Environment International 142 (2020) 105876

Contents lists available at ScienceDirect

Environment International

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

Review article

Short-term exposure to particulate matter (PM_{10} and $PM_{2.5}$), nitrogen dioxide (NO_2), and ozone (O_3) and all-cause and cause-specific mortality: Systematic review and *meta*-analysis

Pablo Orellano[a,*], Julieta Reynoso[b], Nancy Quaranta[c,d], Ariel Bardach[e], Agustin Ciapponi[e]

[a] Centro de Investigaciones y Transferencia San Nicolás, Universidad Tecnológica Nacional (CONICET), San Nicolás, Argentina
[b] Hospital General "San Felipe", San Nicolás, Argentina
[c] Facultad Regional San Nicolás, Universidad Tecnológica Nacional, San Nicolás, Argentina
[d] Comisión de Investigaciones Científicas de la Provincia de Buenos Aires, La Plata, Argentina
[e] Instituto de Efectividad Clínica y Sanitaria (IECS-CONICET), Buenos Aires, Argentina

ARTICLE INFO

Keywords:
Air pollutants
Mortality
Time series studies
Observational study
Systematic review
Meta-analysis

ABSTRACT

Background: Air pollution is a leading cause of mortality and morbidity worldwide. Short-term exposure (from one hour to days) to selected air pollutants has been associated with human mortality. This systematic review was conducted to analyse the evidence on the effects of short-term exposure to particulate matter with aerodynamic diameters less or equal than 10 and 2.5 μm ($PM_{10,}$ $PM_{2.5}$), nitrogen dioxide (NO_2), and ozone (O_3), on all-cause mortality, and PM_{10} and $PM_{2.5}$ on cardiovascular, respiratory, and cerebrovascular mortality.

Methods: We included studies on human populations exposed to outdoor air pollution from any source, excluding occupational exposures. Relative risks (RRs) per 10 $\mu g/m^3$ increase in air pollutants concentrations were used as the effect estimates. Heterogeneity between studies was assessed using 80% prediction intervals. Risk of bias (RoB) in individual studies was analysed using a new domain-based assessment tool, developed by a working group convened by the World Health Organization and designed specifically to evaluate RoB within eligible air pollution studies included in systematic reviews. We conducted subgroup and sensitivity analyses by age, sex, continent, study design, single or multicity studies, time lag, and RoB. The certainty of evidence was assessed for each exposure-outcome combination. The protocol for this review was registered with PROSPERO (CRD42018087749).

Results: We included 196 articles in quantitative analysis. All combinations of pollutants and all-cause and cause-specific mortality were positively associated in the main analysis, and in a wide range of sensitivity analyses. The only exception was NO_2, but when considering a 1-hour maximum exposure. We found positive associations between pollutants and all-cause mortality for PM_{10} (RR: 1.0041; 95% CI: 1.0034–1.0049), $PM_{2.5}$ (RR: 1.0065; 95% CI: 1.0044–1.0086), NO_2 (24-hour average) (RR: 1.0072; 95% CI: 1.0059–1.0085), and O_3 (RR: 1.0043; 95% CI: 1.0034–1.0052). PM_{10} and $PM_{2.5}$ were also positively associated with cardiovascular, respiratory, and cerebrovascular mortality. We found some degree of heterogeneity between studies in three exposure-outcome combinations, and this heterogeneity could not be explained after subgroup analysis. RoB was low or moderate in the majority of articles. The certainty of evidence was judged as high in 10 out of 11 combinations, and moderate in one combination.

Conclusions: This study found evidence of a positive association between short-term exposure to PM_{10}, $PM_{2.5}$, NO_2, and O_3 and all-cause mortality, and between PM_{10} and $PM_{2.5}$ and cardiovascular, respiratory and cerebrovascular mortality. These results were robust through several sensitivity analyses. In general, the level of evidence was high, meaning that we can be confident in the associations found in this study.

1. Introduction

A high proportion of ambient air pollution is generated from combustion processes (Goldberg et al., 2003). Particularly for the most studied and widespread air pollutants, i.e. particles with aerodynamic diameters under 10 and 2.5 μm (PM_{10} and $PM_{2.5}$), ozone (O_3), sulphur

* Corresponding author at: Colon 332, San Nicolas, Provincia de Buenos Aires, ZIP 2900, Argentina.
E-mail address: porellano@frsn.utn.edu.ar (P. Orellano).

https://doi.org/10.1016/j.envint.2020.105876
Received 19 December 2019; Received in revised form 1 June 2020; Accepted 3 June 2020
Available online 23 June 2020

dioxide (SO_2), nitrogen dioxide (NO_2), and carbon monoxide (CO), air quality standards including safe thresholds have been established (Suh et al., 2000).

Air pollution has a widely recognized impact on human health, and there is a broad body of evidence that relates the exposure to air pollution and a wide range of adverse effects (WHO Regional Office for Europe, 2013). Short- and long-term exposure to ambient air pollution contributes to disease burden through an increase in mortality risk, years of life lost, and years lived with disability (Ostro et al., 2018). According to the Global Burden of Disease (GBD) study estimates, around 4.2 million total deaths were directly attributable to particulate matter smaller than 2.5 μm ($PM_{2.5}$) ambient air pollution in 2015 (Cohen et al., 2017). Evidence of the effects of the short-term exposure (from one hour to days) to air pollutants are conclusive for all-cause, respiratory, and cardiovascular mortality, and also for hospital admissions or emergency department visits (WHO|Air quality guidelines - global update 2005," 2018). A number of recent systematic reviews and *meta*-analyses have demonstrated the effect of PM (Atkinson et al., 2015, 2014; Bell et al., 2013; Shah et al., 2015), NO_2, NO_x (Newell et al., 2018; Shah et al., 2013), and O_3 (Bell et al., 2014) on all-cause and cause-specific mortality. In general, time-series studies have been useful to provide substantial information to derive ambient air quality guidelines and standards for acute exposure (Goldberg et al., 2003). In fact, this evidence is in consonance with the air quality guidelines (AQGs) published by the World Health Organization (WHO), a primary reference for air pollution standards worldwide (Landrigan et al., 2018). These are useful documents accessible to health professionals and decision makers, which provide recommendations for key air pollutants, based on global synthesis of scientific evidence. However, numerous recent studies have been published that analyse associations between pollutants and outcomes in the short-term, and the exact shape of the concentration–response functions (CRFs) is unknown or insufficiently defined for many pollutants and outcomes (Landrigan et al., 2018).

As new scientific evidence is generated, AQGs need to be periodically revised and, where necessary, updated. As a result of the vast amount of evidence published in recent years, especially on the lower and upper bounds of the air pollution exposure distribution, WHO has convened a Guideline Development Group to revise the last version of the guidelines published in 2006.

Our systematic review was commissioned by the WHO in order to generate evidence to support the new update of the AQGs, with the aim of providing updated evidence-based numerical concentration levels (i.e. guidelines) and, where possible, an indication of the shape of the CRFs for a number of ambient air pollutants, for relevant averaging times (i.e. long- and short-term exposure duration) and in relation to critical health outcomes. In particular, the objective of this systematic review and *meta*-analysis was to synthetize the worldwide evidence on the effects of short-term exposure to PM ($PM_{2.5}$ and PM_{10}), NO_2, and O_3, on all-cause and/or cause-specific mortality, including cardiovascular, respiratory, and cerebrovascular mortality. We aimed to fill the gap in the knowledge in this field, as previous studies were focused on specific air pollutants or outcomes, publication dates varied, or searches were more restricted. Based on previous evidence, the exposure-outcome combinations that were analysed included PM, NO_2, and O_3 – all-cause mortality, and PM – cardiovascular, respiratory and cerebrovascular mortality. SO_2 was not included in this study, as this pollutant was specifically addressed in a different systematic review. This systematic review was guided by the requirements outlined in the WHO Handbook for Guideline Development, 2nd edition (WHO Handbook for Guideline Development – 2nd Edition, 2014).

2. Methods

2.1. Protocol, registration, and reporting standards

The protocol for this study was registered with PROSPERO (http://www.crd.york.ac.uk/PROSPERO/) under registration number CRD42018087749, before the formal screening of search results (Supplementary File S.1). The reporting complied with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) standards (Moher et al., 2009), with slight adaptations, since these were originally intended for health care intervention evaluation. The PRISMA checklist for this study can be seen in Table A.1 (Appendix).

2.2. Research question

A summary of the Population, Exposure, Comparator, and Outcomes (PECO) question (Morgan et al., 2018) is presented below:

P: Among human population, what is the effect of

E: Short-term exposure to ambient air pollutants (PM_{10}, $PM_{2.5}$, NO_2, O_3) versus

C: Exposure to lower levels of air pollution (difference of 10 μg/m3) on

O: All-cause, cardiovascular, respiratory, and cerebrovascular mortality.

2.3. Eligibility criteria

Our study population consisted of general human population, including subgroups at risk, i.e. children, pregnant women, the elderly, and patients with particular conditions, of all ages, developed and developing areas, both urban and rural settings. We have not imposed geographical restrictions.

We considered only short-term exposures, defined in the order of one hour to 7 days (Shah et al., 2013), to ambient air PM_{10}, $PM_{2.5}$, NO_2, and O_3, from any source, expressed in a concentration unit (e.g. $\mu g/m^3$, ppb). Articles evaluating NO_x were also included in the calculations, using a factor of 0.44 for the conversion from NO_x to NO_2 (Anderson et al., 2013). The exposure to the pollutant of interest had to be via inhalation through outdoor ambient air predominantly. For PM_{10} and $PM_{2.5}$, daily (24-hour) averages were considered, while 8-hour or 24-hour maximum concentrations were included together for O_3. In the case of NO_2, two separate analyses were performed, one for 24-hour average, and one for daily 1-hour maximum. The comparison was with the same population exposed to a lower level of air pollutants, considering a given difference in a standardized concentration (10 $\mu g/m^3$).

The outcomes were classified using the International Statistical Classification of Diseases and Related Health Problems (ICD), and encompassed all-cause natural mortality (ICD-10: A00 to R99), cause-specific mortality including cardio (ICD-10: I01 to I59) and cerebrovascular (ICD-10: I60 to I69), and respiratory mortality (ICD-10: J00 to J99) (WHO|International Classification of Diseases (ICD) Information Sheet," n.d.). We defined an exposure-outcome combination as a pair comprising one of the pollutants selected, and one of the outcomes. Based on previous evidence, the exposure-outcome combinations that were analysed included PM, NO_2, and O_3 – all-cause mortality, and PM – cardiovascular, respiratory, and cerebrovascular mortality.

As for study designs, we included human epidemiological studies, i.e. ecological time-series (ETS), case-crossover (CCO), cohort, and panel studies; systematic reviews of the above studies were used to scan for references. A glossary of definitions used in this study can be seen in

Table A.2 (Appendix). Studies were excluded if they evaluated the exposure in occupational settings, or as a result of indoor exposure exclusively; qualitative studies, reviews, methodological papers, and non-human studies (i.e. in vivo, in vitro), were also excluded. These exclusions were made at the title-abstract screening stage. Other reasons for exclusions were in the case of studies with partial or complete geographical and temporal overlap verified during *meta*-analysis, in order to avoid double-counting participants (Shah et al., 2015). In case of total or partial overlapping data, the article for inclusion was selected according to the following criteria: 1) wider geographical distribution; 2) longer duration of the study period; 3) more recent publishing date. When multiple reports of the same study were detected, all articles were included in a first stage, and previously mentioned criteria were used to prevent double-counting. After the selection of the more informative article, the rest of the articles were excluded from the analysis. Regarding lag times, when multiple lag-estimates were reported in papers, the framework proposed by Atkinson et al. (Atkinson et al., 2014) was followed: if only one lag estimate for a given pollutant/outcome pair is reported, it was included in the analysis. If multiple lag-estimates were reported, the selection algorithm was: 1) the most frequently used lag in all selected studies (0 and 1 days in this systematic review); 2) single lags, but not cumulative/distributed lags.

2.4. Studies search and selection

Studies were searched comprehensively in the following bibliographic databases and citation indexes: Medical Literature Analysis and Retrieval System Online (MEDLINE) via PubMed, and Scopus via Elsevier. Moreover, regional databases in English and other languages as Literatura Latinoamericana y del Caribe en Ciencias de la Salud (LILACS), Western Pacific Region Index Medicus (WPRIM), Index Medicus for South-East Asia Region (IMSEAR), Index Medicus for the Eastern Mediterranean Region (IMEMR), and African Index Medicus (AIM) were searched to retrieve additional articles and grey literature. WHO Regional Databases are open access resources that comprises journal indexes of locally produced information, especially some developing country health journals and other reports, and complement the internationally known bibliographic databases. References of identified relevant articles were scanned to identify additional reports matching the research question.

