor II, IL-6, soluble IL-6 receptor	Standardized-clinical examinations	Ma 1-9 days	PM0.25: IL-6 and TNF alpha nonsignificant positive associated positive associations of IL-6 with 3-day and 5-day PM0.25 averages TNF-alpha was positively associated qUFP ß estimates, IQR/PM0.25: 5.28 (mg/m3),
Wu et al. (2012)	Taiwan, Taipei county, Sin-Jhuang	02/2007-03/2007	Panel (repeated measure)	17, non-smoking mail carriers	Measurement: Mobile	PM<0.25µm PM0.25-1µm	Personal cascade impactor sampler	Age, BMI, SHS, T during working period.	rCAVI, SDNN, rMSSD, HF, LF, LF/HF	Standardized-clinical examinations	Mail delivery	Percent change per 15.3 µg/ml PM0.25, SDNN: -4.7 (−14.5; 6.2), r-MSSD: -5.1 (−12.4; 3.0), HF: -5.7 (−16.5; 6.5), LF: -4.8 (−15.1; 6.8), LF/HF: 1.0 (−2.8; 5.0) rCAVI: -2 (-50; 1.0)
Zanobetti et al. 2014	USA, Boston (Massachussets)	2006-2009	Panel (repeated measure)	64, non-smoking adults, 49-85 yrs, with T2DM	Measurement: Central site	Total	CPC 3022A TSI	BAD at baseline, PM2.5, BC, season,	Endothelial function	Standardized-clinical examinations	ma 0-5	Change in mm: -0.02 (-0.1; 0.07) IQR: 8.180/ml for 24-hour mean

Reference	Country, City	Study period	Study Design	Sample Size, Main study population	Exposure Assessment	Size Fractions	Tech. device	Covariate adjustment	Outcome	Outcome Assessment	Exposure time windows	UFP effect sizes (conficence intervals)
Zhang et al. (2013)	China, Beijing	06/2008-10/2008	Panel (repeated measure)	125, non-smoking young adults	Measurement: Central site	SMPS: PNC14.1-736 TDMPS: PNC13-764.7	SMPS (post-Olympics), TDMPS (pre/during Olympics)	Sex, T, RH, period, DOW	HR, HRV,	Standardized-clinical examinations	Lag days: 0-6	Percent changes per 6,572/ml HR: positive associations for most lag days, although statistical significance was observed only at lag day 3 (0.5%). HRV: similar to HR, not significant: SDNN: inconsistent pattern rMSSD: significant negative associations at lag days 0 and 3 LF, HF, LF/HF no clear pattern Blood Pressure: inconsistent patterns Fibrinogen: inconsistent Red blood cell counts: signif. Negative/protective associations WBC signif. Negative/protective and positive associations Urinary HcG: 24.7% at lag 3. FeNO: significantly and positively associated at most lags Other outcomes: see original article
Zhang et al. (2016a)	USA, Los Angeles Canada, Anaheim	2012-2014	Panel (repeated measure)	97, elderly (>65) non-smoking men, w/o psychiatric disorders, renal failure, active	Measurement for PM: 2 monitoring sites	PM0.18, PM0.18-2.5 (AccMP)	MOUDI, model 100-1, MSP Minneapolis	Heat index, exercise, food intake, sugar/ fat intake, use of gas stoves, trend	EBC, MDA, FeNO, oxLDL, IL-6	Standardized-clinical examinations	Ma 5 days	Percent changes, FeNO: ma 5: stronger estimated associations for ultrafine PM0.18 than larger size-fractions for total mass PM 0.18: 3.0 (0.7; 5.3) per 1.1 µg/m³ AccMP: -0.8 (-3.5; 1.9) per 4.0 µg/m³ (various outcomes (elements and PAHs in PM0.18) in figure 1&2 and supplementary tables) MDA: positively associated with total PM0.18 mass

Reference	Country, City	Study period	Study Design	Sample Size, Main study population	Exposure Assessment	Size Fractions	Tech. device	Covariate adjustment	Outcome	Outcome Assessment	Exposure time windows	UFP effect sizes (conficence intervals)
				cancer, acute infections								
Zhang et al. (2016b)	USA, Los Angeles (California)	07/2012-02/2014	Panel (repeated measure)	93, elderly men	Measurement: Central site	PM0.18, PM0.18-2.5 PM2.5-PM10	MOUDI, model 100-1, MSP Minneapolis	Heat index, exercise, food intake, use of gas stoves	Reactive hyperemia index (RHI)	Standardized-clinical examinations	Ma 5 days	RHI slightly adversely/ not significantly associated with 5-day total mass of PM0.18 or PM0.18–2.5, IQR: 1.13 µg/m3
Scripted exposure												
Bos et al. (2011)	Belgium, Brussels	Not reported/ no reference given	Scripted Exposure	35, physically fit, non-asthmatic adults, mean age: 43 yrs, 26% women	Measurement: Mobile	Total	P-Track UFP Counter (TSI Model 8525)	NA	BDNF (brain-derieved neurotropic factor)	Standardized-clinical examinations	20 min cycling versus filtered room	Serum BDNF concentrations increased significantly after cycling in the clean room (p = 0.02). In contrast, BDNF serum concentrations pre/post cycling along the Antwerp Ring did not differ significantly (p = 0.42). Baseline values of BDNF (before cycling) did not differ significantly between the clean room test and the road trial (p = 0.07). Comparison of the values post-cycling did not show any significant differences
Bos et al. (2013)	Belgium: Brussels/ Mol	02/2011-05/2011	Scripted exposure	24, untrained healthy participants	Measurement: Mobile	PNC20-1000	TSI P-TRAK UFP Counters	NA	eNO, BDNF, leukocyte, neutrophil, lymphocyte, eosinophil, mono-	Other, Standardized-clinical examinations	12 week aerobic training program	eNO levels, urban group: increased significantly, Z = -2.87, P = 0.002, in the urban group, whereas eNO levels did not change, Z = -0.7, P = 0.52, in the rural group. Leukocyte count, urban group increased significantly, t(13) = j2.61, P = 0.02, whereas it did not differ significantly over time in the rural group, t(8) = 0.76, P = 0.47, BDNF levels: no group differences before, U = 54, P = 0.45, and after, U =60, P = 0.68, Cogni-

Reference	Country, City	Study period	Study Design	Sample Size, Main study population	Exposure Assessment	Size Fractions	Tech. device	Covariate adjustment	Outcome	Outcome Assessment	Exposure time windows	UFP effect sizes (conficence intervals)
									cyte, basophil counts			tive testing: Reaction times on the Stroop task improved in the rural group (P = 0.001), but not in the urban group.
Jarjour et al. (2013)	USA, Berkeley	04/2011-06/2011	Scripted exposure	15, healthy, never-smoking regular cyclists, 23-48 yrs	Measurement: Personal	PNC 10 - 1000	CPC	NA	lung function	Standardized-clinical examinations	Post-ride & 4h follow-up difference to baseline	Average changes in lung function ranged from −0.1 liters (low-traffic post-ride FEF25-75%) to +0.24 liters (high-traffic 4-hour FEF25-75%), but all changes in lung function measurements were clinically insignificant, and none of the paired t-tests (by subject) for low-traffic and high-traffic lung function changes had significant p-values.
Janssen et al. (2015), RAPTES study	The Netherlands, Utrecht	03/2009-11/2009	Scripted Exposure	31, healthy non-smoking students	Measurement: Mobile	Total	CPC	FeNO, FVC, FEV: T, RH, season, pollen counts, resp.infections NAL: T, RH, season, endotoxin.	FeNO, lung function; IL-6, protein /lactoferrin in NAL; IL-6/ hCRP, Fibrin., vWF, tPA/PAI-1 in plat.	Standardized-clinical examinations	Ma 5h	Percent change per 23,000/ml after excluding underground: FeNO 2h after exposure: appr. 13.0 (6.0; 21.0) increase IL-6 (nasal) 2h after exposure: appr 15.0 (-11; 50) When the underground site was included in the analysis, FeNO and NAL IL-6 were consistently associated with PNC
Langrish et al. (2012)	China, Beijing	03/2009-05/2009	Scripted Exposure	98, non-smoking adults, mean age: 62 yrs,	Measurement: Mobile	Total	CPC 3007	NA	BP, HR, and 12-lead electrocardiography	Standardized-clinic. examination, Self-reported	2-hr prescribed walks	Group comparison: Mask use vs. no mask: mean arterial pressure (93 ± 10 vs. 96 ± 10 mmHg, p = 0.025), HRV (high-frequency power: 54 vs. 40 msec, p = 0.005; high-frequency normalized power: 23.5 vs. 20.5 msec, p = 0.001; root mean square successive differences: 16.7 vs. 14.8

Reference	Country, City	Study period	Study Design	Sample Size, Main study population	Exposure Assessment	Size Fractions	Tech. device	Covariate adjustment	Outcome	Outcome Assessment	Exposure time windows	UFP effect sizes (conficence intervals)
				history of CAD								msec, p = 0.007)
Laumbach et al. (2014)	US, Piscataway, New Jersey	Not reported/ no reference given	Scripted Exposure	21, non-smoking healthy adults	Measurement: Mobile	PNC10-1000 nm	CPC 3007, TSI	personal covariates and noise by design (cross-over and mixed model). In continuous per particle analysis, adjustment for pre-exposure level of outcome	EBC markers of inflammation; HRV	Standardized-clinical examinations	1.5 h ride in passenger vehicle	At immediately post-exposure, an IQR increase in PNC was associated with statistically significant increases in nitrite (99.4%, 32.1%; 166.7%) and nitrite + nitrate (75.7%, 21.5%; 130.0%) No significant associations between exposure to traffic particles and HRV outcomes at any of the time points. Continuous analysis: non-significant rises of EBC markers per IQR of PN exposure
Kubesch et al. (2015)	Spain, Barcelona	022011-11/2011	Scripted exposure	28, healthy non-smoking adults 18-60 yrs	Measurement: Mobile	PNC10-1000 nm	CPC 3007	Sex, BMI, T, RH, ETS energy expenditure, NO2	BP	Standardized-clinical examinations	2 h exposure	ß-estimates, IQRs not given in main text SBP post exposure 1.13 mmHg (0.28; 2.17) DBP post exposure: 0.89 mmHg (0.29; 1.50)

