Meta-analysis

ISEE ENVIRONMENTAL EPIDEMIOLOGY

OPEN

Choices of morbidity outcomes and concentration–response functions for health risk assessment of long-term exposure to air pollution

Francesco Forastiere[a,b,*], Joseph V. Spadaro[c,d], Carla Ancona[e], Zorana Jovanovic Andersen[f], Ilaria Cozzi[e], Sophie Gumy[c], Dejan Loncar[c], Pierpaolo Mudu[g], Sylvia Medina[h], Roman Perez Velasco[g], Heather Walton[b,i], Jiawei Zhang[f], Michal Krzyzanowski[b]

Background: Air pollution health risk assessment (HRA) has been typically conducted for all causes and cause-specific mortality based on concentration–response functions (CRFs) from meta-analyses that synthesize the evidence on air pollution health effects. There is a need for a similar systematic approach for HRA for morbidity outcomes, which have often been omitted from HRA of air pollution, thus underestimating the full air pollution burden. We aimed to compile from the existing systematic reviews and meta-analyses CRFs for the incidence of several diseases that could be applied in HRA. To achieve this goal, we have developed a comprehensive strategy for the appraisal of the systematic reviews and meta-analyses that examine the relationship between long-term exposure to particulate matter with an aerodynamic diameter smaller than 2.5 μm ($PM_{2.5}$), nitrogen dioxide (NO_2), or ozone (O_3) and incidence of various diseases.

Methods: To establish the basis for our evaluation, we considered the causality determinations provided by the US Environmental Protection Agency Integrated Science Assessment for $PM_{2.5}$, NO_2, and O_3. We developed a list of pollutant/outcome pairs based on these assessments and the evidence of a causal relationship between air pollutants and specific health outcomes. We conducted a comprehensive literature search using two databases and identified 75 relevant systematic reviews and meta-analyses for $PM_{2.5}$ and NO_2. We found no relevant reviews for long-term exposure to ozone. We evaluated the reliability of these studies using an adaptation of the AMSTAR 2 tool, which assesses various characteristics of the reviews, such as literature search, data extraction, statistical analysis, and bias evaluation. The tool's adaptation focused on issues relevant to studies on the health effects of air pollution. Based on our assessment, we selected reviews that could be credible sources of CRF for HRA. We also assessed the confidence in the findings of the selected systematic reviews and meta-analyses as the sources of CRF for HRA. We developed specific criteria for the evaluation, considering factors such as the number of included studies, their geographical distribution, heterogeneity of study results, the statistical significance and precision of the pooled risk estimate in the meta-analysis, and consistency with more recent studies. Based on our assessment, we classified the outcomes into three lists: list A (a reliable quantification of health effects is possible in an HRA), list B+ (HRA is possible, but there is greater uncertainty around the reliability of the CRF compared to those included on list A), and list B– (HRA is not recommended because of the substantial uncertainty of the CRF).

Results: In our final evaluation, list A includes six CRFs for $PM_{2.5}$ (asthma in children, chronic obstructive pulmonary disease, ischemic heart disease events, stroke, hypertension, and lung cancer) and three outcomes for NO_2 (asthma in children and in adults, and acute lower respiratory infections in children). Three additional outcomes (diabetes, dementia, and autism spectrum disorders) for $PM_{2.5}$ were included in list B+. Recommended CRFs are related to the incidence (onset) of the diseases. The International Classification of Diseases, 10th revision codes, age ranges, and suggested concentration ranges are also specified to ensure consistency and applicability in an HRA. No specific suggestions were given for ozone because of the lack of relevant systematic reviews.

Conclusion: The suggestions formulated in this study, including CRFs selected from the available systematic reviews, can assist in conducting reliable HRAs and contribute to evidence-based decision-making in public health and environmental policy. Future research should continue to update and refine these suggestions as new evidence becomes available and methodologies evolve.