Data search included studies from January 1990 up to November 2017. We performed an update of the data search in September 2018, in order to incorporate relevant studies that might have been published shortly before the finalization of the review.

We developed a literature search strategy for each database, using free text and medical subject headings (MeSH) terms, and considering all the eligibility criteria. An example of the search strategy applied to PubMed can be seen in Table A.3 (Appendix). In addition, a manual search in reference lists from other systematic reviews was performed to find additional relevant studies. The strategy was developed by PO, with input from the systematic review team, and reviewed by JR and NQ.

PO and JR independently screened titles and abstracts, and potentially eligible studies were assessed again by the same reviewers based on the full-text, to ensure that those met all the eligibility criteria. Any disagreement on inclusion was resolved by discussion and, if no consensus was reached, a third reviewer (NQ) was consulted. Reasons for excluding articles at this stage were recorded, and registered in an electronic extraction form developed in Microsoft Excel®. When the same study population was used in several publications, only the largest and the most complete study (i.e. multicity studies, or studies with wider temporal or geographical coverage) was included. In all cases, multicity studies were preferred over single-city studies. When full-texts were not available, the original authors were contacted via email, using the author information section. If no response from the author was received, the article was excluded from the analysis.

2.5. Data extraction and process

Data from selected studies was extracted by one reviewer and checked by a second. The data was then transferred to electronic extraction forms developed in Microsoft Excel®. The bibliography was managed using the software Zotero (Ahmed and Al Dhubaib, 2011). Extracted data included study details, exposures, outcomes, and data analysis. Association measures were relative risks (RRs), odds ratios (ORs), and percentage excess (increment) or change in mortality (PercIncr). The data on pollutants concentrations included the mean or median, standard deviation (SD), range, interquartile range (IQR), and percentiles. Due to the relevance of the 5th percentile as a proxy for the lowest level of exposure in a given study, these values were registered, when available. If a given study did not report the 5th percentile, but the SD or the 10th percentile were available, we estimated the 5th percentile using a normal approximation. If only the range was available, we estimated the SD as the range divided by four (Hozo et al., 2005).

All included articles contributed with at least one effect size (i.e. RR) to each exposure-outcome combinations for the main or the subgroup analyses. In some cases, one article contributed with two or more effect sizes to the same combination; for example, one study might have reported independent results from different cities, or from rural and urban settings. In these situations, the number of effect sizes could have been higher than the number of articles, for a given exposure-outcome combination. Moreover, one article might have contributed with evidence for two or more exposure-outcome combinations. These definitions can be seen in Table A.2 (Appendix).

2.6. Risk of bias

The risk of bias (RoB) assessment of included studies was conducted using a new domain-based RoB assessment tool, developed by a group of experts convened by the WHO. A detailed description of this tool can be seen in the WHO website (Risk of bias assessment instrument for systematic reviews informing WHO global air quality guidelines, 2020). The instrument allows rating sections according to 13 items grouped in six domains: confounding, selection bias, exposure assessment, outcome measurement, missing data, and selective reporting (Morgan et al., 2019). Each item can be judged as having low, moderate, or high RoB. For the item of potential confounders that were accounted for in the analysis, four critical confounders (temperature, seasonality, day-of-the-week, long-term trends) and two additional confounders (holidays, influenza epidemics) were considered. If all these confounders were included, the item was classified as having low RoB; if one or two of the additional confounders were not included, but the four critical confounders were incorporated, the item was classified as having moderate RoB; otherwise, a high RoB rating was assigned. Touloumi and colleagues have demonstrated that for cardiovascular diseases, the estimates of models that fail in the inclusion of influenza epidemics as a confounder are consistent with models that include this factor, and thus we considered this potential confounder as not critical, i.e. the absence of that confounder led to moderate RoB (Touloumi et al., 2005). The results for each domain were analysed separately, without considering a single result for the whole article/dataset. If only one item of the same domain was judged as having high RoB, the entire domain was classified as having high RoB. The same logic was applied to moderate vs. low RoB. In the same article, judgments regarding the RoB were assessed separately for each exposure-outcome combination. The assessment of RoB across studies was performed as one of the sensitivity analyses, as will be detailed later.

2.7. Data synthesis and analysis

We used RRs as the common association measure in pooled analyses. Meta-analyses input data were RRs for a standardized increment

in pollutant concentration (10 μg/m^3), assuming a linear exposure-outcome relationship (Shah et al., 2013), and taking into account the original increment of the pollutant, as for the following equation:

$$RR_{(standardized)} = e^{\mathrm{Ln}(RR(original))x10/Increment(original)}$$

The effects expressed as interquartile (or quintile, or percentile differences) were converted into effects per concentration unit increase with the previous equation. When ORs were reported in a study, they were supposed to approach the RRs, under the "rare disease assumption" (Greenland and Thomas, 1982; Knol et al., 2008; Pace and Multani, 2018), given the fact that a cumulative incidence of the outcome lower than 10% was demonstrated or assumed in all articles for all-cause and cause-specific mortality. Effects expressed as Perc-Incr were also recalculated to reflect a RR for a concentration unit increase in the pollutant, assuming a linear relationship, according to the following equation:

$$RR = \frac{Perc - Incr}{100} + 1$$

For the summary measure (pooled RRs), a random-effects (RE) model was employed, assuming that the included studies were a random selection of all possible results. The DerSimonian-Laird estimator was used for the pooled RRs (DerSimonian and Laird, 1986), a straightforward method that allows the incorporation of heterogeneity in the analysis. When the pooled effect size was calculated from 20 or less effect sizes, the Hartung and Knapp adjustment was employed (Hartung and Knapp, 2001).

Heterogeneity between studies was assessed using 80% prediction intervals (PI), this parameter used to estimate the 80% interval in which the true RR in a new air pollution study will lie (Chiolero et al., 2012). We have chosen not to measure heterogeneity using the I^2 parameter, because this statistic is a relative measure, and it is difficult to make a judgement about the absolute amount of heterogeneity. Further, the main problem with the I^2 parameter is that it can be artificially inflated when increasing the sample size (number of included studies), or when increasing the precision of the estimates from primary studies (Rücker et al., 2008). On the contrary, the use of the PI has been strongly advocated in the literature (Borenstein, 2019; Borenstein et al., 2011), as it provides an estimate of the distribution of the true effect sizes. This parameter shows whether the effect is consistent, or if it varies substantially; it also shows if the effect is harmful in all populations, or if there is no effect in some populations. The rule was that when the PI included the null effect (IntHout et al., 2016), some degree of heterogeneity was suspected. In this situation, an attempt was made to explain the source of heterogeneity by subgroup analyses, using readily available study information (age group, sex, and geographical area). Subgroups were statistically compared using the χ^2 test. Chosen subgroups were assumed to explain the heterogeneity if statistical differences were found between subgroup effect sizes. Severe heterogeneity was assumed when PI included the null effect, and at the same time this 80% PI was larger than the 95% CI by a factor of two.

Sensitivity analysis was carried out in order to assess the extent to which model assumptions could have influenced the association measures. This assessment included six analyses, i.e. the inclusion of effect sizes that considered only lags of 0, 1 and 0–1 days, an analysis by epidemiological study design (e.g. ETS, CCO), an analysis excluding papers with declared conflicts of interest (CoI), an analysis comparing articles showing high RoB in some of the 6 domains versus articles with low or moderate RoB, and an analysis comparing multicity versus single-city studies. This last analysis was not previously reported in the protocol. The sixth sensitivity analysis was the evaluation of potential unmeasured confounders through the calculation of the E-value (VanderWeele and Ding, 2017). This type of sensitivity analysis considers a potential unmeasured confounder, which is associated with the exposure and with the disease. The E-value is the minimum strength of association, measured as RR, that an unmeasured confounder would need to have with both the exposure and the outcome, conditional on the measured covariates, to fully explain away a specific exposure–outcome association. It takes into account the association of the unmeasured confounder with the exposure (RR_{EU}) and with the disease (RR_{UD}). We have chosen temperature as an example of unmeasured confounder, and focused on the association between temperature and mortality (RR_U). We selected RR_U values of temperature for each specific mortality cause from a systematic review on temperature and mortality as reported in the review by Song and colleagues (Song et al., 2017). In order to be conservatives, the higher values of different central estimates reported in that paper were selected. Then a comparison was made between the RR_U and the RRs calculated by means of *meta*-analysis in this study ($RR_{observed}$), but considering a wider range of pollutants increase (50 μg/m^3). The E-value was then calculated using the following equation:

$$E - value = RR_{observed} + \sqrt{RR_{observed} \times (RR_{observed} - 1)}$$

For each exposure-outcome combination, the rule was that when the RR_U was higher than the lower confidence limit of the E-value, comparatively weaker confounder associations could explain away the observed association, i.e. the presence of unmeasured confounders is plausible.

The potential for publication bias was assessed through two methods. First, the visual examination of funnel plots asymmetry (Sterne et al., 2011). Second, a numerical evaluation of the potential for publication bias was performed by means of the Egger's regression test (Egger et al., 1997).

The shape of the CRFs was analysed for each exposure-outcome combination, to assess the suitability of linear assumptions regarding the RRs calculations, and the possibility of thresholds occurrence. The CRF can be displayed as a graph that shows the relationship between levels of adverse health responses in exposed populations (vertical axis) and levels of ambient concentrations of a pollutant (horizontal axis), and is widely used to predict the public health impacts of proposed reductions in air pollutants (Cox, 2017).

The interaction between pollutants was analysed by means of the inclusion of co-pollutant models, considering only studies that addressed the effect of the main pollutant controlled by the inclusion of one or more co-pollutants in the regression model.

All analyses and graphics were performed using the "meta" package (version 4.9–2) (Schwarzer et al., 2015) in the statistical software R, version 3.4.4 (https://www.r-project.org/) (Albert and Rizzo, 2012). The script used for the analysis can be seen in Supplementary File S.2. Only one example for each exposure-outcome combination analysis is presented in the script, as the same structure can be replicated with slight modifications to cover the rest of the combinations.

2.8. Certainty of evidence across studies

The certainty of evidence (CoE) for each exposure-outcome combination was judged using an adaptation of the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach (Balshem et al., 2011; Guyatt et al., 2011; Morgan et al., 2016), developed by a group of experts convened by the WHO. The approach is briefly described here, and its full version can be seen in Supplementary File S.3. The CoE was assessed across five domains: limitations in studies, indirectness, inconsistency, imprecision, and publication bias. In short, the procedure was as follows: an evidence related to an exposure-outcome combination based on a number of articles was initially judged as being of moderate CoE, and then was potentially downgrading according to these five domains. The approach implies that there is always a risk of unmeasured confounding in observational studies. Therefore, it starts at moderate certainty. After this first analysis, three other domains were evaluated, allowing the possibility of upgrading the CoE. These additional domains were the large magnitude of the effect, the occurrence of all possible confounding

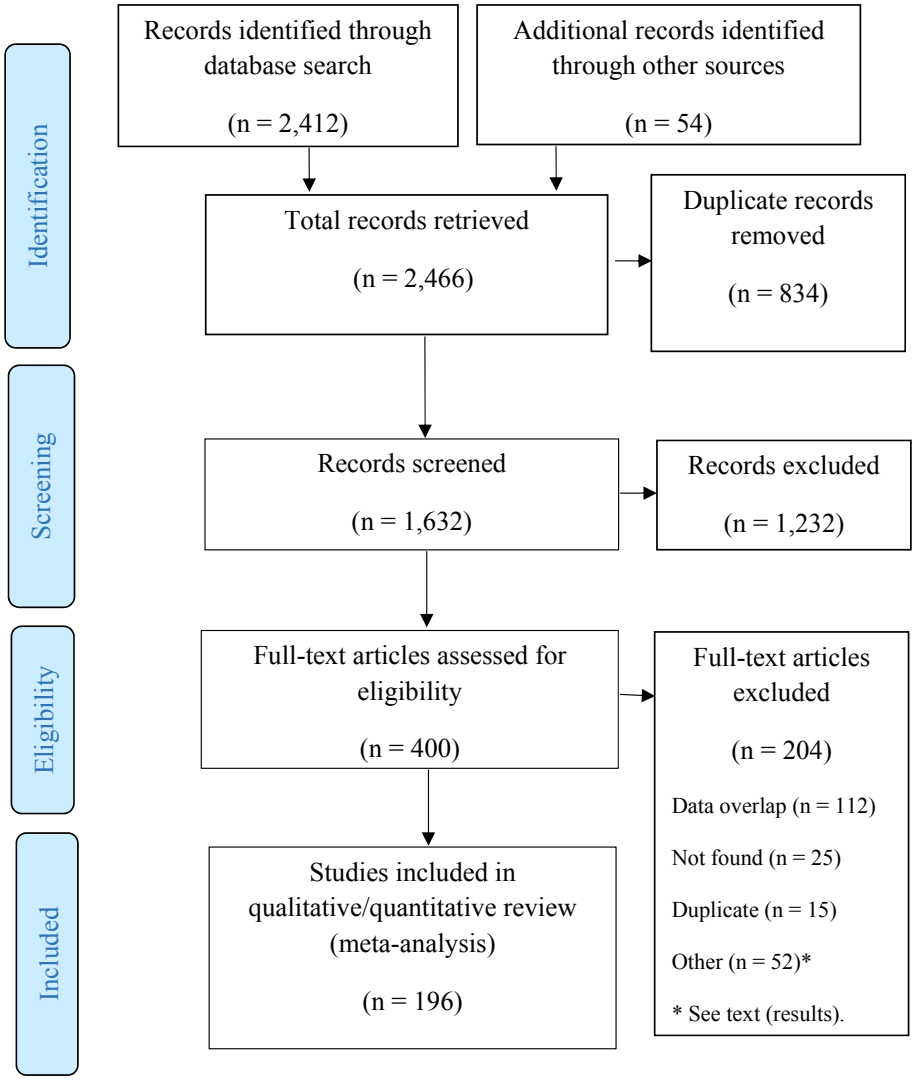


Fig. 1. Flowchart of assessment of eligible studies.