Reference	Country, City	Study period	Study Design	Sample Size, Main study population	Exposure Assessment	Size Fractions	Tech. device	Covariate adjustment	Outcome	Outcome Assessment	Exposure time windows	UFP effect sizes (conficence intervals)
Mirabelli et al. (2015) Atlanta Commuter Exposure Study	USA, Atlanta	12/2009-06/2011	Scripted exposure	39, Non-smoking adults, meadian age: 32 yrs, 19 asthmatic and 21 non-asthmatic	Measurement: Mobile	Total	CPC model3007	NA	Exhaled NO, Malondialdehyde, FEV1 predicted, FVC % predicted, and FEF25–75 % predicted	Standardized-clinical examinations	2h commute by car	Percent changes, 0h, 1h, 2h, 3h post commute Exhaled NO: Non-asthmatics: 2 (-0.2; 0.6), 3 (-1.5; 9), 4 (-1; 10), -1 (-8; 5) Controlled asthmatics: -3.5 (-20; 10), -17 (-28; -3), -17 (-27; -0.5), -17 (-34; 4). Non-controlled asthmatics: 0 (-8; 11), -2 (-13; 9), 3 (-6; 17), 11 (-3; 28) FEV1, categories as above, Non-asthmatics: 1 (-0.2; 3), 1 (-0.2; 3), 1 (-0.5; 1.5), 1.5 (-1.5; 2.5) Controlled asthmatics: -2 (-6; 1.5), -1.5 (-5; 2), -1.5 (-6; 2.5), -1 (-8; 5.5) Non-controlled asthmatics: -1.5 (-3; 1), -1.5 (-3; 1), -1.5 (-4; 1), -3 (-8; 2)
Park et al. (2017)	USA, Sacramento (California)	03/2008 - 06/2008	Scripted exposure	32, healthy adults, frequent bicyclists, mean age: 45.1	Measurement: Mobile	PNC >10	CPC,Model 3007	Sex, age, wind direction, DoW	FVC, FEV1, FEV1/FVC, PEF	Standardized-clinical examinations	Bicycle ride (22km)	ß-estimate changes per 12,225 to 36,833/ml FVC: −0.20 (−0.31; −0.08); FEV1: −0.15 (−0.22; −0.08) FEV1/FVC: 0.00 (−0.01; 0.01), PEF (liters/min): −3.10 (−15.39; 9.18)
Shutt et al. (2017)	Canada, Sault Ste. Marie Ontario	Not reported/ no reference given	Scripted Exposure	60, non-smoking adults, 18–55 yrs	Measurement: Central site	PNC10-1000	TSI Model 3007	HR, age, sex, BMI, T, RH, study site	HRV and components	Standardized-clinical examinations	8h on site stay	Change in ß-estimates per 12,236/ml Heart rate (bpm): 1.10 (0.04; 2.16) HF power(ms^2): -1.89 (-4.38; 0.60) LF power(ms^2): -1.61 (-3.21; -0.01) HF/LF: -0.15 (-0.38; 0.08), SDNN (ms): -7.13* (-12.27; -1.98), RMSSD: -5.03 (-10.63; 0.57), pNN50: -2.20 (-4.24; -0.15)

Reference	Country, City	Study period	Study Design	Sample Size, Main study population	Exposure Assessment	Size Fractions	Tech. device	Covariate adjustment	Outcome	Outcome Assessment	Exposure time windows	UFP effect sizes (conficence intervals)
Steenhof et al. (2013) RAPTES study	The Netherlands, Utrecht	03/2009-11/2009	Scripted Exposure	31, healthy non-smoking students	Measurement: Mobile	Total	CPC	T, RH, season	Cytokine IL-6 and IL-8, protein and lactoferrin in nasal lavage,IL-6 in blood	Standardized-clinical examinations	5h- AP measurement	Changes in ß-estimates per 32,906/ml pre/ 2 h after exposure: NAL IL-6: −2.2 (p> 0.05) NAL protein: 7.9 (p> 0.05), NAL lactoferrin: 4.3 (p> 0.05), serum IL-6: 6.3 (p> 0.05)
Steenhof et al. (2014) RAPTES study	The Netherlands, Utrecht	03/2009-11/2009	Scripted Exposure	31, healthy non-smoking students	Measurement: Mobile	PNC7-3000nm	CPC	T, RH, season	WBC counts: Neutrophils, Monocytes, Lymphocytes, Eosinophile	Standardized-clinical examinations	5h- AP measurement,	Percent changes per 28,100/ml Total WBC, 2h post expo: -2.2 (-5.3; 1.0), 18h post expo: -1.4 (-4.8; 2.2); Neutrophils 2h post expo: -1.3 (-6.2; 3.9), Monocytes 18h post expo: 3.4 (-1.0; 7.9) No robust association between PNC and the number of lymphocytes. No robust association between PNC and the number of eosinophils,
Strak et al. (2012) RAPTES study	The Netherlands, Utrecht	03/2009-11/2009	Scripted Exposure	31, healthy non-smoking students	Measurement: Mobile	NR	CPC	T, RH, season, low/high grasses and birch pollen counts, respiratory infection	FVC, FEV1, FEF25–75%, PEF, FeNO, respire. symptoms	Standardized-clinical examinations	5h- AP measurement	Percent changes per 32,906/ml FeNO (immediately after exposure): 11.24 (5; 17) (p < 0.05), 2h postexpo: 12 (6; 17) next morning: 7 (0.5; 14%) FVC (immediately after exposure): –1.19 (p < 0.05),

Reference	Country, City	Study period	Study Design	Sample Size, Main study population	Exposure Assessment	Size Fractions	Tech. device	Covariate adjustment	Outcome	Outcome Assessment	Exposure time windows	UFP effect sizes (conficence intervals)
Strak et al. (2013a) RAPTES study	The Netherlands, Utrecht	03/2009-11/2009	Scripted Exposure	31, healthy non-smoking students	Measurement: Mobile	PNC7-3000nm	CPC	Temperature, relative humidity, season, use of oral contraceptives	hs-CRP, fibrinogen, platelet counts, vWF, tPA/PAI-1	Standardized-clinical examinations	5h- AP measurement	Percent changes per 32,906/ml 25h post vs. pre: Hs-CRP: -4.31 (-14.35; 6.92) Platelet counts: -1.15 (-2.69; 0.40), vWF: -0.04 (-2.80; 2.80). Exposure of participants to PNC during transport was not associated with changes in acute vascular markers investigated.
Strak et al. (2013b) RAPTES study	The Netherlands, Utrecht	03/2009-11/2009	Scripted Exposure	31, healthy non-smoking students	Measurement: Mobile	Total	CPC	Use of oral contraceptives, T, RH, season	thrombin generation	Standardized-clinical examinations	5h- AP measurement	Percent changes per 32,906/ml endogenous thrombin potential in FXII-mediated thrombin generation pathway: Percent changes per 32,906/ml: all sites: (t=9–t=0): 5.83 (−39.62; 51.29), outdoor sites (t=9–t=0): −0.70 (−52.00; 50.60)all sites (t=25–t=0): −72.40 (−128.56; −16.24), outdoor sites (t=25–t=0): −66.59 (−124.78; −8.40) (post-pre) in ETP two hours after exposure in FXII-mediated thrombin generation pathway,
Weichenthal et al. (2014)	Canada, Montreal	Summer 2013	Scripted Exposure	53, healthy non-smoking women 18-45 yrs, not taking AHM not pregnant or breast-feeding	Measurement: Mobile	PNC 10-100nm	Harvard Impactor and TSI Model 3007	Heart rate, T, caffeine, alcohol, race, age, BMI, recent illness, SHS	HRV (SDNN, RMSSD, pNN50, HF, LF, LF/HF), SBP, DBP, RHI	Standardized-clinical examinations	ma 3h	Percent changes per 10,850/ml, lag 3h: RHI: ma 3h:-4.63 (-8.57; -0.693) SBP: 0.372 (-0.816; 1.56) DBP: 1.29 (-0.329; 2.91) SDNN, ma 3h: 3.61 (0.227; 7.00) numbers from suppl?) Abstract: in UFP exposure was associated with a 4.91% (95% CI: -9.31; -0.512) decrease RHI

[a] AccMP: Accumulation mode particles, AI: Augmentation index, AP: Augmentation pressure, AHM: Antihypertensive medications, AMP: Acuumulation mode particle , BAD: Baseline brachial artery diameter, BC: Black carbon, BDNF: Brain derieved neurotropic factor, BMI: Body mass index, CAD: Coronary artery disease, CD40L: Cluster of differentiation 40 ligand, CD62P: P-selectin (protein), CHF: Chronic heart failure, COPD: Chronic obstructive pulmonary disease, CRP: C-reactive protein, DBP: Diastolic blood pressure, DC: Deceleration capacity, DMA: Differential mobility analyzer, DMPS: Differential mobility particle sizer, DOW: Days of week, EBC: Exhaled breath condensate, eNO: Exhaled nitric oxide, ETS: Enviromental tobacco smoke, FEF 25 - 75: Forced expiratory flow at 25-75% of vital capacity, FeNO: Fractional exhaled nitric oxide , FEV1: Forced expiratory volume in 1 second, FVC: Forced vital capacity, FMPS: Fast mobility particle sizer spectrom, genet. susc.: Genetic susceptibility, HBA1c: Prediabetic marker, HDL: High density lipoprotein, HF: High frequency, HR: Heart rate, HR BP: Heart rate, blood pressure, HRT: Heart rate turbulence, HRV: Heart rate variability, hs-CRP: High-sensitive C-reactive protein, HTM: Hypertensive medication, ICD: International Classification of disease, IHD: Ischaemic heart disease, IGT: Impaired glucose tolerance, IL: Interleukin, LDL: Low density lipoprotein, LF: Low frequency, MA: Mean average, MAC: Mystic Activity Center, MDA: Malondialdehyde, MeanNN: Mean of normal-to-normal intervals , MI: Myocardial infarction, MVF: Microvascular function, MPO: Myeloperoxidase, NAL: Nasal lavage, NAL IL-6: , NR: No reference, OR: Odds ratio, OS: Oxidative stress, oxLDL: Plasma oxidized low-density lipoprotein, PEF: Peak expiratory flow, PEFR: Peak expiratory flow rates, PM: Particulate matter, PNC: Particulate number concentration, PNCacc: PNC accumulation mode particles, PNCait: PNC Aitken mode particles, PNCnuc: PNC nucleation mode particles, PNC::RR: PNC relative risk, PP: Pulse pressure, QT: Q wave and T wave interval, QTc: Heart rate corrected QT-interval, rCAVI: Right cardio-ankle vascular index, RH: Relative humidity, RHI: Reactive hyperemia index, RMSSD: Root mean square of the sucessive differences in ms., SES: Socio-economic status, SBP: Systolic blood pressure, SDNN: Standard deviation of normal-to-normal intervals, SHS: Second hand smoke exposure, SPH site: Harvard school of publich health, T: Temperature, T2DM: Type 2 diabetes mellitus, TNF-alpha: Tumor necrosis factor alpha, TNF-RII: Tumor necrosis factor-a-receptor II, tPA/PAI-1: Tissue plasminogen activator and plasminogen activator inhibitor-1, TpTe: Time from peak to end of T-wave , UFP: Ultrafine particle, vWF: Von Willebrand Factor, WBC: White blood cell counts.

[b] CPC: Condensation particle counter, MOUDI: Micro-Orifice-Uniform-Deposit Impactor, P-TRAK: UFP counter, SMPS: Scanning Mobility Particle Sizer, SMPS+C: Scanning mobility particle sizer and counter, TDMPS: Twin Differential Mobility Particle Sizer.

[c] CAFEH: Community Assessment of Freeway Exposure and Health, HEAPS: Health Effects of Air Pollution in Antwerp Schools, KORA: Cooperative Health Research in the Region Augsburg, RAPTES: Risk of Airborne Particles: a Toxicological–Epidemiological hybrid Study, RUPIOH: Relationship between Ultrafine and fine Particulate matter in Indoor and Outdoor air and respiratory Health.

Table A2a: Primary research articles presenting methods and results of UFP/ quasi-UFP epidemiologic long-term Studies, Mortality

Reference	Country, City	Study period	Study Design	Sample Size, Main study population	Exposure Assessment	Size Fractions	Technical device	Covariate adjustment	Outcome	Outcome Assessment	Exposure time windows	UFP effect sizes (confidence intervals)
Cohort												
Ostro et al. (2015) California Teachers Study	USA, California	01/2001-07/2007	Cohort	101,884 current and former female teachers and administrators, > 30 yrs	CTM	PNC10-100nm	NR	Strata: Age and race, adjusted for smoking status, smoking pack-years, adult SHS exposure, BMI, marital status, alcohol consumption, physical activity, menopausal status and HT use combined, family history of heart disease, hypertension medication/aspirin use, and dietary fat, fiber, and caloric intake	All -cause mortality, CV mortality, IHD mortality, Pulmonary mortality	Administrative database	2000-2007	HRs per 0.969 µg/ml: All-cause mortality: 1.01 (0.98; 1.05), CV mortality: 1.03 (0.97; 1.08), IHD mortality: 1.10 (1.02; 1.18), Pulmonary mortality: 1.01 (0.93; 1.10)

[a] CTM: Chemical transport model, cv: Cardiovascular, HR: Heart rate, IHD: Ischaemic heart disease, NR: No reference.