factors shifting towards the null effect, and the evidence of a concentration–response gradient (equivalent to CRFs). After applying this tool, the overall certainty was rated as: high, meaning that further research is very unlikely to change the confidence in the estimate of the effect; moderate, meaning that further research is likely to have an important impact on the confidence in the estimate of the effect; low, meaning that further research is very likely to have an important impact on the confidence in the estimate of the effect; or very low, meaning that the estimate of the effect is very uncertain. Some domains of this tool were evaluated using results of the RoB, heterogeneity, sensitivity, and publication bias analyses, which were previously described in the methodology.

Limitations in studies: the level of evidence was downgraded if there were statistical differences between studies showing high versus moderate or low RoB in the sensitivity analysis. However, we also analysed the number of studies and the impact they had in the *meta*-analysis. For example, the presence of small studies with high RoB but limited influence on the *meta*-analysis was not a reason to downgrade. If the sensitivity analysis for RoB showed a considerable influence on the pooled effect-size, the conclusions were based on the high-quality studies only, and the evidence was not downgraded. This was a judgement and there were no clear pre-set cut-off points.

Indirectness: the evidence was not downgraded based on this domain, as the research question in the included studies always reflected the original question.

Inconsistency: the evidence was downgraded if severe heterogeneity was detected, i.e. the PI included unity and was more than twice the random effects *meta*-analysis confidence interval.

Imprecision: the evidence was downgraded if the number of mortality cases used when calculating the pooled effect size was below 100,000. This number is lower than the value proposed in the adapted GRADE approach (Supplementary File S.3), because that value was computed for rate ratios in long-term studies. We have decided to use our own cut-off point for short-term studies, using data from selected multi-city studies (Chiusolo et al., 2011; Taneepanichskul et al., 2018; Ueda et al., 2009), in which significant positive effect sizes were found considering approximately 100,000 events (deaths) or less. The idea behind this reasoning is that if the number of events is sufficient for a given study to derive significant effect sizes, the same number will be adequate for *meta*-analysis.

Publication bias: the evidence was downgraded if publication bias was detected by visual inspection of the funnel plot or through the Egger's test. However, if analyses comparing multicity versus single-city studies did not show statistical differences, publication bias was dismissed, and the evidence was not downgraded.

Large effect size: unmeasured confounders and the results of the E-value were used, i.e. when the RR_U between air temperature and mortality was higher than the lower confidence limit of the E-value (the

presence of unmeasured confounders is plausible), the evidence was not upgraded. Otherwise, the evidence was upgraded.

Confounding domain: the evidence was not upgraded using this domain, as several potential confounders could shift the RR in both directions.

Concentration- response gradient domain: the evidence was upgraded when estimations of CRFs were statistically significant, and thus a positive linear or nonlinear association can be assumed.

3. Results

3.1. Studies included

Database searches retrieved 2,412 studies, and 54 additional articles that we identified through reference lists of selected systematic reviews (Achilleos et al., 2017; Atkinson et al., 2014; Bell et al., 2014, 2013; DeVries et al., 2017; Fajersztajn et al., 2017; Li et al., 2016; Lu et al., 2015a; Mills et al., 2015; Newell et al., 2017; Shang et al., 2013; Wang et al., 2014; Zhang et al., 2018). After removing duplicates, 1,632 records were screened by title and abstract, and 400 articles were selected for full-text eligibility assessment. Finally, we included 196 articles in quantitative analysis, showing at least one specific combination of exposure, outcome, age-group and sex (Fig. 1). These 196 articles represented 796 effect sizes. From the 54 additional articles found in selected systematic reviews, only 15 were effectively incorporated in this quantitative analysis, in general due to duplicate data or analysis, to overlapping with more recent articles, or to differences in the selected exposures or outcomes.

Extracted information from included articles can be seen in Supplementary File S.4. These articles comprised a period of publication of 27 years, between 1992 and 2019. The mean duration of the studies was 5.76 years (SD: 4.38, range: 1–31). The majority of studies were carried out in Asia (73), Europe (69), and America (45). Other continents were less represented, for example Oceania (5), or Africa (1), while 3 studies comprised more than one continent. Among selected articles, 62 reported results of multicenter collaborative studies, e.g. APHEA, APHENA, EpiAir, EMECAM, ESCALA, MED-PARTICLES, NMMAPS, PAPA. The most frequently used study designs were ETS studies (1 4 3), followed by CCO (47) and cohort (6) studies. However, the few studies evaluating a cohort of patients were, in all cases, analysed using common methods from ETS or CCO designs, i.e. generalized additive models (GAMs) and conditional logistic regression (CLR). There were no studies self-defined as panel studies. Almost all studies considered mortality in general populations (1 8 9); conversely, 7 studies estimated the influence of air pollution on mortality in specific populations, i.e. persons with cardiovascular diseases, congestive heart failure, hypertension, chronic obstructive pulmonary disease (COPD), or asthma. Some studies reported more than one effect size per exposure-outcome combination, for the same sex and age groups. In these cases, different effect sizes represented separate areas (e.g. rural vs. urban settings), cities, or countries, and were taken as independent samples in the *meta*-analysis.

Regarding individual characteristics of cases considered in the exposure-outcome combinations, 550 effect sizes were retrieved for all-ages or adults, 207 for elderly people, and 38 for children or newborns. Only 183 effect sizes were sex-specific, while 612 effect sizes considered both sexes. All combinations of selected pollutants and outcomes were well represented, with 7 or more effect sizes (range: 7–66 effect sizes). Conversely, in the full-text selection stage, 204 articles were excluded due to different reasons, i.e. total or partial geographical or temporal data overlap (1 1 2), full-text articles not found (25), duplicate data or analysis (15), differences in the applied methodology (e.g. multiple Pearson correlation, principal component analysis, instrumental variable models) (13), pollutants or exposures other than selected (11), lack of data on pollutant's increase (6), a different association measure or analytic model (5), lack of reporting on error terms (5), long-term exposure only (4), incomplete data (4), not a paper or scientific publication (2), and only protocol presented (2) (Supplementary File S.5). Regarding exclusions due to overlapping, in all cases multicity studies were replaced by other multicity studies with wider geographical or temporal coverage (See Supplementary File S.6 for details on replacements due to overlap between multicity studies).

The mean/median concentration of pollutants ranged from 14.0 to 245.0 $\mu g/m^3$ (PM_{10}), from 5.7 to 176.7 $\mu g/m^3$ ($PM_{2.5}$), from 18.4 to 99.2 $\mu g/m^3$ (NO_2 24-hour average), from 40.0 to 161.2 $\mu g/m^3$ (NO_2 1-hour max.), and from 15.2 to 206.0 $\mu g/m^3$ (O_3). Fifth (5th) percentiles reported in the articles or estimated in this study ranged from 0.0 to 221.0 $\mu g/m^3$ (PM_{10}), from 0.0 to 26.15 $\mu g/m^3$ ($PM_{2.5}$), from 0.0 to 63.8 $\mu g/m^3$ (NO_2 24-hour average), from 0.0 to 65.8 $\mu g/m^3$ (NO_2 1-hour max.), and from 0.0 to 93.1 $\mu g/m^3$ (O_3). All these data regarding included studies can be seen in Supplementary File S.4.

3.2. Risk of bias

In three out of 6 domains, the RoB was found to be only low or moderate. These domains were selection bias, exposure assessment, and selective responding. For example, in the selective responding domain, 21% (77) of the exposure-outcome combinations (31 articles) were classified as moderate RoB because they reported a subset or re-analysis of data that was already published in a previous article, or because reported results are preliminary analyses (in both cases, regardless of being non-overlapping data). In the other three domains, a variable proportion of articles were found to have high RoB. The domain with a higher proportion of high RoB was missing data (59% of exposure-outcome combinations presenting high RoB), while confounding and outcome measurement showed low proportions of high RoB (7% and 2% of exposure-outcome combinations with high RoB, respectively). The reasons for the high RoB in the missing data domain were related to the lack of information on the number of missing values in the exposure, or to the absence of information regarding imputation methods. The same judgment was applied when the number of missing data was higher than 5%. A summary of the results of the RoB analysis can be seen in Fig. 2. The description of the RoB analysis per item and domain in individual studies, together with the rationale to justify each judgment, are presented in Supplementary File S.7.

3.3. Meta-analysis

The detailed results regarding pooled effect sizes (RRs) for a 10 $\mu g/m^3$ increase in pollutants, p-values, heterogeneity, and funnel plot asymmetry can be seen in Table 1. We found positive associations between pollutants and all-cause mortality for PM_{10} (RR: 1.0041; 95% CI: 1.0034–1.0049), $PM_{2.5}$ (RR: 1.0065; 95% CI: 1.0044–1.0086), NO_2 (24-hour average) (RR: 1.0072; 95% CI: 1.0059–1.0085), and O_3 (RR: 1.0043; 95% CI: 1.0034–1.0052). On the contrary, the association was non-significant for NO_2 (1-hour max.) (RR: 1.0024; 95% CI: 0.9995–1.0053). PM_{10} and $PM_{2.5}$ were also positively associated with cardiovascular, respiratory, and cerebrovascular mortality (Table 1). The forest plots for these analyses are shown in Figs. A.1 to A.11 (Appendix).

In the majority of the exposure-outcome combinations, the 80% PIs excluded the null effect; this means that the heterogeneity between studies was not substantial. All these values can be seen in Table 1. Conversely, in the associations between $PM_{2.5}$ and respiratory or cerebrovascular mortality, and between NO_2 (1-hour max.) and all-cause mortality, the 80% PIs included the null effect. This suggests heterogeneity to some extent, sufficient to find some populations in which the effects of these pollutants on the outcome could be null. The three exposure-outcome combinations that displayed some degree of heterogeneity between studies in the main analysis did not show statistical differences between subgroups in the subgroup analysis by age, sex, and continent. In this sense, heterogeneity could not be explained by means

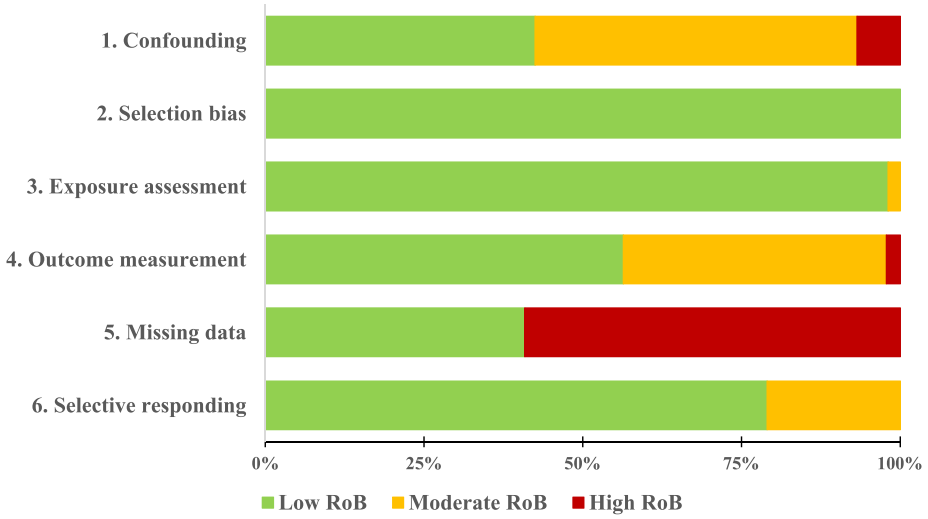


Fig. 2. Summary of the RoB assessment.

of differences in subpopulations, at least in the subgroups that were considered in this review. The results of these subgroup analyses can be seen in Table A.4 (Appendix).

The visual inspection of funnel plots gave, in the majority of cases, subtle to pronounced indication of asymmetry for PM_{10}, with the exception of the combination PM_{10} - respiratory mortality. In the case of $PM_{2.5}$, the asymmetry in the funnel plot was only evident for all-cause mortality. We also found evidence of asymmetric funnel plots for O_3 and NO_2 (24-hour average). All these results were confirmed by the Egger's test (Table 1), when this test could be performed (exposure-outcome combinations with 10 or more effect sizes). Funnel plots for exposure-outcome combinations with 10 or more effect sizes can be seen in Fig. 3.

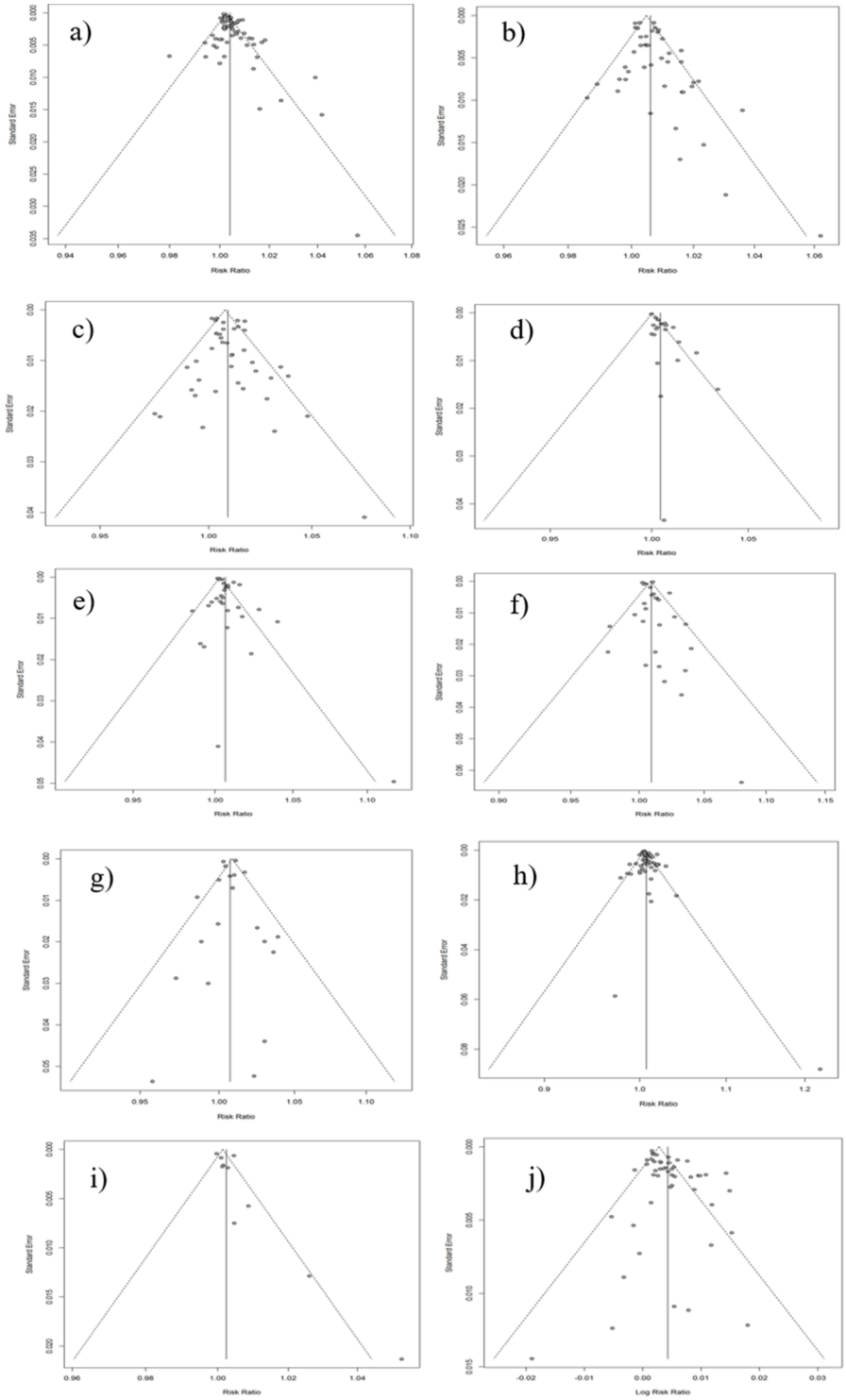
The sensitivity analysis by lag showed positive associations in the same exposure-outcome combinations as in the main analysis, with two exceptions, i.e. $PM_{2.5}$ and cerebrovascular mortality, and NO_2 (1-hour max.) and all-cause mortality (Table A.5 of the Appendix). The same occurred with the analysis by study design for ETS (Table A.6 of the Appendix), but not for CCO (Table A.7 of the Appendix), probably due to the small number of studies. This design also showed in general higher values of heterogeneity. When considering only multicity studies, the associations were positive in the same exposure-outcome combinations as in the main analysis, with the exception of two (PM_{10} and $PM_{2.5}$ with cerebrovascular mortality), probably due to the small number of effect sizes (Table A.8 of the Appendix). NO_2 (1-hour max.) was not analysed because the number of studies was too small. Differences between multicity and single-city studies were observed in only one exposure-outcome combination (PM_{10} – respiratory mortality). In the analysis using only studies with low or moderate RoB, the results were similar to the main analysis (Table A.9 of the Appendix). Still, statistical differences between studies with low/moderate versus high RoB were found in the combinations $PM_{2.5}$ – all-cause mortality, $PM_{2.5}$ – cerebrovascular mortality, and NO_2 (24-hour average) – all-cause mortality. The analysis by declared CoI was not performed, as only one of the included papers declared potential conflicts.

E-values with 95% CIs and RR_Us can be seen in (Table A.10 of the

Table 1
Exposures, outcomes and pooled effect sizes.

Pollutant	Outcome	Number of effect sizes	RR (95% CI)	p-value	PI	Egger's test (p-value)
PM_{10}	All-cause mortality	66	1.0041 (1.0034–1.0049)	< 0.0001	1.0013–1.0070	< 0.001
PM_{10}	Cardiovascular mortality	44	1.0060 (1.0044–1.0077)	< 0.0001	1.0016–1.0105	0.024
PM_{10}	Respiratory mortality	41	1.0091 (1.0063–1.0119)	< 0.0001	1.0017–1.0166	0.209
PM_{10}	Cerebrovascular mortality	20	1.0044 (1.0022–1.0066)	0.0005	1.0001–1.0087	< 0.001
$PM_{2.5}$	All-cause mortality	29	1.0065 (1.0044–1.0086)	< 0.0001	1.0017–1.0114	0.015
$PM_{2.5}$	Cardiovascular mortality	28	1.0092 (1.0061–1.0123)	< 0.0001	1.0026–1.0158	0.803
$PM_{2.5}$	Respiratory mortality	20	1.0073 (1.0029–1.0116)	0.0023	0.9998–1.0148	0.606
$PM_{2.5}$	Cerebrovascular mortality	7	1.0072 (1.0012–1.0132)	0.0257	0.9953–1.0192	N/A
NO_2 (24-hour average)	All-cause mortality	54	1.0072 (1.0059–1.0085)	< 0.0001	1.0031–1.0113	0.048
NO_2 (1-hour max.)	All-cause mortality	10	1.0024 (0.9995–1.0053)	0.0892	0.9985–1.0064	0.154
O_3	All-cause mortality	48	1.0043 (1.0034–1.0052)	< 0.0001	1.0013–1.0073	0.001

RR, pooled relative risks; 95% CI, 95% confidence interval; p-value, significance of the association or statistical tests; PI, 80% prediction interval; N/A, not applicable (< 10 studies).



a) PM_{10} – all-cause mortality; b) PM_{10} – cardiovascular mortality; c) PM_{10} – respiratory mortality; d) PM_{10} – cerebrovascular mortality; e) $PM_{2.5}$ – all-cause mortality; f) $PM_{2.5}$ – cardiovascular mortality; g) $PM_{2.5}$ – respiratory mortality; h) NO_2 (24-hour average) – all-cause mortality; i) NO_2 (1-hour max.) – all-cause mortality; j) O_3 – all-cause mortality.

Fig. 3. Funnel plots to explore publication bias for each exposure-outcome combination. (a) PM_{10} – all-cause mortality; (b) PM_{10} – cardiovascular mortality; (c) PM_{10} – respiratory mortality; (d) PM_{10} – cerebrovascular mortality; (e) $PM_{2.5}$ – all-cause mortality; (f) $PM_{2.5}$ – cardiovascular mortality; (g) $PM_{2.5}$ – respiratory mortality; (h) NO_2 (24-hour average) – all-cause mortality; (i) NO_2 (1-hour max.) – all-cause mortality; (j) O_3 – all-cause mortality.

Appendix). In all exposure-outcome combinations except two, the RR_Us were below the lower limit of the E-value, meaning that unmeasured confounders are not supposed to have a major influence on the association. The exceptions were the combinations $PM_{2.5}$ – respiratory mortality and NO_2 (1-hour) – all-cause mortality, which showed the RR_U above the lower limit of the E-value, and thus the presence of unmeasured confounders that explain away the exposure-outcome associations cannot be ruled out.

Table 2
Certainty of evidence profile for each exposure-outcome combination.

Exposure - Outcome	Limitations in studies	Indirectness	Inconsistency	Imprecision	Publication bias	Large effect size	Confounding	Concentration-response gradient	Certainty of evidence
	Downgrade					Upgrade			
PM_{10} – all-cause mortality	(0) No differences between studies with low/moderate versus high RoB	(0) The research question in the studies reflects the original question.	(0) 80% prediction interval did not include unity.	(0) Number of mortality cases higher than 100,000.	(0) Publication bias detected, but no difference between multicity and single-city studies was observed.	(+1) Unmeasured confounding would not suffice to explain away the effect estimate.	(0) Several potential confounders that would shift the RR in both directions.	(+1) Significant positive association detected in the main analysis.	High ⊗⊗⊗⊗
PM_{10} – cardiovascular mortality	(0) No differences between studies with low/moderate versus high RoB	(0) The research question in the studies reflects the original question.	(0) 80% prediction interval did not include unity.	(0) Number of mortality cases higher than 100,000.	(0) Publication bias detected, but no difference between multicity and single-city studies was observed.	(+1) Unmeasured confounding would not suffice to explain away the effect estimate.	(0) Several potential confounders that would shift the RR in both directions.	(+1) Significant positive association detected in the main analysis.	High ⊗⊗⊗⊗
PM_{10} – respiratory mortality	(0) No differences between studies with low/moderate versus high RoB	(0) The research question in the studies reflects the original question.	(0) 80% prediction interval did not include unity.	(0) Number of mortality cases higher than 100,000.	(0) Publication bias was not detected.	(+1) Unmeasured confounding would not suffice to explain away the effect estimate.	(0) Several potential confounders that would shift the RR in both directions.	(+1) Significant positive association detected in the main analysis.	High ⊗⊗⊗⊗
PM_{10} –cerebrovascular mortality	(0) No differences between studies with low/moderate and high RoB	(0) The research question in the studies reflects the original question.	(0) 80% prediction interval did not include unity.	(0) Number of mortality cases higher than 100,000.	(0) Publication bias detected, but no difference between multicity and single-city studies was observed.	(+1) Unmeasured confounding would not suffice to explain away the effect estimate.	(0) Several potential confounders that would shift the RR in both directions.	(+1) Significant positive association detected in the main analysis.	High ⊗⊗⊗⊗
$PM_{2.5}$ – all-cause mortality	(0) Statistical differences between studies with low/moderate versus high RoB, but studies showing high RoB had small weight on the results	(0) The research question in the studies reflects the original question.	(0) 80% prediction interval did not include unity.	(0) Number of mortality cases higher than 100,000.	(0) Publication bias detected, but no difference between multicity and single-city studies was observed.	(+1) Unmeasured confounding would not suffice to explain away the effect estimate.	(0) Several potential confounders that would shift the RR in both directions.	(+1) Significant positive association detected in the main analysis.	High ⊗⊗⊗⊗
$PM_{2.5}$ – cardiovascular mortality	(0) No differences between studies with low/moderate versus high RoB	(0) The research question in the studies reflects the original question.	(0) 80% prediction interval did not include unity.	(0) Number of mortality cases higher than 100,000.	(0) Publication bias was not detected.	(+1) Unmeasured confounding would not suffice to explain away the effect estimate.	(0) Several potential confounders that would shift the RR in both directions.	(+1) Significant positive association detected in the main analysis.	High ⊗⊗⊗⊗
$PM_{2.5}$ – respiratory mortality	(0) No differences between studies with low/moderate versus high RoB	(0) The research question in the studies reflects the original question.	(0) 80% prediction interval included unity, but is not twice the confidence interval.	(0) Number of mortality cases higher than 100,000.	(0) Publication bias was not detected.	(0) Unmeasured confounding could influence the effect estimate.	(0) Several potential confounders that would shift the RR in both directions.	(+1) Significant positive association detected in the main analysis.	High ⊗⊗⊗⊗
$PM_{2.5}$ – cerebrovascular mortality	(-1) Statistical differences between studies with low/moderate versus high RoB	(0) The research question in the	(0) 80% prediction interval included unity, but is not twice	(0) Number of mortality cases	(0) Publication bias was not detected.	(+1) Unmeasured confounding would not suffice to explain	(0) Several potential confounders that	(+1) Significant positive association detected in the main analysis.	High ⊗⊗⊗⊗