Table A2b: Primary research articles presenting methods and results of UFP/ quasi-UFP epidemiologic long-term Studies, Morbidity

Reference	Country, City	Study period	Study Design	Sample Size, Main study population	Exposure Assessment	Size Fractions	Technical device	Covariate adjustment	Outcome	Outcome Assessment	Exposure time windows	UFP effect sizes (confidence intervals)
Cross-sectional												
Li et al. (2017)	USA, Somerville Malden, Boston, Dorchester (Massachusets)	2009-2012	Cross-sectional	704 adults, ≥ 40 yrs,	LUR: Spatio-temporal, Microscale personal exposure model	Total (>4 nm)	CPC TSI Model 3775	A) Age, sex, race, smoking status, education, income, time of residence at current address, percieved stress, work status, marital status, sample type, physical activity B) Plus BMI in subgroup C) Plus diagnoses (sensi.anal)	IHD, stroke, CHF; Self-report or medication for DM and/or hypertension	Self-reported	12 months, assumed to be stable over 7-11 years	ORs, increments NR Stroke/ IHD: 1.35 (0.83; 2.22) Diabetes: 0.71 (0.46; 1.10) Hypertension: 1.14 (0.81; 1.62)
Laurent et al. (2014)	USA, California	01/2001-12/2008	Cross-sectional	960,945	CTM	PM0.1	Not reported/ reference given	Maternal race/ethnicity, education, parity, trimester primary care beginning, infant's gender, maternal age, length of gestation and median income	Term low birth weight	Administrative database	2000-2006	ORs per 0,4271 µg/m³: 1.03 (1.02; 1.03)
Case-control												
Laurent et al. (2016 b)	USA, California	01/2001-12/2008	Case-cohort	363,160, 72,632 cases, 290,528 controls	CTM, LUR: Spatio-temporal	CTM: <100nm (PM0.1) CALINE4: PNC (unclear)	CTM: CPC Model 3786	Race/ethnicity and educational level, maternal age and median household income at Census block	Term low birth weight	Administrative database	2000-2006 (PM0.1)	ORs per 6,444/ml PNC: 1.001 (0.989; 1.014) Primary PM0.1, per 1.359 µg/m³: 0.996 (0.981; 1.011) Onroad gasoline PM0.1 per 0.083µg/m³: 1.051 (1.015; 1.089)

Reference	Country, City	Study period	Study Design	Sample Size, Main study population	Exposure Assessment	Size Fractions	Technical device	Covariate adjustment	Outcome	Outcome Assessment	Exposure time windows	UFP effect sizes (confidence intervals)
Laurent et al. (2016 a)	USA, California	01/2001-12/2008	Case-control	1,105,970, 442,314 cases, 710,360 controls	CTM, LUR: Spatio-temporal	CTM: <100nm (PM0.1) CALINE4: unclear (UFP)	CTM: CPC Model 3786	Race/ethnicity, educational level, maternal age, median household income	Preterm birth	Administrative database	2000-2008	ORs PM0.1 per 1.389µg/m³: 1.021 (1.015; 1.028) PNC per 6,480/ml: 0.995 (0.988; 1.000) (geocoded at tax parcel level): 1.028 (1.021; 1.036)

[a] BMI: Body mass index, CALINEA: California Line Source Dispersion Model Version 4, CHF: Chronic heart failure, CTM: Chemical transport model, DM: Diabetes mellitus, IHD: Ischaemic heart disease, LUR: Land use regression, NR: No reference, OR: Odds ratio, PM: Particulate matter, PNC: Particulate number concentration, UFP: Ultrafine particle.

[b] CPC: Condensation particle counter.

Table A2c: Primary research articles presenting methods and results of UFP/ quasi-UFP epidemiologic long-term Studies, Subclinical Outcomes

Reference	Country, City	Study period	Study Design	Sample Size, Main study population	Exposure Assessment	Size Fractions	Technical device	Covariate adjustment	Outcome	Outcome Assessment	Exposure time windows	UFP effect sizes (confidence intervals)
Cross-sectional analysis within cohort												
Aguilera et al. (2016)	Switzerland, Basel/ Geneva/ Lugano/	2001/02 - 2010/11	cross-sectional analysis within cohort	1,503 Adults, ≥ 50 yrs, partici-	LUR	PNC10-300 nm	miniDISC	sex, age, sex–age interaction, educational level, smoking	CIMT	Standardized-clinical examinations	2011-2012	Percent change per 10.-90. percentil PNC: 2.06 (0.03; 4.10) LDSA: 2.32 (0.23; 4.48)

Refer-er-ence	Country, City	Study period	Study Design	Sample Size, Main study popula-tion	Exposure Assessment	Size Frac-tions	Technical device	Covariate adjustment	Outcome	Outcome Assess-ment	Exposure time win-dows	UFP effect sizes (confi-cence intervals)
SAPAL PAL-DIA study	Wald			pants of Sapaldia 2 & 3				status at SAPALDIA2 (S2), smoking pack-years between S2 +SAPALDIA3 (S3), smoking pack-years between S2 and S3)2, BMI at S2, (BMI at S2)2, BMI at S3 and (BMI at S3)				
Cohort												

Reference	Country, City	Study period	Study Design	Sample Size, Main study population	Exposure Assessment	Size Fractions	Technical device	Covariate adjustment	Outcome	Outcome Assessment	Exposure time windows	UFP effect sizes (confidence intervals)
Sunyer et al. (2015) BREATHE	Spain, Barcelona	01/2012-03/2013	Cohort	2,715, children from schools in low vs. high polluted areas	Central measurement at schools plus LUR for exposure assessment at home.	PNC10-700nm	miniDISC	Age, sex, maternal education, residential neighborhood SES, AP exposure at home, school and individual, traffic around school	Working memory, Superior working memory, Inattentiveness	Standardized-clinical examinations	Two weekly measurement campaigns averaged as long-term AP expos	Difference in cognitive development/ß-estimates, per 6,110/ml increase at baseline and 12-mo change Working memory: Baseline: -6.5 (-14; 1.5) 12-mo change: -4.9 (-10; 0.22) Superior working memory: Baseline: -0.95 (-7.4; 5.6), 12-mo change: -5 (-9.1; -0.96) Inattentiveness: Baseline: 4.5 (-4.0; 13) 12-mo change: 3.9 (0.31; 7.6)
Repeated measure within Cohort study												
Viehmann et al. (2015)	Germany, Essen/ Mülheim/ Bochum	2000-2002 (BL), 2006-2008 (FU)	Repeated measure within Cohort study	3,275 with baseline data, 3213 with follow-up data	CTM	PNC5-2200	NR	Sex, (BMI), education, smoking, temperature (1–5 days moving average), season, short-term air pollutant (1–3	hs-CRP, Fibrinogen, WCC, Platelets	Standardized-clinical examinations	365 days	Percent change per 27,000/ml hs-CRP: 3.8 (−0.6; 8.4), Fibrinogen: 1.0 (0.0; 2.0), WCC: 1.0 (−0.1; 2.1), Platelets: 0.6 (−0.4; 1.7)

Refer-er-ence	Country, City	Study period	Study Design	Sample Size, Main study popula-tion	Exposure Assessment	Size Frac-tions	Technical device	Covariate adjustment	Outcome	Outcome Assess-ment	Exposure time win-dows	UFP effect sizes (confi-cence intervals)
								days moving average), time trend and time point.				
Cross-sectional												
Lane et al (2015) CAFEH	USA, Somervil-le/ Boston (Massa-chusets)	07/2009-09/2010	Cross-sectional	140 Adults, ≥ 40 yrs	LUR: Spatio-temporal, Microscale personal exposure model	Total	CPC TSI Model 3775	Age, sex, BMI, smoking status (or SES instead of smoking status)	hs-CRP, IL-6	Stand-ardized-clinical examina-tions	Annual average	ß-estimates, increment unclear, Personal exposure model: Residential annual aver-age+ work+ oth-er+highway+Aircondition: LN hsCRP: 1.26 (−0.02; 2.75) LN IL-6: 0.65 (−0.26; 1.55)
Lane et al. (2016) CAFEH	USA, Boston (Massa-chussets)	07/2009-02/2012	Cross-sectional	408 Adults, ≥ 40 yrs	LUR: Spatio-temporal, Microscale personal exposure model	PNC4-3,000 nm	CPC TSI Model 3775	a) Age, sex, continuous BMI, smoking status and education. B) Age, sex, continuous BMI, smoking status, educa-tion and race/ethnicity . C) Age, sex, BMI, smoking status, educa-tion and nativity.	hsCRP, IL-6, TNFRII, Fibri-ongen	Stand-ardized-clinical examina-tions	Annual average	Percent change per 10,000/ml (IQR) a) hsCRP: 9.8 (-8.3; 31.4), IL-6: 5.8 (-5.6; 18.5), TNFRII: 3.6 (-1.9; 9.4), Fibr.: -1.9 (-5.5; 1.6) b) hsCRP: 14.0 (-4.6; 36.2), IL-6: 8.9 (-2.6; 21.8), TNFRII: 5.1 (-0.4; 10.9), Fibr: -1.9 (-5.5; 1.6) c) ähnlich wie b) White non-Hispanic, a) hsCRP: 32.7 (3.7; 67.2), IL6: 22.6 (-0.2; 45.5), TNFRII: 16.8 (5.8; 27.7), Fibr. -0.02 (-0.7; 0.7), East-Asian: a) hsCRP: 6.1 (-18.3; 31.0), IL6: 2.6 (-12.2; 17.3), TNFRII: 0.1

Refer-er-ence	Country, City	Study period	Study Design	Sample Size, Main study popula-tion	Exposure Assessment	Size Frac-tions	Technical device	Covariate adjustment	Outcome	Outcome Assess-ment	Exposure time win-dows	UFP effect sizes (confi-cence intervals)
												(-1.2; 1.4), Fibr. -0.06 (-5.4; 4.2),

[a] AP: Augmentation pressure, BMI: Body mass index, CIMT: Carotid intima-media thickness, CTM: Chemical transport model, Fibr.: Fibrinogen, hs-CRP: High-sensitive C-reactive protein, IL: Interleukin, LDSA: Lung deposited surface area, LN: Natural log, LUR: Land use regression, PNC: Particulate number concentration, SES: Socio-economic status, TNFRII: Tumor necrosis factor-a-receptor II, WCC: white blood cell count.

[b] CPC: Condensation particle counter, minidisc: Miniature diffusion size classifiers.

[c] BREATHE: Brain Development and Air Pollution Ultrafine Particles in School Children, CAFEH: Community Assessment of Freeway Exposure and Health, SAPALDIA: Swiss Co-hort Study on Air Pollution and Lung and Heart Diseases in Adults.