(*continued on next page*)

Table 2 (*continued*)

Exposure - Outcome	Limitations in studies	Indirectness	Inconsistency	Imprecision	Publication bias	Large effect size	Confounding	Concentration-response gradient	Certainty of evidence
	Downgrade					Upgrade			
		studies reflects the original question.	the confidence interval.	higher than 100,000.		away the effect estimate.	would shift the RR in both directions.		
NO_2 (24-hour average) – all-cause mortality	(-1) Statistical differences between studies with low/ moderate versus high RoB	(0) The research question in the studies reflects the original question.	(0) 80% prediction interval did not include unity.	(0) Number of mortality cases higher than 100,000.	(0) Publication bias detected, but no difference between multicity and single-city studies was observed.	(+1) Unmeasured confounding would not suffice to explain away the effect estimate.	(0) Several potential confounders that would shift the RR in both directions.	(+1) Significant positive association detected in the main analysis.	High ⊗⊗⊗⊗
NO_2 (1-hour max.) – all-cause mortality	(0) No differences between studies with low/moderate versus high RoB	(0) The research question in the studies reflects the original question.	(0) 80% prediction interval did not include unity.	(0) Number of mortality cases higher than 100,000.	(0) Publication bias was not detected.	(0) Unmeasured confounding could influence the effect estimate.	(0) Several potential confounders that would shift the RR in both directions.	(0) No significant association detected in the main analysis.	Moderate ⊗⊗○○
O_3 – all-cause mortality	(0) No differences between studies with low/moderate versus high RoB	(0) The research question in the studies reflects the original question.	(0) 80% prediction interval did not include unity.	(0) Number of mortality cases higher than 100,000.	(0) Publication bias detected, but no difference between multicity and single-city studies was observed.	(+1) Unmeasured confounding would not suffice to explain away the effect estimate.	(0) Several potential confounders that would shift the RR in both directions.	(+1) Significant positive association detected in the main analysis.	High ⊗⊗⊗⊗

Certainty of evidence, starting from moderate certainty (⊗⊗⊗○); CRFs, concentration–response functions; (), between brackets is the downgrading of levels in that domain; RoB, risk of bias in individual studies; RR, relative risk.

3.4. Concentration-response functions

In all papers, the description regarding concentrations only included full-period averages, ranges and dispersion of the pollutants, but not daily values. Accordingly, only a description regarding the shape of the concentration–response curves and potential thresholds in a daily basis was carried out, in those papers in which this analysis was performed. The linearity assumptions were investigated in 40 of the included articles, in general by means of semi-parametric regressions (GAMs). The existence of non-linear behaviour and potential thresholds were mainly analysed through visual inspection of the graphics produced by the GAMs, and occasionally through statistical tests (e.g. using the Akaike Information Criterion). The papers analysing the CRFs are mentioned in Supplementary File S.4, where a specific column was added to show in which articles evidence of deviation from linearity was found. The individual studies that analysed the CRF in specific cities reported, in general, linear associations between daily (24-hour) mean PM_{10} and all-cause and cause-specific mortality (18 out of 21 articles that analysed linearity assumptions), and no threshold was detected. The exceptions were three papers that showed some non-linearity for all-cause, cardiovascular, respiratory, and cerebrovascular mortality (Chen et al., 2008; Dong et al., 2018; Li et al., 2013), with no references to potential thresholds. In the same line, the CRF between daily (24-hour) mean $PM_{2.5}$ and all-cause and cause-specific mortality was analysed in 11 articles, all of them showing indication of a linear behaviour, and with no indication of potential thresholds. The behaviour of the NO_2 (24-hour average) CRF was analysed in 16 papers, among which 13 did not show evidence of deviations from linearity. In three papers, the curve was found to be non-linear (Guo et al., 2017), especially at higher concentrations (Lu et al., 2015b), with a potential threshold at 37.6 μg/m^3 in the daily concentration of this pollutant (Moolgavkar et al., 2013). There were no papers dealing with CRFs for NO_2 (1-hour max.). As for O_3 (8-hour or 24-hour max.), the CRF was analysed in 11 articles, showing evidence of non-linear behaviour in 5 of them (Collart et al., 2018; Maji et al., 2017; Pascal et al., 2012; Qian et al., 2007; Williams et al., 2014). Non-linear models with thresholds at 100 μg/m^3 (Pascal et al., 2012) and at 60 to 80 μg/m^3 (Collart et al., 2018) were hypothesized, as well as linear models with thresholds at 65 μg/m^3 (in urban settings) (Atkinson et al., 2012) and 85 μg/m^3 (Collart et al., 2018).

3.5. Co-pollutant models

In general, two-pollutant models showed non-significant associations for PM_{10}, $PM_{2.5}$, and O_3 and mortality, when adjusting for a second air pollutant. This fact could be related to the low number of studies analysing co-pollutant models, which in turn impacts the statistical power of the tests. However, some exposure-outcome combinations showed positive associations, i.e. the association between PM_{10} and all-cause mortality and between $PM_{2.5}$ and respiratory mortality when adjusted by NO_2, and the association between NO_2 (24-hour average) and all-cause mortality when adjusted by PM or by O_3 (Table A.11 of the Appendix). Association values were higher than single-pollutant models in some combinations, and lower in others, irrespective of the statistical significance. Three-pollutant models were not evaluated, as in all the articles we haven't found more than two effect sizes reporting adjustments by the same combination of two pollutants, while a minimum of three effect sizes was required for *meta*-analysis. On the other hand, in a high proportion of articles the correlations between multiple pollutants were moderate to high (correlation coefficient > 0.4) (Dai and Zhou, 2017), which might threaten the validity of co-pollutant models (Supplementary File S.4).

3.6. Certainty of evidence

The only reason for downgrading the CoE in the different exposure-outcome combinations was the limitations regarding RoB in the articles (in two combinations). On the other hand, the evidence was upgraded due to large effect size (9 combinations), and the assumption of concentration–response curves in all but one of the exposure-outcome combinations, i.e. the relation between NO_2 (1-hour max.) and all-cause mortality. As for the final judgment regarding the CoE, it was high in 10 combinations, and moderate in one combination (NO_2 (1-hour max.) – all-cause mortality). The descriptions associated with this analysis, together with explanations of the rationale behind the judgements made, can be seen in Table 2.

4. Discussion

4.1. Summary of evidence

This systematic review and *meta*-analysis found evidence of a positive association between short-term exposure to PM_{10}, $PM_{2.5}$, NO_2 (24-hour average) and O_3 and all-cause mortality in humans, and between PM_{10}, $PM_{2.5}$ and cardiovascular, respiratory and cerebrovascular mortality. Conversely, NO_2 (1-hour max.) has shown a positive but non-significant association with all-cause mortality. These analyses, showing an increased risk of all-cause and cause-specific mortality associated with short-term exposure to air pollutants, are in consonance with previous evidence of published *meta*-analyses for PM (Atkinson et al., 2015, 2014; Bell et al., 2013; Shah et al., 2015), O_3 (Bell et al., 2014), and NO_2 or NO_x (Newell et al., 2018; Shah et al., 2015). In the same line, the Integrated Science Assessment for Particulate Matter, a comprehensive review of the United States Environmental Protection Agency, found that the newly published evidence from multiple high quality studies reinforces the causal relationship between $PM_{2.5}$ exposure and mortality, with biological plausibility for $PM_{2.5}$-related cardiovascular and respiratory morbidity (U.S. EPA, 2019). They also found evidence of an independent effect of $PM_{2.5}$ from co-pollutant models, and a linear CRF with no-threshold. Additionally, a recent study that evaluated short-term associations of PM_{10} and $PM_{2.5}$ with all-cause mortality in 652 cities has reported evidence of a positive relationship, with RRs very close to the values found in our study (Liu et al., 2019). In fact, that study analysed data from many of the cities that were included in our study, but the authors accessed the original data, while our study used RRs from published results. Anyway, the obtained RRs are similar, although less precise in our study. By contrast, we were able to perform many subgroup and sensitivity analyses, which were inputs to further evaluate heterogeneity, publication bias, and the CoE. The same occurred with a recent study analysing the effects of O_3 on mortality in 406 locations, which also found a positive effect (Vicedo-Cabrera et al., 2020).

Comparing to previous systematic reviews and multicity studies, our research had the advantage of including international and local information published in several databases, which contains articles communicating national or sub-national data. This comprehensive search allowed us a better control of publication bias, even though this issue could not be completely controlled. Furthermore, our study comprehensively analysed several air pollutants and outcomes, considering numerous subgroup analyses to evaluate the source of heterogeneity, and sensitivity analyses to assess the influence of model assumptions in the associations.

The magnitude of the associations was, as expected, lower than the associations between mortality and the exposure to these same air pollutants in the long-term; however, in epidemiology small risks applied to large populations are likely to represent a major health problem. Larger magnitudes in long-term studies might be attributable to cumulative effects (Beverland et al., 2012), due to the fact that short-term exposure studies only take into account a small proportion of the health effects; in addition, adverse effects are dependent of both concentration and length of the exposure (Pope, 2007). As for the study design, time series studies capture cases in which air pollution increases

both the risk of underlying diseases leading to frailty and the short-term risk of death among the frail, and cases in which air pollution is unrelated to risk of chronic diseases but short-term exposure increases mortality among persons who are frail (Künzli et al., 2001). In this sense, short-term exposure to air pollution might be acting as a trigger of long-term exposure deaths, or as a trigger of deaths related to underlying susceptibility from other causes.

Heterogeneity between studies was detected in some of the exposure-outcome combinations. However, heterogeneity in a certain degree is expected in epidemiological studies on air pollution, due to differences in populations, exposures, and study conditions (Goodman et al., 2015). Further, if the CRFs between the pollutants and the outcomes are non-linear, e.g. if the size of the effect depends on the mean level of air pollution in the area, as shown in the article by Liu and colleagues for PM (Liu et al., 2019), some unexplained heterogeneity may emerge. In addition, although the large majority of studies were conducted in the general population, at least one study among patients with co-morbidities was included in quantitative analysis for almost all the exposure-outcome combinations. This might have been another source of heterogeneity. Subgroup analyses were carried-out to explain potential sources of heterogeneity, in those exposure-outcome combinations that showed some degree of heterogeneity. These combinations included the associations between $PM_{2.5}$ and respiratory and cerebrovascular mortality, and the association between NO_2 (1-hour max.) and all-cause mortality. In none of these associations the heterogeneity was reduced after control for some co-variables, i.e. age group, sex, and continent, meaning that other unmeasured factors might possibly be acting as moderators.

The RoB in individual studies was high in a considerable proportion of articles, but only in the missing data domain. In this domain, several studies did not report methods used to impute missing data, nor declared the proportion or number of missing days for the exposure. Other domains were more stable, mainly due to the high standardization in time-series studies analysing short-term exposure to air pollutants. Other domain that influenced the RoB was the confounding domain, but in a small proportion of studies. However, the direction of the associations did not change after excluding articles showing high RoB, as demonstrated in the sensitivity analysis. A known confounder of the association between air pollution and mortality is ambient air temperature. This confounder has a recognized influence both on air pollution levels and mortality. In this line, we considered this factor as a critical confounder in the RoB domain, and in the E-value assessment.

As for publication bias, in 6 out of 11 combinations of outcomes and pollutants some asymmetries in the funnel plots were detected. This suggests that small studies showing non-significant effects could remain unpublished, and thus the true effect could be overestimated (Rothstein et al., 2008). However, these methods are indications of publication bias, but in some circumstances funnel plots can show misleading evidence of publication bias that is in fact related to heterogeneity (Levine et al., 2009; Peters et al., 2007; Terrin et al., 2003), as commonly observed in air pollution studies (Anderson et al., 2005). It is possible that the cause behind funnel plots asymmetry in this *meta*-analysis can be partially related to this explanation. A procedure that was implemented to reduce the potential effect of publication bias was to carry out a comprehensive literature review, which encompassed an inclusive search strategy and the search in regional databases, and allowed us the possibility of including other sources commonly considered gray literature (Rothstein et al., 2008). Nevertheless, the effect of publication bias was not completely avoided in all exposure-outcome combinations. However, the RRs of the exposure-outcome combinations were consistent in the main analysis and in the multicity sensitivity analysis, i.e. the associations were positive in all multicity analyses with the exception of cerebrovascular mortality. In this sense, the asymmetry of funnel plots could be more related to heterogeneity than to publication bias, as the probability of not publishing multicity studies due to adverse results is assumed to be minimal. All things considered, publication bias could have inflated the size of the true effect, but it could not have affected the general conclusion, i.e. a positive effect of air pollutants in mortality, as previously concluded in an article that assessed publication bias for PM (Anderson et al., 2005).

We found that the sensitivity analysis did not modify the results of the main analysis substantially, indicating that structural modelling decisions have not influenced the associations. The exception was for CCO designs, which showed less evident associations. This fact could be related to small statistical power, as the number of CCO studies is notably lower than ETS studies. However, it should be noted that CCO designs could be less prone to bias when compared with Poisson time-series designs (Carracedo-Martínez et al., 2010), and thus these differences could be reflecting more than merely methodological issues. This point deserves future confirmatory research.

In general, linear CRFs were found for PM_{10} and $PM_{2.5}$ associated with all-cause and cause-specific mortality. In contrast, some articles found a non-linear behaviour for NO_2 (24-hour average), with a potential threshold at 37.6 $\mu g/m^3$ average daily concentration. For O_3, a number of articles also found a non-linear behaviour, with potential thresholds in the range of 60–100 $\mu g/m^3$. The linear behaviour of some of the associations is consistent with the idea of a negative effect of pollutants even at low or background ambient concentrations, as was previously observed for $PM_{2.5}$, PM_{10}, and NO_2 (Kelly and Fussell, 2015; Li et al., 2015). This apparent absence of a safe level of air pollution below which health detrimental effects are negligible has deep implications for the development of ambient concentration limits in air quality guidelines, as even small reductions in air pollution levels might have a considerable impact in preventing mortality (Medina et al., 2004). A linear association and the absence of thresholds were found for O_3 in a recent paper, although levels of this pollutant below 70 $\mu g/m^3$ might be attributed to non-anthropogenic sources (Vicedo-Cabrera et al., 2020).

4.2. Co-pollutant models

Commonly, co-pollutant models show problems in the validity of associations between pollutants and mortality or morbidity outcomes, due to multicollinearity between pollutants. This is particularly recognized when high values of correlations between pollutants, measured as Pearson or Spearman correlation coefficients, are observed (Dominici et al., 2010). In many cases the multicollinearity between predictors has the effect of inflating the standard errors, leading to type II errors (Mason and Perreault, 1991). However, in this study the associations were observed even after the adjustment by a second pollutant, at least in some combinations of pollutants and outcomes. In almost all the associations, problems might have arisen with statistical power, as only a small proportion of studies analysed two or three-pollutant models. In fact, the analysis of three-pollutant models was not possible, due to the reduced number of articles assessing these associations (two or fewer studies per combination). Moreover, the results of these co-pollutant models should be interpreted with caution, given the high correlation between pollutants, and the known influence that multicollinearity has in model estimates (Stafoggia et al., 2017). Several statistical methods have been developed recently to analyse the effect of co-pollutant models, and at the same time dealing with multicollinearity issues (Stafoggia et al., 2017); nevertheless, these methods are not currently widespread, and none of the articles that were included in this review adopted these approaches. Independently of statistical issues, the challenges of multi-pollutant models are, first, to differentiate the direct effect of each pollutant from the spurious associations due to pollutants acting as surrogate of others (Billionnet et al., 2012). Second, multi-pollutant exposures might affect multiple target organs, leading to higher health risks. To achieve a better understanding of the effect of multiple pollutants, more research on the biological mechanisms of each pollutant should be carried out, preferably grouping pollutants according to their mode of action.

4.3. Strengths and limitations of the review

There are a few strengths that are worth being mentioned in relation to this review. First, the efforts made at different stages of the review, in order to overcome publication bias. This encompasses the inclusive search strategy, the query in regional databases that included gray literature, and the consideration of different study designs, i.e. ETS, CCO, panel and cohort studies. Second, sensitivity analyses showed that the associations reported in the main analyses were stable in different situations, which contributed to the reliability of results. In the majority of cases, the main results were stable in relation to methodological choices, i.e. the lag structure, the study design, and the RoB in some domains. Third and more relevant, the CoE was high in 10 out of 11 combinations, and moderate in one combination.

On the other hand, this review was subject to several limitations. First, the use of non-randomized observational studies made the analyses more prone to bias, related to failure in the control of potential confounders. This issue is common with other reviews that assess non-interventional studies. Second, several articles were potentially selected for inclusion after reading the title or the abstract, but the full-texts were not found. However, it should be mentioned that in all the missing articles in which the abstract was accessible, the associations between pollutants and mortality were positive. In this sense, it is reasonable to consider that the inclusion of these articles in quantitative analyses would not modify the direction of the associations. Third, the use of the E-value as an indication for the presence of an unmeasured confounder in the context of air pollution epidemiology is debatable, because the RRs estimated from these studies are typically lower than the association between outcomes and potential confounders. The use of the original cut-off point proposed in the standard GRADE approach ($RR > 2$) is not feasible for the same reason. Fourth, in some judgements for the CoE assessment we had no other option than using arbitrary thresholds. An example of this was the cut-off point of 100,000 deaths for the imprecision domain. More assessments to develop suitable criteria are to be developed in the future, in order to judge the minimum number of cases in short-term studies for precision evaluation. Finally, there were differences in the representation between continents. For example, the number of studies from North America, Europe and Asia was much higher than the number of studies from Latin America or Africa. Given the influence of several moderators in the associations, it is difficult that global estimates can be extrapolated to all regions. It is worth noting that low and middle economies were not fully represented in studies, and it is not clear if some variables, characteristic of these areas, could have an influence on the associations between pollutants and mortality. However, this fact is not a limitation of this particular study, but rather a shortcoming of research in this field. To overcome this deficiency, more resources should be allocated to research in these areas, preferably in the form of multi-collaborative projects.

The results of this systematic review contribute evidence on the influence of selected air pollutants on general and specific mortality, which is meant to be used for the update of the WHO Air Quality Guidelines. First, this information includes numerical values for the associations between air pollutants and specific risks, in order to be used in studies on the economic and disease burden attributable to air pollution, risk assessments, and other analyses. Second, consideration about single and multiple-pollutant exposures can be considered to discuss about unifactorial and multifactorial causation. Third, the results about CRFs and thresholds are useful for determination of air pollution limits in international recommendations, and national or local legislations. Fourth, subgroup analyses contribute to evaluation of the differential risk associated with different subpopulations. Finally, the assessment of the CoE of the different exposure-outcome combinations is relevant to understand the quality of the evidence presented in this review and in future guidelines.

5. Conclusions

Despite the limitations of this review, it was shown that an increase in outdoor concentrations of PM_{10}, $PM_{2.5}$, NO_2 and O_3 increases the risk of all-cause and cause-specific mortality in humans. These associations were proved to be stable through a number of sensitivity analyses, which enhance the validity of the conclusions presented here. As for CRFs, these curves have shown for NO_2 and O_3 non-linearity effects, and some evidence of thresholds, in line with previous estimates. The high consistency in the direction of the associations, and the high or moderate CoE reinforce the hypothesis of a positive association between air pollution and human mortality.

Declaration of Competing Interest

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

Acknowledgements

The systematic review was funded by the World Health Organization Regional Office for Europe, supported by the European Commission (DG Environment), Federal Ministry for the Environment, Nature Conservation and Nuclear Safety (Germany), Federal Ministry of Health (Germany), Government of the Republic of Korea, Federal Office for the Environment (Switzerland) and United States Environmental Protection Agency, and delivered as part of the evidence base that informs the ongoing development of WHO global air quality guidelines. All rights in the work, including ownership of the original work and copyright thereof, are vested in WHO. The authors alone are responsible for the views expressed in this publication and they do not necessarily represent the decisions or the stated policy of the World Health Organization. The authors would like to recognize the valuable contribution of Jos Verbeek, Román Pérez Velasco, Hanna Yang, Dorota Jarosińska, Rebecca Morgan, and experts from the Guideline Development Group (GDG). They all provided a consistent support during the full extension of the project, along with guidance, evaluation, and useful suggestions. The authors also recognize the contribution of Jiankun Fang, Yutong Zhu, and Ingu Kim, for the translation of four articles in Chinese and Korean.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envint.2020.105876.

References

Achilleos, S., Kioumourtzoglou, M.-A., Wu, C.-D., Schwartz, J.D., Koutrakis, P., Papatheodorou, S.I., 2017. Acute effects of fine particulate matter constituents on mortality: A systematic review and meta-regression analysis. Environ. Int. 109, 89–100. https://doi.org/10.1016/j.envint.2017.09.010.

Ahmed, K.K.M., Al Dhubaib, B.E., 2011. Zotero: A bibliographic assistant to researcher. J. Pharmacol. Pharmacother. 2, 303–305. https://doi.org/10.4103/0976-500X.85940.

Albert, J., Rizzo, M., 2012. R by Example, Use R!. Springer-Verlag, New York.

Anderson, H.R., Atkinson, R.W., Peacock, J.L., Sweeting, M.J., Marston, L., 2005. Ambient particulate matter and health effects: publication bias in studies of short-term associations. Epidemiol. Camb. Mass 16, 155–163.

Anderson, H.R., Favarato, G., Atkinson, R.W., 2013. Long-term exposure to air pollution and the incidence of asthma: meta-analysis of cohort studies. Air Qual. Atmosphere Health 6, 47–56. https://doi.org/10.1007/s11869-011-0144-5.

Atkinson, R.W., Kang, S., Anderson, H.R., Mills, I.C., Walton, H.A., 2014. Epidemiological time series studies of PM2.5 and daily mortality and hospital admissions: a systematic review and meta-analysis. Thorax 69, 660–665. https://doi.org/10.1136/thoraxjnl-2013-204492.

Atkinson, R.W., Mills, I.C., Walton, H.A., Anderson, H.R., 2015. Fine particle components and health–a systematic review and meta-analysis of epidemiological time series studies of daily mortality and hospital admissions. J. Expo. Sci. Environ. Epidemiol. 25, 208–214. https://doi.org/10.1038/jes.2014.63.

Atkinson, R.W., Yu, D., Armstrong, B.G., Pattenden, S., Wilkinson, P., Doherty, R.M., Heal, M.R., Anderson, H.R., 2012. Concentration-response function for ozone and daily mortality: results from five urban and five rural U.K. populations. Environ. Health Perspect. 120, 1411–1417. https://doi.org/10.1289/ehp.1104108.

Balshem, H., Helfand, M., Schünemann, H.J., Oxman, A.D., Kunz, R., Brozek, J., Vist, G.E., Falck-Ytter, Y., Meerpohl, J., Norris, S., Guyatt, G.H., 2011. GRADE guidelines: 3. Rating the quality of evidence. J. Clin. Epidemiol. 64, 401–406. https://doi.org/10.1016/j.jclinepi.2010.07.015.

Bell, M.L., Zanobetti, A., Dominici, F., 2014. Who is more affected by ozone pollution? A systematic review and meta-analysis. Am. J. Epidemiol. 180, 15–28. https://doi.org/10.1093/aje/kwu115.

Bell, M.L., Zanobetti, A., Dominici, F., 2013. Evidence on vulnerability and susceptibility to health risks associated with short-term exposure to particulate matter: a systematic review and meta-analysis. Am. J. Epidemiol. 178, 865–876. https://doi.org/10.1093/aje/kwt090.