Table A3a: Short-term studies with adjustment for co-pollutants, mortality

Reference	Exposure time window	Outcome	UFP effect w/o co-pollutant adjustment	PM10 adjusted UFP effect	PM2.5 adjusted UFP effect	NO2 adjusted UFP effect
Lanzinger et al. (2016)	ma 2-5	CV mortality, ma 2-5	RRs/ PNC20-100 per 2750/ml ma 2-5: −0.5 (−5.3; 4.5)	-	RRs/ PNC20-100 Ma 2-5: 0.5 (-0.5; 2)	RRs/ PNC20-100 Ma: 2-5: -5 (-7; 0.5)
	ma 2-5	Resp. mortality, ma 2-5	RRs/ PNC20-100 per 2750/ml ma 2-5: 8.5 (−4.8; 23.7)	-	Ma 2-5: 7 (-10; 30)	Ma 2-5: 14 (2.5; 26)

Reference	Exposure time window	Outcome	UFP effect w/o co-pollutant adjustment	PM10 adjusted UFP effect	PM2.5 adjusted UFP effect	NO2 adjusted UFP effect	
Leitte et al. (2012)	ma 0-3 ma 0-4 lag2	Resp. mortality	Percentage change/ PNC300–1000 per 840/ml ma 0-3: 8.9 (1.3;17) ma 0-4: 11.5 (3.0;20.7) PNC total per 14,000/ml lag 2: 9.3 (1.3;17.9)	PNC300–1000 ma 0-3: 3 (-8; 15) ma 0-4: 8 (-5; 21) PNC total lag 2: 10 (2; 19)		PNC300–1000 ma 0-3: 2 (-9; –13) ma 0-4: 6 (-7; 18) PNC total lag 2: 9 (1.4; 17.9)	SO2: PNC300–1000 ma 0-3: 4 (-5; 15) ma 0-4:.7 (-4; 18) PNC total lag 2: 9 (1; 17.7)
Meng et al. (2013)	ma 01	All-natural-cause mortality	Percent change, all periods, per 2,600/ml PNC250–280: 2.41 (1.23; 3.58) per 63/ml PNC650-1000 0.12 (-0.22; 0.45)	PNC250–280 1.75 (0.26; 3.24) PNC650-1000 –0.12 (–0.56; 0.32)	PNC250–280 2.18 (0.81; 3.55) PNC650-1000 –0.06 (–0.40; 0.29)	PNC250–280 1.66 (0.14; 3.17) PNC650-1000: 0.15 (–0.54; 0.25)	SO2, PNC250–280 2.04 (0.53; 3.54) PNC650-1000 ma 0-1: –0.07 (–0.47; 0.33) PM2.5-10, PNC250-280: 2.52 (1.34 ; 3.71), PNC 650-1000: 0.10 (–0.24; 0.44)
Samoli et al. (2016a)	lag 1d lag 1d lag 2d	non-accidental mortality CV mortality respiratory mortality	−0.06 (−1.16; 1.06) −2.04 (−3.94; −0.10) −1.86 (−4.50; 0.86)				Effect estimates were generally robust to co-source adjustment, alt-hough mutual adjustment for all sources generally exerted greater influence on the estimates com-pared with estimates from two sources models.

Reference	Exposure time window	Outcome	UFP effect w/o co-pollutant adjustment	PM10 adjusted UFP effect	PM2.5 adjusted UFP effect	NO2 adjusted UFP effect	
Stafoggia et al. (2017)	lag 5 lag 6 lag 7	Non-accidental mortality:	Percent increases PNC per 10,000/ml lag 5: 0.32 (-0.08; 0.72) lag 6: 0.35 (-0.05; 0.75) lag 7: 0.37 (-0.03; 0.7%)	Percent increases PNC per 10,000/ml lag 5: 0.16 (-0.25; 0.57) lag 6: 0.22 (-0.18; 0.63) lag 7: 0.28 (-0.13; 0.68)	Percent increases PNC per 10,000/ml lag 5: -0.14 (-0.80; 0.53) lag 6: -0.04 (-0.70; 0.62) lag 7: 0.01 (-0.74; 0.76)	Percent increases PNC per 10,000/ml lag 5: -0.08 (-0.55; 0.40) lag 6: -0.15 (-0.69; 0.38) lag 7: -0.25 (-0.72; 0.22)	PM2.5-10: similar to PM2.5 CO: lag 5: 0.22 (-0.25; 0.70) lag 6: 0.30 (-0.16; 0.77) lag 7: 0.13 (-0.35; 0.60) O3 lag 5: 0.40 (-0.02; 0.82) lag 6: 0.27 (-0.14; 0.69) lag 7: 0.30 (-0.12; 0.72)
Su et al. (2015)	ma 05	overall CVD mortality	Percent increase per 8,328/ml PN3-100: ma 05: 8.8 (2.7; 15.2)	ma 05: 7.5 (-3; 14)	ma 05: 7 (1; 13)	ma 05: 5 (-2; 12)	

[a] CO: Carbon monoxide, cv: Cardiovascular, CVD: Cardiovascular, MA: Mean average, NO2: Nitrogen dioxide, O3: Ozone, PM: Particulate matter, PNC: Particulate number concentration,
RR: relative risk, SO2: Sulfur dioxide, UFP: Ultrafine particle.

Table A3b: Short-term studies with adjustment for co-pollutants, emergency/hospital visits/admissions

Reference	Exposure time window	Outcome	UFP effect w/o co-pollutant adjustment	PM10 adjusted UFP effect	PM2.5 adjusted UFP effect	NO2 adjusted UFP effect	
Evans et al. (2014)	unclear	Number of pediatric asthma visits	**PNC/ORs** Lag 1: 0.89 (0.64; 1.24)per 3,007/ml Lag 4: 1.27 (0.9; 1.79) per 2,088/ml **AccMP/ORs** Lag 1: 0.73 (0.50;1.08) per 874/ml Lag 4: 1.00 (0.71;1.4) per 638/ml	Two-pollutant models using the pollutants shown to be associated with asthma exacerbation (ultrafine particles, carbon monoxide, and ozone). The effect estimates in these models did not differ substantially from those in the single-pollutant models (data not shown)			
Iskandar et al. (2012)	ma 0-4	Hospital admissions due to asthma	**ORs** per 3,812.86/ml: 1.06 (0.98-1.14)	0.99 (0.92; 1.08)	0.99 (0.91; 1.08)	0.97 (0.89; 1.06)	NOx: 1 (0.91; 1.08)
Lanzinger et al. (2016)	ma 2-5	CV hospital admissions	RRs/ UFP per 2,750/ml ma 2-5: 0.3 (-1.7; 2.4)		-0.5 (-2; 1.5)	-0.7 (-2.1; 1)	
	ma 0-5	Resp. hospital admissions	RRs/ UFP per 2,750/ml ma 0-5: 3.4 (-3.2; 7.3)		-3 (-11; 5)	-4 (-85; 2)	
Rosenthal et al. (2013)	lag 0d lag 2d lag 0d	Out-of hospital card. arrest, MI	ORs/ PNC per 10,624/ml: lag 0d: 1.27 (1.05; 1.54) lag 3d: 0.97 (0.80; 1.05) ORs/ AccMP per 1,007/ml lag 0d: 1.19 (1.04; 1.54) lag 2d: 0.96 (0.84; 1.10)		PNC lag 0d: 1.20 lag 3d: 0.99 AccMP lag 0d: 1.02 lag 2d: 0.93 Cis not reported		O3, PNC: lag 0d: 0.89 lag 3d: 1.10 AccMP lag 0d: 1.00 lag 2d: 0.98 Cis not reported
Samoli et al. (2016a)	lag 1d	CV hospital admissions, 15-65y 65y+	0.81 (−0.78; 2.42) −0.07 (−1.27; 1.15)				Adjustment of co-source estimates: Effect estimates of background urban NSD

Reference	Exposure time window	Outcome	UFP effect w/o co-pollutant adjustment	PM10 adjusted UFP effect	PM2.5 adjusted UFP effect	NO2 adjusted UFP effect	
	lag 2d	Respiratory hospital admissions, 0-14y 15-64y 65y+	1.86 (−0.28; 4.05) −1.14 (−2.66; 0.41) −1.09 (−2.42; 0.27)				with either adult CVD or pediatric hospitalizations re-mained robust as did the estimates between nucleation NSD and pediatric hospital ad-missions.
Samoli et al. (2016b)	lag days 0-7	Respiratory hospital admissions	Percentage change per 10,000/ml lag 0: -0.44 (-1.73; 0.87) lag 1: -0.58 (-1.93; 0.79) lag 2: -0.22 (-0.92; 0.38) lag 5: 0.43 (-0.58; 1.45) lag 7: -0.37 (-1.39; 0.66)	lag 0: -0.73 (-2.21; 0.77) lag 1: -1.09 (-2.50; 0.34) lag 2: -0.58 (-1.24; 0.08) lag 5: 0.26 (-0.82; 1.36) lag 7: -0.24 (-1.36; 0.89)	lag 0: -0.51 (-2.12; 1.14) lag 1: -0.70 (-2.39; 1.02) lag 2: -0.65 (-1.77; 0.49) lag 5: 0.33 (-1.17; 1.84) lag 7: -0.68 (-1.96; 0.62)	lag 0: -0.42 (-2.08; 1.28) lag 1: -0.55 (-2.16; 1.09) lag 2: 0.04 (-0.67; 0.75) lag 5: -0.82 (-1.57; -0.07) lag 7: -0.83 (-2.09; 0.45)	O3: lag 0: -0.05 (-1.14; 1.34) lag 1: 0.08 (-1.61; 1.80) lag 2: -0.14 (-0.76; 0.49) lag 5: 0.35 (-0.35; 1.29) lag 7: -0.30 (-1.27; 0.69)

[a] AccMP: Accumulation mode particles, CVD: Cardiovascular, NO2: Nitrogen dioxide, NOx: Nitrogen oxides, NSD: Size distributions of ultrafine particles, O3: Ozone, OR: Odds ratio, PM: Particulate matter, RR: Relative risk, UFP: Ultrafine particle.

Table A3c: Short-term studies with adjustment for co-pollutants, subclinical outcomes

Reference	Exposure time window	Outcome	UFP effect w/o co-pollutant adjustment	PM10 adjusted UFP effect	PM2.5 adjusted UFP effect	NO2 adjusted UFP effect	adjusted for different pollutants
Croft et al. (2017)	48h	Fibrinogen	UFP, percent changes, 1.90 (0.86; 2.95) per 1743/ml	-	2.46 (0.96; 3.96)	-	Delta-C: 2.76 (1.09; 4.42) BC: 2.50 (0.84; 4.16)
	12h	MPO	AccMP, per 452/ml, percent changes −2.80 (−4.68; −0.92)	-	-2.2 (-4.78;0.38)	-	Delta-C: -2.37(-5.18; 0.45) BC: -1.83(-4.44; 0.79)
	96h	MPO	UFP, per 1434/ml, percent changes −5.55 (−8.51; −2.59)		-6.11 (-10.02; -2.20)		Delta-C: -5.77 (-9.99; -1.55) BC: -9.54 (-14.12; -4.95)
Gong et al. (2014)	depending on outcome	FeNO, EBC pH, EBC nitrite, WBC, urinary MDA, 8-OHdG	% changes FeNO, lag 0: 25.34 (12.96; 39.09) EBC pH, lag 1: 1.54 (0.79; 2.28) EBC nitrite, lag 6: 25.64 (16.12; 35.94), WBC, lag 0: 3.5 (1,7) urinary MDA,lag 3: 10.89 (0.56; 22.28 8-OHdG, lag 3: 28.56 (4.08; 59.53			FeNO: 26 (13; 40) EBC pH: similar EBC nitrite: 3 (-2; 18) WBC: similar urinary MDA, lag 3: 6.5 (-4; 17) urine 8-OHdG: 19 (-7; 44)	SO2 (further adjustments: see article) FeNO: 20 (6; 34) EBC pH: slightly lower EBC nitrite: 10 (0; 20) WBC: similar urinary MDA: 8 (-3; 19) urine 8-OHdG: 24 (0; 49)
Han et al. (2016)	ma 8h	FeNO	PNCait 12 (6; 20)		19 (9; 29)	17 (5; 28)	BC: 15 (7; 24) SO2: 11 (1;20)
Janssen et al. (2015)	2h after exposure:	FeNO	Percent change per 23,000/ml PNC after excluding underground: 13.0 (6.0; 21.0) increase	6 (-0.5;12)	8 (-1.5; 18)	17 (6; 27)	O3: 18 (8; 27)

Reference	Exposure time window	Outcome	UFP effect w/o co-pollutant adjustment	PM10 adjusted UFP effect	PM2.5 adjusted UFP effect	NO2 adjusted UFP effect	adjusted for different pollutants
	2h after exposure	IL-6 (nasal)	15.0 (-11; 50) When the underground site was included in the analysis, FeNO and NAL IL-6 were consistently associated with PNC	8 (-18; 39)	11 (-13; 45)	-5 (-31; 30)	O3: 2 (-24,38)
	2h after exposure	FEV1	-1.5**	-1.6**	-1.5**	-0.4	
Li et al. (2016)	lag 1d	FEF 50%	ß-estimates per 5646.4/ml UFP: 0.40 (0.24; 0.56)				O3: 0.25 (0.01; 0.48)
	lag 1d	FEF 75%	UFP: 0.29 (0.19; 0.39)				O3: 0.16 (0.01; 0.30)
	lag 1d	FVC	UFP: 0.08 (-0.01; 0.18) AccMP: 0.00 (-0.10; 0.09)				O3: 0.14 (0.00; 0.28) O3: -0.04 (-0.17; 0.09)
	lag 1d	FEV1	UFP: 0.11 (0.02; 0.20) AccMP: 0.03 (-0.06; 0.12)				O3: 0.13 (-0.01; 0.26) O3: -0.03 (-0.15; 0.10)
Peters et al. (2015)	various	HR SDNN RMSSD	Percent changes per 16,000/ml personal PNC: SDNN, concurrent −0.56 (-1.02; −0.09), HR, lag 0-4 min: 0.23 (0.11; 0.36) lag 5-0 min: 0.16 (0.04; 0.28) RMSSD: estimates close to 0		lag unclear: estimates remain nearly the same		
Pieters et al. (2015)	0d	SBP	ß-estimates (mmHg): PN20-30nm: 6.35 (1.56; 11.47) per 860/cm3 30–50 nm: 1.18 (0.05; 2.31), per 712/ml, 50–70 nm, 0.92 (–0.05; 1.89) per 540/ml, 70–100 nm: 0.86 (0.05; 1.68) per 358/ml, Total UFP: 2.92 (0.30; 5.61) per 1,666/ml	Similar results (see figure 3)			
Rich et al. (2012)	24-47h	TpTe (msec):	UFP: 0.33 (-0.32; 0.98) AccMP: 1.05 (0.28; 1.82) per 897/ml		(AccMP) 1.28 (0.25; 2.31)		AccMP: -0.26 (-1.06; 0.53) UFP: 1.23 (0.29; 2.17)

Reference	Exposure time window	Outcome	UFP effect w/o co-pollutant adjustment	PM10 adjusted UFP effect	PM2.5 adjusted UFP effect	NO2 adjusted UFP effect	adjusted for different pollutants
	0-5h	rMSSD (ms)	UFP: -3.19: (–5.32; –1.05) AccMP: -1.91 (-4.31; 0.49)				AccMP: -3.63 (-6.47; -0.79) UFP: -0.76 (-2.42; 3.94)
	72-95h	HRT (ms/RR)	UFP: 0.06 (-0.43; 0.55) AccMP: -0.67 (-1.18; -0.15)		(AccMP) -0.65 (-1.39; 0.07)		AccMP: 0.62 (0.04; 1.21) UFP: -1.05 (-1.68; -0.42)
	0-5h	SBP (mmHg)	AccMP per 897/ml 0.63 (-0.27; 1.53)		(AccMP) 0.32 (-0.94; 1.57)		
	24-47	Fibrinogen (g/L)	UFP: 0.08 (0.02; 0.14) AccMP: 0.12 (0.04; 0.20)		(AccMP) 0.12 (0.01; 0.23)		AccMP: 0.034 (-0.05; 0.11) UFP: 0.10 (-0.003; 0.19)
Rückerl et al. (2014)	ma 05	CRP	percent change per 5,722/ml PNC (3-100) 12 (2; 23):		3 (-8; 17)		
Rückerl et al. (2016)	ma 05	hsCRP	Percent change per 22.3 mm2/cm³ SC(DCPS) ma 5: 29.8 [15.9;45.3] per 168.9 mm²/cm³ SC10-800, ma 5: 9.2 (0.8; 18.3) per 0.06 SC3-10nm, ma 5: 9.6 (1.9; 18.0) per 5.7 SC30-50nm, ma 5, 3.2 (-3.9; 10.9) per 24.7 SC50-100nm, ma 5, 4.2 (-2.5; 11.4),	similar results, slightly weaker with SC(DCPS)			only adjusted for RHO2.5: apparent particle density of particulate matter with aerodynamic diameter <2.5µm and <10µm, respectively
	ma 05	other outcomes		MPO and IL-6 associations similar. Few associations slightly			

Reference	Exposure time window	Outcome	UFP effect w/o co-pollutant adjustment	PM10 adjusted UFP effect	PM2.5 adjusted UFP effect	NO2 adjusted UFP effect	adjusted for different pollutants
				stronger, e.g. IL-6 with LC(EAD) and SC(DCPS), , some slightly weaker. For fibrinogen, associations were somewhat inconclusive for lag 4: associations for both, LC (EAD) and SC(DCPS) turned from positive to negative, when adjusted for ρ2.5 or ρ10.			
Steenhof et al. (2013)	pre/ 2 h after exposure	NAL IL-6	Changes in ß-estimates per 32,906/ml NAL IL-6: −2.2 (p< 0.05)	−3.6	−3.6	−13.3	
			NAL protein: 7.9 (p< 0.05)	-	7.8	−1.3	
			NAL lactoferrin: 4.3 (p< 0.05),	−0.8	1.2	0.6	
			serum IL-6: 6.3 (p< 0.05)	7.2	6.8	5.8	
Steenhof et al. (2014)	2 h & 18 after exposure	total WBC	2h: -2.2 (-5.3; 1.0), 18h after expo: -1.4 (-4.8; 2.2);	-2.71 (p <0.1) -2.00	-2.50 -1.70	-2.04 -1.08	
	2 h & 18 after exposure	Neutrophils	2h: -1.3 (-6.2; 3.9) 18h: -0.46	-1.97 -0.76	-1.70 -0.57	-2.01 -1.03	
	2 h & 18 after exposure	Monocytes	2h: -0.31 18h: 3.4 (-1.0; 7.9)	-0.44 2.69	-0.48 3.04	-0.13 1.76	

Reference	Exposure time window	Outcome	UFP effect w/o co-pollutant adjustment	PM10 adjusted UFP effect	PM2.5 adjusted UFP effect	NO2 adjusted UFP effect	adjusted for different pollutants
Strak et al. (2012)	immediately after exposure	FeNO	11.2 (p < 0.05)	11.3 (p < 0.05)	11.3 (p < 0.05)		O3: 12.0 (p < 0.05)
	immediately after exposure	FVC	-1.19 (p < 0.05)	-1.26 (p < 0.05)	-1.26 (p < 0.05)	-1.19 (p < 0.05)	O3: -1.15 (p < 0.05)
Strak et al. (2013)	25h post-pre	Hs-CRP	Percent changes per 32,906/ml -4.31 (-14.35; 6.92)	-2.75 (-15.66; 5.32)	-5.23 (-15.15; 5.85)	-11.23* (-21.75; 0.71)	EC(fine): -9.91 (-20.56; 2.17) OC (coarse): -2.23 (-12.51; 9.26)
		Fibrinogen	-0.92 (-2.98; 1.19)	-1.12 (-3.19; 0.99)	-1.06 (-3.11; 1.05)	-1.40 (-3.77; 1.04)	
		Platelet counts	-1.15 (-2.69; 0.40),	-1.26 (-2.80; 0.32)	-1.21 (-2.75; 0.36)	-0.51 (-2.29; 1.30)	
		von-Willebrandt-Faktor:	-0.04 (-2.80; 2.80).	0.16 (-2.70; 3.09)	0.28 (-2.56; 3.20)	-0.73 (-3.88; 2.52)	
Strak et al. (2013)	2h after exposure (t9-t0),	endogenous thrombin potential in FXII-mediated thrombin generation pathway	Percent changes per 32,906/ml: all sites: (t9–t0): 5.83 (−39.62; 51.29), outdoor sites (t9–t0): −0.70 (−52.00; 50.60)	all sites: 3.17 (−43.10; 49.44) outdoor sites: −0.70 (−52.00, 50.60)	all sites: 3.40 (−42.14; 48.95) outdoor sites (t9–t0): 7.80 (−45.65; 61.25)	all sites: −27.76 (−79.32; 23.81) outdoor sites: 8.79 (−44.62; 62.20)	
	next morning (t25-t0)		all sites (t25–t0): −72.40 (−128.56, −16.24), outdoor sites (t25–t0): −66.59 (−124.78, −8.40)	all sites: −71.38 (−129.02, −13.73) outdoor: -80.02 (−139.74, −20.29)	all sites: −71.38 (−129.02; −13.73) outdoor: -79.46 ((−139.10; −19.82)	all sites: −47.39 (−114.60, 19.82) outdoor: -46.48 ((−112.47; 19.51)	
Sun et al. (2015)	ma 04	SDNN	Percent change PNC5-560: -7.9 (-9.7;-6.1) PNC 10-20: -7 (-8.9;-5.1)	-	-	PNC5-560: -7.73 (-9.57; -5.85) PNC 10-20: -7.21 (-	O3: PNC5-560: -7.47 (-

Reference	Exposure time window	Outcome	UFP effect w/o co-pollutant adjustment	PM10 adjusted UFP effect	PM2.5 adjusted UFP effect	NO2 adjusted UFP effect	adjusted for different pollutants
			PNC 20-50: -6.6 (-8.1;-5) PNC 50-100: -5.4 (-7.3;-3.4) PNC 100-200: -3 (-4.6;-1.3) PNC 200-560: -0.45 (-2.43;1.56).			9.14; -5.24) PNC 20-50: -6.36 (-7.92; -4.77) PNC 50-100: -5.65 (-7.69; -3.56) PNC 100-200: -2.53 (-4.26; -0.77) PNC 200-560: 0.09 (-1.96;2.18)	9.65; -5.24) PNC 10-20: -6.73 (-8.65; -4.77) PNC 20-50: -6.07 (-7.77; -4.33) PNC 50-100: -3.49 (-6.03;-0.89) PNC 100-200: 0.3 (-1.84; 2.49) PNC 200-560: 3.25 (0.97; 5.59)
Weichenthal et al. (2014)	lag 3h	RHI	Percent changes per 10,850/ml: -4.91 (-9.31; -0.512)	+ Adjustment for Exposures during Previous Visits and Regional Air Quality	-4.74 (-9.21; -0.26)	-5.03 (-9.52; -0.54)	-4.62 (-9.07; -0.168)
	lag 3h	SBP	0.377 (-0.900; 1.65)		0.421 (-0.862; 1.70)	0.590 (-0.683; 1.86)	O3: 0.565 (-0.698; 1.83)
	lag 3h	DBP	1.61 (-0.155; 3.38)		1.65 (-0.115; 3.42	2.00 (0.253; 3.74)	O3: 1.88 (0.126; 3.64)
	lag 3h	SDNN	9.86 (0.245; 19.5)		3.74 (0.346; 7.14)	4.20 (0.855; 7.55)	O3: 4.05 (0.721; 7.38)
Wu et al. (2012)			PM0.25, SDNN: -4.7 (−14.5; 6.2), r-MSSD: -5.1 (−12.4; 3.0), HF: -5.7 (−16.5; 6.5), LF: -4.8 (−15.1; 6.8), LF/HF: 1.0 (−2.8; 5.0)	Appendix not available, waiting for author's response			
Zhang et al. (2013)	lag 3d	HR	Percent changes per 6,572/ml 0.5 (0.1; 1.0)			0.7 (-0.2; 1.3)	O3: 0.6 ((0.2; 1.2)