Beverland, I.J., Cohen, G.R., Heal, M.R., Carder, M., Yap, C., Robertson, C., Hart, C.L., Agius, R.M., 2012. A Comparison of Short-term and Long-term Air Pollution Exposure Associations with Mortality in Two Cohorts in Scotland. Environ. Health Perspect. 120, 1280–1285. https://doi.org/10.1289/ehp.1104509.

Billionnet, C., Sherrill, D., Annesi-Maesano, I., GERIE study, 2012. Estimating the health effects of exposure to multi-pollutant mixture. Ann. Epidemiol. 22, 126–141. https://doi.org/10.1016/j.annepidem.2011.11.004.

Borenstein, M., 2019. Common Mistakes in Meta-Analysis: And How to Avoid Them. Biostat, Incorporated.

Borenstein, M., Hedges, L.V., Higgins, J.P.T., Rothstein, H.R., 2011. Introduction to Meta-Analysis Edición: 1, first ed. Wiley.

Carracedo-Martínez, E., Taracido, M., Tobias, A., Saez, M., Figueiras, A., 2010. Case-crossover analysis of air pollution health effects: a systematic review of methodology and application. Environ. Health Perspect. 118, 1173–1182. https://doi.org/10.1289/ehp.0901485.

Chen, G., Song, G., Jiang, L., Zhang, Y., Zhao, N., Chen, B., Kan, H., 2008. Short-term effects of ambient gaseous pollutants and particulate matter on daily mortality in Shanghai. China. J. Occup. Health 50, 41–47.

Chiolero, A., Santschi, V., Burnand, B., Platt, R.W., Paradis, G., 2012. Meta-analyses: with confidence or prediction intervals? Eur. J. Epidemiol. 27, 823–825. https://doi.org/10.1007/s10654-012-9738-y.

Chiusolo, M., Cadum, E., Stafoggia, M., Galassi, C., Berti, G., Faustini, A., Bisanti, L., Vigotti, M.A., Dessì, M.P., Cernigliaro, A., Mallone, S., Pacelli, B., Minerba, S., Simonato, L., Forastiere, F., EpiAir Collaborative Group, 2011. Short-Term Effects of Nitrogen Dioxide on Mortality and Susceptibility Factors in 10 Italian Cities: The EpiAir Study. Environ. Health Perspect. 119, 1233–1238. https://doi.org/10.1289/ehp.1002904.

Cohen, A.J., Brauer, M., Burnett, R., Anderson, H.R., Frostad, J., Estep, K., Balakrishnan, K., Brunekreef, B., Dandona, L., Dandona, R., Feigin, V., Freedman, G., Hubbell, B., Jobling, A., Kan, H., Knibbs, L., Liu, Y., Martin, R., Morawska, L., Pope, C.A., Shin, H., Straif, K., Shaddick, G., Thomas, M., van Dingenen, R., van Donkelaar, A., Vos, T., Murray, C.J.L., 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, 1907–1918. https://doi.org/10.1016/S0140-6736(17)30505-6.

Collart, P., Dramaix, M., Levêque, A., Mercier, G., Coppieters, Y., 2018. Concentration-response curve and cumulative effects between ozone and daily mortality: an analysis in Wallonia. Belgium. Int. J. Environ. Health Res. 28, 147–158. https://doi.org/10.1080/09603123.2018.1453050.

Cox, L.A.T., 2017. Do causal concentration-response functions exist? A critical review of associational and causal relations between fine particulate matter and mortality. Crit. Rev. Toxicol. 47, 603–631. https://doi.org/10.1080/10408444.2017.1311838.

Dai, Y.-H., Zhou, W.-X., 2017. Temporal and spatial correlation patterns of air pollutants in Chinese cities. PloS One 12, e0182724. https://doi.org/10.1371/journal.pone.0182724.

DerSimonian, R., Laird, N., 1986. Meta-analysis in clinical trials. Control. Clin. Trials 7, 177–188.

DeVries, R., Kriebel, D., Sama, S., 2017. Outdoor Air Pollution and COPD-Related Emergency Department Visits, Hospital Admissions, and Mortality: A Meta-Analysis. COPD 14, 113–121. https://doi.org/10.1080/15412555.2016.1216956.

Dominici, F., Peng, R.D., Barr, C.D., Bell, M.L., 2010. Protecting human health from air pollution: shifting from a single-pollutant to a multipollutant approach. Epidemiol. Camb. Mass 21, 187–194. https://doi.org/10.1097/EDE.0b013e3181cc86e8.

Dong, H., Yu, Y., Yao, S., Lu, Y., Chen, Z., Li, G., Yao, Y., Yao, X., Wang, S.-L., Zhang, Z., 2018. Acute effects of air pollution on ischaemic stroke onset and deaths: a time-series study in Changzhou. China. BMJ Open 8, e020425. https://doi.org/10.1136/bmjopen-2017-020425.

Egger, M., Davey Smith, G., Schneider, M., Minder, C., 1997. Bias in meta-analysis detected by a simple, graphical test. BMJ 315, 629–634.

Fajersztajn, L., Saldiva, P., Pereira, L.A.A., Leite, V.F., Buehler, A.M., 2017. Short-term effects of fine particulate matter pollution on daily health events in Latin America: a systematic review and meta-analysis. Int. J. Public Health 62, 729–738. https://doi.org/10.1007/s00038-017-0960-y.

Goldberg, M.S., Burnett, R.T., Stieb, D., 2003. A review of time-series studies used to evaluate the short-term effects of air pollution on human health. Rev. Environ. Health 18, 269–303.

Goodman, J.E., Petito Boyce, C., Sax, S.N., Beyer, L.A., Prueitt, R.L., 2015. Rethinking Meta-Analysis: Applications for Air Pollution Data and Beyond. Risk Anal. Off. Publ. Soc. Risk Anal. 35, 1017–1039. https://doi.org/10.1111/risa.12405.

Greenland, S., Thomas, D.C., 1982. On the need for the rare disease assumption in case-control studies. Am. J. Epidemiol. 116, 547–553.

Guo, Y., Ma, Y., Zhang, Y., Huang, S., Wu, Y., Yu, S., Zou, F., Cheng, J., 2017. Time series analysis of ambient air pollution effects on daily mortality. Environ. Sci. Pollut. Res. Int. 24, 20261–20272. https://doi.org/10.1007/s11356-017-9502-7.

Guyatt, G., Oxman, A.D., Akl, E.A., Kunz, R., Vist, G., Brozek, J., Norris, S., Falck-Ytter, Y., Glasziou, P., DeBeer, H., Jaeschke, R., Rind, D., Meerpohl, J., Dahm, P., Schünemann, H.J., 2011. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J. Clin. Epidemiol. 64, 383–394. https://doi.org/10.1016/j.jclinepi.2010.04.026.

Hartung, J., Knapp, G., 2001. A refined method for the meta-analysis of controlled clinical trials with binary outcome. Stat. Med. 20, 3875–3889.

Hozo, S.P., Djulbegovic, B., Hozo, I., 2005. Estimating the mean and variance from the median, range, and the size of a sample. BMC Med. Res. Methodol. 5, 13. https://doi.org/10.1186/1471-2288-5-13.

IntHout, J., Ioannidis, J.P.A., Rovers, M.M., Goeman, J.J., 2016. Plea for routinely presenting prediction intervals in meta-analysis. BMJ Open 6, e010247. https://doi.org/10.1136/bmjopen-2015-010247.

Kelly, F.J., Fussell, J.C., 2015. Air pollution and public health: emerging hazards and improved understanding of risk. Environ. Geochem. Health 37, 631–649. https://doi.org/10.1007/s10653-015-9720-1.

Knol, M.J., Vandenbroucke, J.P., Scott, P., Egger, M., 2008. What do case-control studies estimate? Survey of methods and assumptions in published case-control research. Am. J. Epidemiol. 168, 1073–1081. https://doi.org/10.1093/aje/kwn217.

Künzli, N., Medina, S., Kaiser, R., Quénel, P., Horak, F., Studnicka, M., 2001. Assessment of deaths attributable to air pollution: should we use risk estimates based on time series or on cohort studies? Am. J. Epidemiol. 153, 1050–1055. https://doi.org/10.1093/aje/153.11.1050.

Landrigan, P.J., Fuller, R., Acosta, N.J.R., Adeyi, O., Arnold, R., Basu, N.N., Baldé, A.B., Bertollini, R., Bose-O'Reilly, S., Boufford, J.I., Breysse, P.N., Chiles, T., Mahidol, C., Coll-Seck, A.M., Cropper, M.L., Fobil, J., Fuster, V., Greenstone, M., Haines, A., Hanrahan, D., Hunter, D., Khare, M., Krupnick, A., Lanphear, B., Lohani, B., Martin, K., Mathiasen, K.V., McTeer, M.A., Murray, C.J.L., Ndahimananjara, J.D., Perera, F., Potočnik, J., Preker, A.S., Ramesh, J., Rockström, J., Salinas, C., Samson, L.D., Sandilya, K., Sly, P.D., Smith, K.R., Steiner, A., Stewart, R.B., Suk, W.A., van Schayck, O.C.P., Yadama, G.N., Yumkella, K., Zhong, M., 2018. The Lancet Commission on pollution and health. Lancet Lond. Engl. 391, 462–512. https://doi.org/10.1016/S0140-6736(17)32345-0.

Levine, T.R., Asada, K.J., Carpenter, C., 2009. Sample Sizes and Effect Sizes are Negatively Correlated in Meta-Analyses: Evidence and Implications of a Publication Bias Against NonSignificant Findings. Commun. Monogr. 76, 286–302. https://doi.org/10.1080/03637750903074685.

Li, G.X., Zhou, M.G., Zhang, Y.J., Cai, Y., Pan, X.C., 2013. Seasonal effects of PM10 concentrations on mortality in Tianjin, China: a time-series analysis. J. Public Health 21, 135–144. https://doi.org/10.1007/s10389-012-0529-4.

Li, H., Chen, R., Meng, X., Zhao, Z., Cai, J., Wang, C., Yang, C., Kan, H., 2015. Short-term exposure to ambient air pollution and coronary heart disease mortality in 8 Chinese cities. Int. J. Cardiol. 197, 265–270. https://doi.org/10.1016/j.ijcard.2015.06.050.

Li, M.-H., Fan, L.-C., Mao, B., Yang, J.-W., Choi, A.M.K., Cao, W.-J., Xu, J.-F., 2016. Short-term Exposure to Ambient Fine Particulate Matter Increases Hospitalizations and Mortality in COPD: A Systematic Review and Meta-analysis. Chest 149, 447–458. https://doi.org/10.1378/chest.15-0513.

Liu, C., Chen, R., Sera, F., Vicedo-Cabrera, A.M., Guo, Y., Tong, S., Coelho, M.S.Z.S., Saldiva, P.H.N., Lavigne, E., Matus, P., Valdes Ortega, N., Osorio Garcia, S., Pascal, M., Stafoggia, M., Scortichini, M., Hashizume, M., Honda, Y., Hurtado-Díaz, M., Cruz, J., Nunes, B., Teixeira, J.P., Kim, H., Tobias, A., Íñiguez, C., Forsberg, B., Åström, C., Ragettli, M.S., Guo, Y.-L., Chen, B.-Y., Bell, M.L., Wright, C.Y., Scovronick, N., Garland, R.M., Milojevic, A., Kyselý, J., Urban, A., Orru, H., Indermitte, E., Jaakkola, J.J.K., Ryti, N.R.I., Katsouyanni, K., Analitis, A., Zanobetti, A., Schwartz, J., Chen, J., Wu, T., Cohen, A., Gasparrini, A., Kan, H., 2019. Ambient Particulate Air Pollution and Daily Mortality in 652 Cities. N. Engl. J. Med. 381, 705–715. https://doi.org/10.1056/NEJMoa1817364.

Lu, F., Xu, D., Cheng, Y., Dong, S., Guo, C., Jiang, X., Zheng, X., 2015a. Systematic review and meta-analysis of the adverse health effects of ambient PM2.5 and PM10 pollution in the Chinese population. Environ. Res. 136, 196–204. https://doi.org/10.1016/j.envres.2014.06.029.

Lu, F., Zhou, L., Xu, Y., Zheng, T., Guo, Y., Wellenius, G.A., Bassig, B.A., Chen, X., Wang, H., Zheng, X., 2015b. Short-term effects of air pollution on daily mortality and years of life lost in Nanjing. China. Sci. Total Environ. 536, 123–129. https://doi.org/10.1016/j.scitotenv.2015.07.048.

Maji, S., Ahmed, S., Siddiqui, W.A., Ghosh, S., 2017. Short term effects of criteria air pollutants on daily mortality in Delhi. India. Atmos. Environ. 150, 210–219. https://doi.org/10.1016/j.atmosenv.2016.11.044.