Reference	Exposure time window	Outcome	UFP effect w/o co-pollutant adjustment	PM10 adjusted UFP effect	PM2.5 adjusted UFP effect	NO2 adjusted UFP effect	adjusted for different pollutants
	lag 3d	HF	-5 (-1; -8)			(lag 1): -0.5 (-9; 1)	O3 (lag 4) -7 (-9; -3)
		CRP, fibrinogen, BCC & differentials, 8OHdG, FeNO, EBC pH, nitrate, nitrite, +nitrate, 8-isoprostane), CD62P sCD40L], platelet aggregation, vWF, BP				mostly similar	mostly similar
Zhang et al. (2016a)		FeNO	Percent changes, FeNO: ma 5: stronger estimated associations for ultrafine PM0.18 than larger size-fractions for total mass PM 0.18: 3.0 (0.7; 5.3) per 1.1 µg/m³ AccMP: -0.8 (-3.5; 1.9) per 4.0 µg/m³ (various outcomes (elements and PAHs in PM0.18) in figure 1&2 and supplementary tables) MDA: positively associated with total PM0.18 mass	The estimates of association from two-pollutant models of O3 with a primary air pollutant (BC, NOx and PAHs) became largely nonsignificant with effect estimates attenuating toward the null for airway biomarkers, except for associations between FeNO and PAHs in PM0.18 that remained significant. The results for the systemic biomarkers were similar using two-pollutant			

Reference	Exposure time window	Outcome	UFP effect w/o co-pollutant adjustment	PM10 adjusted UFP effect	PM2.5 adjusted UFP effect	NO2 adjusted UFP effect	adjusted for different pollutants
				models and single-pollutant models (data not shown)			
Zhang et al. (2016b)	ma 5d	RHI	PM0.18 per 1.13 µg/m3 -0.01 (-0.05; 3) (only figures)				O3: 0.15 (0.04; 0.06)

[a] 8-OHdG: Urinary 8-hydroxy-2'-deoxyguanosine, AccMP: Accumulation mode particles, BC: Black carbon, BCC: Blood cell counts, BP: Blood pressure, CD62P: P-selectin (protein) sCD40L: soluble CD40 ligand, CRP: C-reactive protein, DBP: Diastolic blood pressure, Delta-C: Estimate of wood smoke pollution, EBC: Exhaled breath condensate pH, EC(fine): Elemental carbon, FeNO: Fractional exhaled nitric oxide, FEF 50-75: Forced expiratory flow at 50-75% of vital capacity, FEV1: Forced expiratory volume in 1 second, FVC: Forced vital capacity, HF: High frequency, HR: Heart rate, HRT: Heart rate turbulence, hs-CRP: High-sensitive C-reactive protein, IL: Interleukin, LC(EAD): Particle length concentration measured by Electric Aerosol Detector, MA: Mean average, MDA: Malondialdehyde, MPO: Myeloperoxidase, NAL: Nasal lavage, NO2: Nitrogen dioxide, NOx: Nitrogen oxides, O3: Ozone, OC(coarse), PAHs: Polycyclic aromatic hydrocarbons, PM: Particulate matter, PNCait: PNC Aitken mode particles, RHI: Reactive hyperemia index, RMSSD: Root mean square of the sucessive differences in ms., SBP: systolic blood pressure, SC(DCPS): Particle surface concentration measured by Diffusion charging particle sensor, SDNN: Standard deviation of normal-to-normal intervals, SO2: Sulfur dioxide, TpTe: Time from peak to end of T-wave, UFP: Ultrafine particle, vWF: Von Willebrand Factor, WBC: White blood cell counts.

Table A4a: Long-term studies with adjustment for co-pollutants, mortality

Reference	Exposure time window	Outcome	UFP effect w/o co-pollutant adjustment	PM10 adjusted UFP effect	PM2.5 adjusted UFP effect	NO2 adjusted UFP effect	
Ostro et al. (2015)		IHD mortality	HRs per 0.969 µg/ml: 1.10 (1.02; 1.18),				CU: 1.39 (1.05; 1.83), adj. for other consituents: no sign. Assoc.

[a] CU: Copper, HR: Heart rate, IHD: Ischaemic heart disease, NO2: Nitrogen dioxide, PM: Particulate matter, UFP: Ultrafine particle.

Table A4b: Long-term studies with adjustment for co-pollutants, subclincal outcomes

Reference	Exposure time window	Increment	UFP effect w/o co-pollutant adjustment	PM10 adjusted UFP effect	PM2.5 adjusted UFP effect	NO2 adjusted UFP effect
Aguilera et al. (2016)	2011-2012	CIMT	PNC (main model): 2.06 (0.03; 4.10) LDSA (main model): 2.32 (0.23; 4.48)			LDSA: 3.35 (–0.20; 6.90) PNC: –1.53 (–4.99; 1.93)

[a] CIMT: Carotid intima-media thickness, LDSA: Lung deposited surface area, NO2: Nitrogen dioxide, PM: Particulate matter, UFP: Ultrafine particle.

Table A5a: Objective quality indicators, short-term studies, mortality

Reference	Study population specified	Sample type of study population	Response Rate [%]	Subjects recruited from same or similar populations	Subjects recruited from same time period	Sample representative for general population	Lost to follow-up after Baseline provided	Losses to follow-up likely to introduce bias	exposure assignment	Complete or partial residential address historyprovided	Description of size ranges	QA/QC for UFP measures described?	Exposure assessment implemented consistently across all study participants	Outcome measures clearly defined and implemented consistently	Outcome assessors blinded to exposure status resp. Case-control status of participants	Analysis adjusted for other air pollutants
Lanzinger et al. (2016)	Yes	NA	NA	No	No	CR	NA	NA	NA	NA	Yes	Yes	Yes	Yes	NA	Yes
Leitte et al. (2016)	Yes	NA	NA	Yes	Yes	CR	NA	NA	NA	NA	Yes	No	Yes	Yes	Yes	Yes
Meng et al. (2013)	Yes	NA	NA	Yes	Yes	CR	NA	NA	NA	NA	Yes	Yes	Yes	Yes	Yes	Yes
Samoli et al (2016a)	Yes	CS	NA	Yes	Yes	CR	NA	NA	NA	NA	Yes	No	Yes	Yes	Yes	Yes
Stafoggia et al.	Yes	NA	NA	Yes	No	CR	NA		City	NA	Yes	Yes	No	Yes	Yes	Yes

Refer-ence	Study population specified	Sample type of study popula-tion	Response Rate [%]	Subjects recruited from same or similar populations	Subjects recruited from same time period	Sample representative for general population	Lost to follow-up after Base-line provided	Losses to follow-up likely to introduce bias	exposure assignment	Complete or partial residential address historyprovided	Description of size ranges	QA/QC for UFP measures de-scribed?	Exposure assessment imple-mented consistently across all study participants	Outcome measures clearly defined and implemented consistently	Outcome assessors blinded to exposure status resp. Case-control status of participants	Analysis adjusted for other air pollutants
(2017)																
Su et al. (2015)	Yes	NA	NA	Yes	Yes	CR	NA	NA	NA	NA	Yes	No	Yes	Yes	Yes	Yes
Wolf et al. 2015	Yes	Other	NA	Yes	Yes	CR	NA	NA	NA	NA	Yes	No	No	Yes	Yes	No

CS: Convenience Sample, CR: Completely representative, NA: Not applicable,

Table A5b: Objective quality indicators, short-term studies, morbidity

Reference	Study population specified	Sample type of study population	Response Rate [%]	Subjects recruited from same or similar populations	Subjects recruited from same time period	Sample representative for general population	Lost to follow-up after Baseline provided	Losses to follow-up likely to introduce bias	Exposure assignment	Complete or partial residential address historyprovided a	Description of size ranges	QA/QC for UFP measures described?	Exposure assessment implemented consistently across all	Outcome measures clearly defined and implemented	Outcome assessors blinded to exposure status resp. Case-control status	Analysis adjusted for other air pollutants
Cole-Hunter et al. (2013)	Not specified/ RG	NR/NR	NA	Yes	Yes	SG	NA	NA	Mobile personal	NA	Yes	Yes	Yes	Yes	NR/ NR	No
Karakatsani (2012)	Yes	CS	NA	No	Yes	SG	NR/NR	Cannot det.	City	NA	Total	Yes	Yes	Yes	Yes	No
Langrish et al. (2012)	Yes	CS	NA	Yes	Yes	SG	NA	NA	Mobile personal	NA	No	No	Yes	Yes	NA	No
Link et al. (2013)	Yes	CS	NA	Yes	Yes	SG	Yes	Yes	NA	NA	No	No	Yes	Yes	Yes	No
Mehta et al. (2015)	Yes	Other	Subgoup of larger cohort	Yes	Yes	SG	Yes	Yes	City	NR/ RG	Yes	Yes	Yes	Yes	Yes	No

Reference	Study population specified	Sample type of study population	Response Rate [%]	Subjects recruited from same or similar populations	Subjects recruited from same time period	Sample representative for general population	Lost to follow-up after Baseline provided	Losses to follow-up likely to introduce bias	Exposure assignment	Complete or partial residential address historyprovided a	Description of size ranges	QA/QC for UFP measures described?	Exposure assessment implemented consistently across all	Outcome measures clearly defined and implemented	Outcome assessors blinded to exposure status resp. Case-control status	Analysis adjusted for other air pollutants
Wang et al. (2014)	Yes	RS	NR/ RG	Yes	No	SG	NR/ NR	CD	NA	NA	No	No	Yes	Yes	Yes	No
Wolf et al. 2015	Yes	Other	NA	Yes	Yes	CR	NA	NA	NA	NA	Yes	No	Yes	Yes	Yes	No

CD: Cannot determin , CS: Convenience Sample, CR: Completely representative, SR: Somewhat representative, SG: selected group, NA: Not applicable, NR/NR: Not reported/ no reference given, NR/RG: Not reported/ reference given

Table A5c: Objective quality indicators, short-term studies, emergency/hospital admissions

Reference	Study population specified	Sample type of study population	Response Rate [%]	Subjects recruited from same or similar populations	Subjects recruited from same time period	Sample representative for general population	Lost to follow-up after Baseline provided	Losses to follow-up likely to introduce bias	Exposure assignment	Complete or partial residential address historyprovided	Description of size ranges	QA/QC for UFP measures described?	Exposure assessment implemented consistently across all study participants	Outcome measures clearly defined and implemented consistently	Outcome assessors blinded to exposure status resp. Case-control status of participants	Analysis adjusted for other air pollutants
Delfino et al. (2014)	Yes	CS	NA	Yes	Yes	SR	NA	NA	Geo-coded addresses	No complete residential address history	NR/RG	No	Yes	Yes	Yes	No
Diaz-Robles et al. (2014)	NS/ NR	CS	NA	Yes	Yes	SR	NA	NA	NA	NA	Yes	No	Yes	Yes	Yes	No
Evans et al. (2014)	Yes	CS	NA	Yes	Yes	SG	NR/NR	No	NA	NA	Yes	NR/NR	Yes	Yes	NA	Yes
Gardner et al. (2014)	Yes	NA	NA	Yes	Yes	SG	NA	NA	City	NA	Yes	Yes	Yes	Yes	Yes	No

Reference	Study population specified	Sample type of study population	Response Rate [%]	Subjects recruited from same or similar populations	Subjects recruited from same time period	Sample representative for general population	Lost to follow-up after Baseline provided	Losses to follow-up likely to introduce bias	Exposure assignment	Complete or partial residential address historyprovided	Description of size ranges	QA/QC for UFP measures described?	Exposure assessment implemented consistently across all study participants	Outcome measures clearly defined and implemented consistently	Outcome assessors blinded to exposure status resp. Case-control status of participants	Analysis adjusted for other air pollutants
Iskandar et al. (2012)	Yes	CS	NA	Yes	Yes	SR	NA	NA	NA	NA	Yes	NR/NR	Yes	Yes	NA	Yes
Lanzinger et al. (2016)	Yes	NA	NA	No	No	CR	NA	NA	NA	NA	Yes	Yes	Yes	Yes	NA	Yes
Liu et al. (2013)	Yes	NA	NA	Yes	Yes	CR	NA	NA	NA	NA	Yes	No	Yes	Yes	NA	No
Rosenthal et al. (2013)	Yes	NA	NA	Yes	Yes	SG	NA	NA	NA	NA	Yes	No	Yes	Yes	Yes	Yes
Samoli et al (2016a)	Yes	CS	NA	Yes	Yes	CR	NA	NA	NA	NA	Yes	No	Yes	Yes	Yes	Yes

Reference	Study population specified	Sample type of study population	Response Rate [%]	Subjects recruited from same or similar populations	Subjects recruited from same time period	Sample representative for general population	Lost to follow-up after Baseline provided	Losses to follow-up likely to introduce bias	Exposure assignment	Complete or partial residential address historyprovided	Description of size ranges	QA/QC for UFP measures described?	Exposure assessment implemented consistently across all study participants	Outcome measures clearly defined and implemented consistently	Outcome assessors blinded to exposure status resp. Case-control status of participants	Analysis adjusted for other air pollutants
Samoli et al. (2016b)	Yes	NA	NA	Yes	No	CR	NA	NA	NA	NA	Yes	Partly	No	Yes	Yes	Yes
Wichmann et al. (2013)	Yes	NA	NA	Yes	Yes	SG	NA	NA	NA	NA	Yes	No	Yes	Yes	NA	No

CR: Completely representative, CS: Convenience Sample, NA: Not applicable, NR/NR: Not reported/ no reference given, NR/RG: Not reported/ reference given, NS/NR: Not specified/ no reference given, SG: selected group, SR: Somewhat representative.

Table A5d: Objective quality indicators, short-term studies, subclinical outcomes

Reference	Study population specified	Sample type of study population	Response Rate [%]	Subjects recruited from same or similar popula-tions	Subjects recruited from same time period	Sample representative for general population	Lost to follow-up after Baseline provided	Losses to follow-up likely to introduce bias	Exposure assignment	Complete or partial residential address history-provided a	Description of size ranges	QA/QC for UFP measures described?	Exposure assessment implemented consistently across all participants	Outcome measures clearly defined and imple-mented	Outcome assessors blinded to exposure status resp. Case-control status of participants	Analysis adjusted for other air pollutants
Bartell et al. (2013)	Yes	CS	NA	Yes	Yes	SG	Yes	No	Micro-environ ron-ments	NA	Yes	No	Yes	Yes	Yes	No
Bind et al. (2016)	Yes	CS	NA	Yes	Yes	SG	NR/ RG	Yes	NA	Yes	Yes	No	Yes	Yes	Yes	No
Bos et al. (2011)	Not speci-fied/ RG	CS	NA	Yes	Yes	SG	NA	NA	Mobile per-sonal	NA	No	No	Yes	Yes	No	No
Bos et al. (2013)	Yes	CS	NA	No	Yes	SG	NA	NA	NA	NA	Yes	No	Yes	Yes	No	No

Reference	Study population specified	Sample type of study population	Response Rate [%]	Subjects recruited from same or similar populations	Subjects recruited from same time period	Sample representative for general population	Lost to follow-up after Baseline provided	Losses to follow-up likely to introduce bias	Exposure assignment	Complete or partial residential address history-provided a	Description of size ranges	QA/QC for UFP measures described?	Exposure assessment implemented consistently across all participants	Outcome measures clearly defined and implemented	Outcome assessors blinded to exposure status resp. Case-control status of participants	Analysis adjusted for other air pollutants
Chung M. et al. (2015)	Yes	Random + CS	NR/ RG	Yes	Yes	SG	NR/ RG	CD	NA	NA	No	NR/ NR	Yes	Yes	Yes	No
Cole-Hunter et al. (2013)	Not specified/ RG	NR/NR	NA	Yes	Yes	SG	NA	NA	Mobile personal	NA	Yes	Yes	Yes	Yes	NR/ NR	No
Cole-Hunter et al. (2016)	Yes	NR/ NR	NR/ NR	Yes	Yes	SG	NR/ NR	NA	Mobile personal	NA	Yes	Yes	Yes	Yes	No	No
Croft et al. (2017)	Yes	CS	NA	Yes	Yes	SG	NA	NA	NA	NA	Yes	No	Yes	Yes	Yes	Yes
Frampton et al. (2012)	Yes	NR/ NR	NR/ NR	Yes	Yes	SG	Yes	No	City	NA	Yes	Yes	Yes	Yes	Yes	No

Reference	Study population specified	Sample type of study population	Response Rate [%]	Subjects recruited from same or similar popula-tions	Subjects recruited from same time period	Sample representative for general population	Lost to follow-up after Baseline provided	Losses to follow-up likely to introduce bias	Exposure assignment	Complete or partial residential address history-provided a	Description of size ranges	QA/QC for UFP measures described?	Exposure assessment implemented consistently across all participants	Outcome measures clearly defined and imple-mented	Outcome assessors blinded to exposure status resp. Case-control status of participants	Analysis adjusted for other air pollutants
Fuller et al. (2015)	Yes	NR/ NR	NA	No	NR/ NR	SR	NA	NA	city, geo-coded adress	NR/ NR	No	No	Yes	Yes	NA	No
Gong et al. (2014)	Yes	NR/NR	NR/NR	Yes	Yes	SG	NR/RG	NA	City	No	Yes	Yes	Yes	Yes	Yes	Yes
Hampel et al. (2012)	Yes	NR/ NR	NA	Yes	Yes	SG	NA	Yes	City	NR/ NR	Yes	No	Yes	Yes	NA	No
Hampel et al. (2014)	Yes	NA	NA	Yes	NA	SG	NA	NA	Mobile per-sonal	NA	Yes	No	Yes	Yes	NR/ NR	NR/ NR
Han et al. (2016)	Yes	CS	NA	Yes	Yes	SG	NR/ NR	CD	NA	NA	Yes.	Yes	Yes	Yes	Yes	Yes

Reference	Study population specified	Sample type of study population	Response Rate [%]	Subjects recruited from same or similar popula-tions	Subjects recruited from same time period	Sample representative for general population	Lost to follow-up after Baseline provided	Losses to follow-up likely to introduce bias	Exposure assignment	Complete or partial residential address history-provided a	Description of size ranges	QA/QC for UFP measures described?	Exposure assessment implemented consistently across all participants	Outcome measures clearly defined and imple-mented	Outcome assessors blinded to exposure status resp. Case-control status of participants	Analysis adjusted for other air pollutants
Hoff-mann et al. (2012)	Yes	CS	NA	Yes	Yes	SG	NR/ NR	No	NA	NA	No	No	Yes	Yes	Yes	No
Huttun en et al. (2012)	Yes	CS	84	Yes	Yes	SG	NR/ NR	No	NA	NA	Yes	No	Yes	Yes	Yes	No
Janssen et al. (2015)	Yes	CS	NA	Yes	Yes	SG	NA	NA	Micro-environ ron-ments	NA	NO	Yes	Yes	Yes	Yes	Yes
Jarjour et al. (2013)	Yes	CS	NA	Yes	Yes	SG	NA	NA	Mobile per-sonal	NA	Yes	Yes	Yes	Yes	No	No
Ka-rottki et al (2015)	Yes	CS	NA	Yes	Yes	SG	Yes	No	Micro-environ ron-ments,	NR/ NR	Yes	No	Yes	Yes	Yes	No

Reference	Study population specified	Sample type of study population	Response Rate [%]	Subjects recruited from same or similar populations	Subjects recruited from same time period	Sample representative for general population	Lost to follow-up after Baseline provided	Losses to follow-up likely to introduce bias	Exposure assignment	Complete or partial residential address history-provided a	Description of size ranges	QA/QC for UFP measures described?	Exposure assessment implemented consistently across all participants	Outcome measures clearly defined and implemented	Outcome assessors blinded to exposure status resp. Case-control status of participants	Analysis adjusted for other air pollutants
									NA							
Karottki et al. (2014)	Yes	NA	NA	Yes	Yes	SG	NA	NA	Micro-environments	Yes, complete RH	Yes	Yes	Yes	Yes	Yes	No
Kubesch et al. (2015)	Yes	CS	NA	NR/ NR	Yes	SG	Yes	No	Micro-environments	NA	Yes	Yes	Yes	Yes	No	No
Langrish et al. (2012)	Yes	CS	NA	Yes	Yes	SG	NA	NA	Mobile personal	NA	No	No	Yes	Yes	NA	No
Laumbach et al. (2014)	Yes	CS	NA	Yes	Yes	SG	NR/ NR	CD	Mobile personal	Yes	Yes	No	Yes	Yes	NR/ NR	No

Reference	Study population specified	Sample type of study population	Response Rate [%]	Subjects recruited from same or similar populations	Subjects recruited from same time period	Sample representative for general population	Lost to follow-up after Baseline provided	Losses to follow-up likely to introduce bias	Exposure assignment	Complete or partial residential address history-provided a	Description of size ranges	QA/QC for UFP measures described?	Exposure assessment implemented consistently across all participants	Outcome measures clearly defined and implemented	Outcome assessors blinded to exposure status resp. Case-control status of participants	Analysis adjusted for other air pollutants
Li et al. (2016)	Yes	Random + CS	NR/ NR	Yes	Yes	SG	NR/ NR	CD	City	NA	Yes	No	Yes	Yes	Yes	Yes
Ljungman et al. (2014)	Yes	RS	NR/ NR	No	No	SR	NA	NA	NA	NA	No	No	Yes	Yes	Yes	No
Manney et al. (2012)	Yes	CS	NA	No	Yes	SG	NR/ NR	CD	Geo-coded addresses	NA	Yes	Yes	Yes	Yes	Yes	No
Mehta et al. (2014)	Yes	Other	Subgoup of larger cohort	Yes	Yes	SG	Yes	Yes	City	NR/ RG	Yes	Yes	Yes	Yes	Yes	No
Mirabelli et al. (2015)	Yes	CS	NA	Yes	Yes	SG	NA	NA	Mobile personal	NA	No	Yes	Yes	Yes	No	No