Mason, C.H., Perreault, W.D., 1991. Collinearity, Power, and Interpretation of Multiple Regression Analysis. J. Mark. Res. 28, 268–280. https://doi.org/10.2307/3172863.

Medina, S., Plasencia, A., Ballester, F., Mücke, H.G., Schwartz, J., 2004. Apheis: public health impact of PM10 in 19 European cities. J. Epidemiol. Community Health 58, 831–836. https://doi.org/10.1136/jech.2003.016386.

Mills, I.C., Atkinson, R.W., Kang, S., Walton, H., Anderson, H.R., 2015. Quantitative systematic review of the associations between short-term exposure to nitrogen dioxide and mortality and hospital admissions. BMJ Open 5, e006946. https://doi.org/10.1136/bmjopen-2014-006946.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., PRISMA Group, 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ 339https://doi.org/10.1136/bmj.b2535. b2535.

Moolgavkar, S.H., McClellan, R.O., Dewanji, A., Turim, J., Luebeck, E.G., Edwards, M., 2013. Time-series analyses of air pollution and mortality in the United States: a subsampling approach. Environ. Health Perspect. 121, 73–78. https://doi.org/10.

1289/ehp.1104507.
Morgan, R.L., Thayer, K.A., Bero, L., Bruce, N., Falck-Ytter, Y., Ghersi, D., Guyatt, G., Hooijmans, C., Langendam, M., Mandrioli, D., Mustafa, R.A., Rehfuess, E.A., Rooney, A.A., Shea, B., Silbergeld, E.K., Sutton, P., Wolfe, M.S., Woodruff, T.J., Verbeek, J.H., Holloway, A.C., Santesso, N., Schünemann, H.J., 2016. GRADE: Assessing the quality of evidence in environmental and occupational health. Environ. Int. 92–93, 611–616. https://doi.org/10.1016/j.envint.2016.01.004.
U.S. EPA. Integrated Science Assessment (ISA) for Particulate Matter (Final Report, 2019). U.S. Environmental Protection Agency, Washington, DC, EPA/600/R-19/188, 2019.
Morgan, R.L., Thayer, K.A., Santesso, N., Holloway, A.C., Blain, R., Eftim, S.E., Goldstone, A.E., Ross, P., Ansari, M., Akl, E.A., Filippini, T., Hansell, A., Meerpohl, J.J., Mustafa, R.A., Verbeek, J., Vinceti, M., Whaley, P., Schünemann, H.J., GRADE Working Group, 2019. A risk of bias instrument for non-randomized studies of exposures: A users' guide to its application in the context of GRADE. Environ. Int. 122, 168–184. https://doi.org/10.1016/j.envint.2018.11.004.
Morgan, R.L., Whaley, P., Thayer, K.A., Schünemann, H.J., 2018. Identifying the PECO: A framework for formulating good questions to explore the association of environmental and other exposures with health outcomes. Environ. Int. 121, 1027–1031. https://doi.org/10.1016/j.envint.2018.07.015.
Newell, K., Kartsonaki, C., Lam, K.B.H., Kurmi, O., 2018. Cardiorespiratory health effects of gaseous ambient air pollution exposure in low and middle income countries: a systematic review and meta-analysis. Environ. Health Glob. Access Sci. Source 17, 41. https://doi.org/10.1186/s12940-018-0380-3.
Newell, K., Kartsonaki, C., Lam, K.B.H., Kurmi, O.P., 2017. Cardiorespiratory health effects of particulate ambient air pollution exposure in low-income and middle-income countries: a systematic review and meta-analysis. Lancet Planet. Health 1, e368–e380. https://doi.org/10.1016/S2542-5196(17)30166-3.
Ostro, B., Spadaro, J.V., Gumy, S., Mudu, P., Awe, Y., Forastiere, F., Peters, A., 2018. Assessing the recent estimates of the global burden of disease for ambient air pollution: Methodological changes and implications for low- and middle-income countries. Environ. Res. 166, 713–725. https://doi.org/10.1016/j.envres.2018.03.001.
Pace, N.D., Multani, J.K., 2018. On the Reporting of Odds Ratios and Risk Ratios. Nutrients 10. https://doi.org/10.3390/nu10101512.
Pascal, M., Wagner, V., Chatignoux, E., Falq, G., Corso, M., Blanchard, M., Host, S., Larrieu, S., Pascal, L., Declercq, C., 2012. Ozone and short-term mortality in nine French cities: Influence of temperature and season. Atmos. Environ. 62, 566–572. https://doi.org/10.1016/j.atmosenv.2012.09.009.
Peters, J.L., Sutton, A.J., Jones, D.R., Abrams, K.R., Rushton, L., 2007. Performance of the trim and fill method in the presence of publication bias and between-study heterogeneity. Stat. Med. 26, 4544–4562. https://doi.org/10.1002/sim.2889.
Pope, C.A., 2007. Mortality effects of longer term exposures to fine particulate air pollution: review of recent epidemiological evidence. Inhal. Toxicol. 19 (Suppl 1), 33–38. https://doi.org/10.1080/08958370701492961.
Qian, Z., He, Q., Lin, H.-M., Kong, L., Liao, D., Yang, N., Bentley, C.M., Xu, S., 2007. Short-term effects of gaseous pollutants on cause-specific mortality in Wuhan. China. J. Air Waste Manag. Assoc. 1995 (57), 785–793.
Risk of bias assessment instrument for systematic reviews informing WHO global air quality guidelines (2020) [WWW Document], 2020. URL http://www.euro.who.int/en/health-topics/environment-and-health/air-quality/publications/2020/risk-of-bias-assessment-instrument-for-systematic-reviews-informing-who-global-air-quality-guidelines-2020 (accessed 3.17.20).
Rothstein, H.R., Sutton, A.J., Borenstein, M. (Eds.), 2008. Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments, 1 edition. ed. Wiley.
Rücker, G., Schwarzer, G., Carpenter, J.R., Schumacher, M., 2008. Undue reliance on I(2) in assessing heterogeneity may mislead. BMC Med. Res. Methodol. 8, 79. https://doi.org/10.1186/1471-2288-8-79.
Schwarzer, G., Carpenter, J.R., Rücker, G., 2015. Meta-Analysis with R, Use R!. Springer International Publishing.
Shah, A.S.V., Langrish, J.P., Nair, H., McAllister, D.A., Hunter, A.L., Donaldson, K., Newby, D.E., Mills, N.L., 2013. Global association of air pollution and heart failure: a systematic review and meta-analysis. Lancet Lond. Engl. 382, 1039–1048. https://doi.org/10.1016/S0140-6736(13)60898-3.
Shah, A.S.V., Lee, K.K., McAllister, D.A., Hunter, A., Nair, H., Whiteley, W., Langrish, J.P., Newby, D.E., Mills, N.L., 2015. Short term exposure to air pollution and stroke: systematic review and meta-analysis. BMJ 350, h1295. https://doi.org/10.1136/bmj.h1295.
Shang, Y., Sun, Z., Cao, J., Wang, X., Zhong, L., Bi, X., Li, H., Liu, W., Zhu, T., Huang, W., 2013. Systematic review of Chinese studies of short-term exposure to air pollution and daily mortality. Environ. Int. 54, 100–111. https://doi.org/10.1016/j.envint.2013.01.010.
Song, X., Wang, S., Hu, Y., Yue, M., Zhang, T., Liu, Y., Tian, J., Shang, K., 2017. Impact of ambient temperature on morbidity and mortality: An overview of reviews. Sci. Total Environ. 586, 241–254. https://doi.org/10.1016/j.scitotenv.2017.01.212.
Stafoggia, M., Breitner, S., Hampel, R., Basagaña, X., 2017. Statistical Approaches to Address Multi-Pollutant Mixtures and Multiple Exposures: the State of the Science. Curr. Environ. Health Rep. 4, 481–490. https://doi.org/10.1007/s40572-017-0162-z.
Sterne, J.A.C., Sutton, A.J., Ioannidis, J.P.A., Terrin, N., Jones, D.R., Lau, J., Carpenter, J., Rücker, G., Harbord, R.M., Schmid, C.H., Tetzlaff, J., Deeks, J.J., Peters, J., Macaskill, P., Schwarzer, G., Duval, S., Altman, D.G., Moher, D., Higgins, J.P.T., 2011. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. BMJ 343, d4002. https://doi.org/10.1136/bmj.d4002.
Suh, H.H., Bahadori, T., Vallarino, J., Spengler, J.D., 2000. Criteria air pollutants and toxic air pollutants. Environ. Health Perspect. 108 (Suppl 4), 625–633. https://doi.org/10.1289/ehp.00108s4625.
Taneepanichskul, N., Gelaye, B., Grigsby-Toussaint, D.S., Lohsoonthorn, V., Jimba, M., Williams, M.A., 2018. Short-term effects of particulate matter exposure on daily mortality in Thailand: a case-crossover study. Air Qual. Atmosphere Health 11, 639–647. https://doi.org/10.1007/s11869-018-0571-7.
Terrin, N., Schmid, C.H., Lau, J., Olkin, I., 2003. Adjusting for publication bias in the presence of heterogeneity. Stat. Med. 22, 2113–2126. https://doi.org/10.1002/sim.1461.
Touloumi, G., Samoli, E., Quenel, P., Paldy, A., Anderson, R.H., Zmirou, D., Galan, I., Forsberg, B., Schindler, C., Schwartz, J., Katsouyanni, K., 2005. Short-term effects of air pollution on total and cardiovascular mortality: the confounding effect of influenza epidemics. Epidemiol. Camb. Mass 16, 49–57. https://doi.org/10.1097/01.ede.0000142152.62400.13.
Ueda, K., Nitta, H., Ono, M., 2009. Effects of fine particulate matter on daily mortality for specific heart diseases in Japan. Circ. J. Off. J. Jpn. Circ. Soc. 73, 1248–1254. https://doi.org/10.1253/circj.cj-08-1149.
VanderWeele, T.J., Ding, P., 2017. Sensitivity Analysis in Observational Research: Introducing the E-Value. Ann. Intern. Med. 167, 268–274. https://doi.org/10.7326/M16-2607.
Vicedo-Cabrera, A.M., Sera, F., Liu, C., Armstrong, B., Milojevic, A., Guo, Y., Tong, S., Lavigne, E., Kyselý, J., Urban, A., Orru, H., Indermitte, E., Pascal, M., Huber, V., Schneider, A., Katsouyanni, K., Samoli, E., Stafoggia, M., Scortichini, M., Hashizume, M., Honda, Y., Ng, C.F.S., Hurtado-Diaz, M., Cruz, J., Silva, S., Madureira, J., Scovronick, N., Garland, R.M., Kim, H., Tobias, A., Íñiguez, C., Forsberg, B., Åström, C., Ragettli, M.S., Röösli, M., Guo, Y.-L.L., Chen, B.-Y., Zanobetti, A., Schwartz, J., Bell, M.L., Kan, H., Gasparrini, A., 2020. Short term association between ozone and mortality: global two stage time series study in 406 locations in 20 countries. BMJ 368, m108. https://doi.org/10.1136/bmj.m108.
Wang, Y., Eliot, M.N., Wellenius, G.A., 2014. Short-term changes in ambient particulate matter and risk of stroke: a systematic review and meta-analysis. J. Am. Heart Assoc. 3. https://doi.org/10.1161/JAHA.114.000983.
WHO | Air quality guidelines - global update 2005 [WWW Document], 2018. . WHO. URL https://www.who.int/phe/health_topics/outdoorair/outdoorair_aqg/en/ (accessed 11.12.18).
WHO|International Classification of Diseases (ICD) Information Sheet [WWW Document], n.d. . WHO. URL http://www.who.int/classifications/icd/factsheet/en/ (accessed 11.12.18).
WHO Regional Office for Europe, 2013. Review of evidence on health aspects of air pollution – REVIHAAP Project: Technical Report. WHO Regional Office for Europe, Copenhagen.
Williams, M.L., Atkinson, R.W., Anderson, H.R., Kelly, F.J., 2014. Associations between daily mortality in London and combined oxidant capacity, ozone and nitrogen dioxide. Air Qual. Atmosphere Health 7, 407–414. https://doi.org/10.1007/s11869-014-0249-8.
Zhang, Z., Wang, J., Lu, W., 2018. Exposure to nitrogen dioxide and chronic obstructive pulmonary disease (COPD) in adults: a systematic review and meta-analysis. Environ. Sci. Pollut. Res. Int. 25, 15133–15145. https://doi.org/10.1007/s11356-018-1629-7.
WHO Handbook for Guideline Development – 2nd Edition, 2014. URL https://apps.who.int/iris/handle/10665/145714 (accessed 3.25.20).