Reference	Study population specified	Sample type of study population	Response Rate [%]	Subjects recruited from same or similar popula-tions	Subjects recruited from same time period	Sample representative for general population	Lost to follow-up after Baseline provided	Losses to follow-up likely to introduce bias	Exposure assignment	Complete or partial residential address history-provided a	Description of size ranges	QA/QC for UFP measures described?	Exposure assessment implemented consistently across all participants	Outcome measures clearly defined and imple-mented	Outcome assessors blinded to exposure status resp. Case-control status of participants	Analysis adjusted for other air pollutants
Olsen et al. (2014)	Yes	Ran-dom + CS	20	Yes	Yes	SR	NA	NA	Micro-environ ron-ments, Mobile per-sonal	Yes	Yes	Yes	Yes	Yes	Yes	No
Park et al. (2017)	Yes	CS	NA	No	Yes	SG	NA	NA	Mobile per-sonal	NA	No	Yes	Yes	Yes	No	No
Peng et al. (2016)	Yes	CS	NA	Yes	Yes	SG	NR/ NR	No	NA	NA	No	No	Yes	Yes	Yes	No
Peters et al. (2015)	Yes	NR/ NR	NR/ NR	Yes	Yes	SG	NR/ NR	NA	NA	NA	Yes	Yes	Yes	Yes	NR/ NR	Yes

Reference	Study population specified	Sample type of study population	Response Rate [%]	Subjects recruited from same or similar popula-tions	Subjects recruited from same time period	Sample representative for general population	Lost to follow-up after Baseline provided	Losses to follow-up likely to introduce bias	Exposure assignment	Complete or partial residential address history-provided a	Description of size ranges	QA/QC for UFP measures described?	Exposure assessment implemented consistently across all participants	Outcome measures clearly defined and imple-mented	Outcome assessors blinded to exposure status resp. Case-control status of participants	Analysis adjusted for other air pollutants
Pieters et al. (2015)	Yes	CS	NA	No	Yes	SG	Yes	No	Micro-environ ron-ments	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Rich et al. (2012)	Yes	CS	NA	Yes	Yes	SG	NR/ NR	NA	NA	NA	Yes	No	Yes	Yes	Yes	Yes
Rückerl et al. (2014)	Yes	Other	NA	Yes	Yes	SG	NR/ NR	NA	NA	NA	Yes	No	Yes	Yes	Yes	Yes
Rückerl et al. (2016)	Yes	Other	NA	Yes	Yes	SG	NR/ NR	NA	NA	NA	Yes	Yes	Yes	Yes	Yes	Yes
Sarnat et al. (2014)	Yes	CS	NA	Yes	Yes	SG	Yes	No	Mobile per-sonal	NA	No	No	Yes	Yes	Yes	NR/ NR
Shutt et al. (2017)	Yes	CS	NA	Yes	Yes	SR	NA	NA	NA	NA	Yes	No	Yes	Yes	No	No

Reference	Study population specified	Sample type of study population	Response Rate [%]	Subjects recruited from same or similar popula-tions	Subjects recruited from same time period	Sample representative for general population	Lost to follow-up after Baseline provided	Losses to follow-up likely to introduce bias	Exposure assignment	Complete or partial residential address history-provided a	Description of size ranges	QA/QC for UFP measures described?	Exposure assessment implemented consistently across all participants	Outcome measures clearly defined and imple-mented	Outcome assessors blinded to exposure status resp. Case-control status of participants	Analysis adjusted for other air pollutants
Song et al. (2013a)	Yes	CS	NA	Yes	Yes	SG	Yes	No	Micro-environ ron-ments	NA	Yes	No	Yes	Yes	Yes	No
Song et al. (2013b)	Yes	CS	NA	Yes	Yes	SG	NR/ NR	NA	Micro-environ ron-ments	NA	Yes	No	Yes	Yes	Yes	No
Steen-hof et al. (2013)	Yes	CS	NA	Yes	Yes	SG	NA	NA	Micro-environ ron-ments	NA	No	No	Yes	Yes	No	Yes
Steen-hof et al. (2014)	Yes	CS	NA	Yes	Yes	SG	NA	NA	Micro-environ ron-ments	NA	Yes	Yes	Yes	Yes	Yes	Yes
Strak et al. (2012)	Yes	CS	NA	Yes	Yes	SG	NA	NA	Micro-environ ron-ments	NA	No	No	Yes	Yes	No	Yes

Reference	Study population specified	Sample type of study population	Response Rate [%]	Subjects recruited from same or similar popula-tions	Subjects recruited from same time period	Sample representative for general population	Lost to follow-up after Baseline provided	Losses to follow-up likely to introduce bias	Exposure assignment	Complete or partial residential address history-provided [a]	Description of size ranges	QA/QC for UFP measures described?	Exposure assessment implemented consistently across all participants	Outcome measures clearly defined and imple-mented	Outcome assessors blinded to exposure status resp. Case-control status of participants	Analysis adjusted for other air pollutants
Strak et al. (2013a)	Yes	CS	NA	Yes	Yes	SG	NA	NA	Micro-environ ron-ments	NA	Yes	No	Yes	Yes	No	Yes
Strak et al. (2013b)	Yes	CS	NA	Yes	Yes	SG	NA	NA	Micro-environ ron-ments	NA	No	No	Yes	Yes	No	Yes
Sun et al. (2015)	Yes	CS	NA	Yes	Yes	SG	NR/ NR	CD	NA	NA	Yes	No	Yes	Yes	Yes	Yes
Wang et al. (2016)	Yes	CS	NA	Yes	Yes	SG	NR/ NR	NA	NA	NA	Yes	No	Yes	Yes	Yes	No
Weiche nthal et al. (2014)	Yes	CS	NA	Yes	Yes	SG	NA	NA	Mobile per-sonal	NA	Yes	No	Yes	Yes	No	Yes

Reference	Study population specified	Sample type of study population	Response Rate [%]	Subjects recruited from same or similar populations	Subjects recruited from same time period	Sample representative for general population	Lost to follow-up after Baseline provided	Losses to follow-up likely to introduce bias	Exposure assignment	Complete or partial residential address history-provided [a]	Description of size ranges	QA/QC for UFP measures described?	Exposure assessment implemented consistently across all participants	Outcome measures clearly defined and implemented	Outcome assessors blinded to exposure status resp. Case-control status of participants	Analysis adjusted for other air pollutants
Wittkopp et al. (2013)	Yes	CS	NA	Yes	Yes	SG	NR/ NR		Retirement community	NA	Yes	No	Yes	Yes	NA	No
Wu et al. (2012)	Yes	CS	NA	Yes	Yes	SG	NR/ NR	NA	Mobile personal	NA	Yes	No	Yes	Yes	Yes	Yes
Zanobetti et al. 2014	Yes	CS	NA	Yes	Yes	SG	NR/ NR	No	NA	NA	No	No	Yes	Yes	Yes	NR/ NR
Zhang et al. (2013)	Yes	CS	NA	Yes	Yes	SG	Yes	No	NA	NA	Yes	Yes	Yes	Yes	Yes	Yes
Zhang et al. (2016a)	Yes	CS	NA	Yes	Yes	SG	NR/ NR	CD	Mobile personal	NA	Yes	Yes	Yes	Yes	Yes	Yes

Reference	Study population specified	Sample type of study population	Response Rate [%]	Subjects recruited from same or similar populations	Subjects recruited from same time period	Sample representative for general population	Lost to follow-up after Baseline provided	Losses to follow-up likely to introduce bias	Exposure assignment	Complete or partial residential address history-provided a	Description of size ranges	QA/QC for UFP measures described?	Exposure assessment implemented consistently across all participants	Outcome measures clearly defined and implemented	Outcome assessors blinded to exposure status resp. Case-control status of participants	Analysis adjusted for other air pollutants
Zhang et al. (2016b)	Yes	CS	NA	NR/ NR	Yes	SG	NR/ NR	CD	NA	NA	Yes	Yes	Yes	Yes	Yes	Yes

CD: Cannot determine, CR: Completely representative, CS: Convenience Sample, NA: Not applicable, NR/NR: Not reported/ no reference given, NR/RG: Not reported/ reference given, SG: selected group, SR: Somewhat representative.

Table A6a: Objective quality indicators, long-term studies, mortality

Reference	Study population specified	Sample type of study population	Response Rate [%]	Subjects recruited from same or similar populations	Subjects recruited from same time period	Sample representative for general population	Lost to follow-up after Baseline provided	Losses to follow-up likely to introduce bias	Exposure assignment	Complete or partial residential address history provided [a]	Description of size ranges	QA/QC for UFP measures described?	Exposure assessment implemented consistently across all	Outcome measures clearly defined and implemented	Outcome assessors blinded to exposure status resp. Case-control status of participants	Analysis adjusted for other air pollutants
Ostro et al. (2015)	Yes	RS	40	Yes	Yes	SG	Yes	CD	Micro-environments	Yes	Yes	Yes	Yes	Yes	Yes	Yes

CD: Cannot determine, RS: Random sample, SG: selected group.

Table A6b: Objective quality indicators, long-term studies, morbidiy

Reference	Study population specified	Sample type of study population	Response Rate [%]	Subjects recruited from same or similar populations	Subjects recruited from same time period	Sample representative for general population	Lost to follow-up after Baseline provided	Losses to follow-up likely to introduce bias	Exposure assignment	Complete or partial residential address history provided a	Description of size ranges	QA/QC for UFP measures described?	Exposure assessment implemented consistently across all	Outcome measures clearly defined and implemented	Outcome assessors blinded to exposure status resp. Case-control status of participants	Analysis adjusted for other air pollutants
Laurent et al. (2014)	Yes	Other	NA	Yes	Yes	SG	NA	NA	Geocoded addresses	NR/ NR	Yes	Yes	Yes	Yes	Yes	No
Laurent et al. (2016a)	Yes	Other	NA	Yes	Yes	SG	NA	NA	Geocoded addresses	NR/ NR	Partly	Yes	Yes	Yes	Yes	No
Laurent et al. (2016b)	Yes	Other	NA	Yes	Yes	SG	NA	NA	Geocoded addresses	NR/ NR	Partly	No	Yes	Yes	Yes	No
Li et al. (2017)	Yes	Random + CS	NR/ NR	No	Yes	SR	NA	NA	Microenvironments	Yes	Yes	No	Yes	Yes	Yes	No

CS: Convenience Sample, NA: Not applicable, NR/NR: Not reported/ no reference given, SG: selected group, SR: Somewhat representative.

Table A6d: Objective quality indicators, long-term studies, subclinical outcomes

Reference	Study population specified	Sample type of study population	Response Rate [%]	Subjects recruited from same or similar populations	Subjects recruited from same time period	Sample representative for general population	Lost to follow-up after Baseline provided	Losses to follow-up likely to introduce bias	Exposure assignment	Complete or partial residential address historyprovided a	Description of size ranges	QA/QC for UFP measures described?	Exposure assessment implemented consistently across all	Outcome measures clearly defined and implemented	Outcome assessors blinded to exposure status resp. Case-control status of participants	Analysis adjusted for other air pollutants
Aguilera et al. (2016)	Yes	RS	NR/ RG	Yes	Yes	SR	NR/ RG	Yes	Microen-viron-ments	Yes	Yes	No	Yes	Yes	Yes	Yes
Lane et al (2015)	Yes	RS + CS	NR/ NR	No	Yes	SR	NA	NA	Microen-viron-ments	NR/ NR	No	Yes	Yes	Yes	Yes	No
Lane et al. (2016)	Yes	RS + CS	NR/ RG	No	No	SR	NA	NA	Microen-viron-ments	NR/ NR	Yes	Yes	Yes	NR/ RG	Yes	No

Sunyer et al. (2015)	Yes	RS	59	Yes	Yes	SR	Yes	No	NA	No	Yes	No	Yes	Yes	No	No
Viehmann et al. (2015)	Yes	RS	NR/ NR	Yes	Yes	SR	NR/ RG	Can-not deter-ter-mine	Geocoded addresses	Yes	Yes	No	Yes	Yes	Yes	No

CS: Convenience Sample, NA: Not applicable, NR/NR: Not reported/ no reference given, NR/RG: Not reported/ reference given, RS: Random sample, SG: selected group, SR: Somewhat representative.