Keywords: Air pollution; Particulate matter; Nitrogen dioxide; Morbidity; Systematic review; Meta-analysis

[a]National Research Council, IFT, Palermo, Italy; [b]Environmental Research Group, Imperial College, London, United Kingdom; [c]World Health Organization, Headquarters, Geneva, Switzerland; [d]Spadaro Environmental Research Consultants (SERC), Philadelphia, Pennsylvania; [e]Department of Epidemiology, Lazio Regional Health Service, Local Health Unit Roma 1, Rome, Italy; [f]Department of Public Health, University of Copenhagen, Copenhagen, Denmark; [g]World Health Organization (WHO), European Center for Environment and Health, Bonn, Germany; [h]Santé Publique France, Paris, France; and [i]National Institute of Health Research Health Protection Research Unit on Environmental Exposures and Health at Imperial College London, London, United Kingdom

The content of this article represents the personal views of the authors and does not represent the views of the authors' employers and associated institutions, that is, the NIHR, UK Health Security Agency, UK Department of Health and Social Care or World Health Organization. Where authors are identified as staff of the World Health Organization, the authors alone are responsible for the views

What this study adds:

This study determines reliable concentration–response functions (CRFs) for morbidity outcomes to be applied in health risk assessments (HRAs) of long-term exposure to air pollution. It offers a comprehensive and rigorous approach to selecting CRFs based on causality considerations, reliability appraisal of the systematic reviews and meta-analyses providing the CRFs, and confidence in the epidemiological evidence as the sources of the CRFs. The conclusions of the study contribute to the accuracy and reliability of HRAs, providing guidance for authorities and researchers in assessing the impact of air pollution on morbidity. This research fills a gap in the literature by integrating and systemizing the available evidence and facilitating informed decision-making in the public health and environmental policy arena.

Introduction

Long-term exposure to ambient air pollution has been convincingly linked to mortality. concentration–response functions (CRFs), mathematical relationships between air pollutant concentrations and mortality, have been proposed based on systematic reviews[1,2] and complemented by more recent studies such as the large European ELAPSE (Effects of Low-Level Air Pollution: A Study in Europe) study.[3] However, health risk assessment (HRA), including quantitative health impact assessment, traditionally oriented toward mortality impact, should be supplemented by evaluating the impact on morbidity to calculate the overall health burden of air pollution. Several morbidity outcomes have been linked to long-term exposure to air pollution, but a comprehensive, systematic, and recently updated evaluation of suitable CRFs is not available. In addition, it should be noted that a series of recent empirical studies suggest that air pollution, through its effects on morbidity, also has significant negative impacts on the economy through direct and indirect costs, including medical expenditures, labor productivity losses, and welfare costs.[4]

A review of epidemiological studies published before 2013 with recommendations concerning the assessment of the impacts of ambient air pollution on mortality and morbidity was formulated in the World Health Organization (WHO) HRAPIE (Health Risks of Air Pollution in Europe) project.[5,6] The causality assessment of the pollutant/outcome associations in HRAPIE was provided by the "Review of EVIdence on Health Aspects of Air Pollution (REVIHAAP)" project.[7] The HRAPIE report contained a list of CRFs for cost–benefit analyses in Europe, but they have been applied in various applications, including burden and impact analyses (burden provides a snapshot of the overall air pollution health effect at a location; impact analyses compare the burden of pollution across different policies). The recommendations from the HRAPIE report regarding the morbidity outcomes related to long-term exposure were few and included "prevalence of bronchitis in children aged 6–12 (or 6–18) years" and "incidence of chronic bronchitis in adults (age, 18+ years)" for particulate matter with an aerodynamic diameter smaller than 10 μm (PM_{10}) and "prevalence of bronchitis symptoms in asthmatic children aged 5–14 years" for nitrogen dioxide (NO_2). All these pollutant/morbidity pairs were placed in a category described as containing pollutant/outcome pairs for which there is significant uncertainty about the quality of the evidence used to quantify effects.

Since 2013, new evidence on the association between air pollutants and morbidity has been published in the scientific literature, and several official documents have proposed various choices of CRFs for HRA. In the United Kingdom, Public Health England published a comprehensive assessment of the burden of air pollution in 2018, accompanied by an extensive list of morbidity outcomes for HRA.[8] In 2023, the UK Committee on the Medical Effects of Air Pollutants (COMEAP) provided recommendations for quantifying health effects associated with air pollutants, including morbidity.[9] Recent statements from COMEAP on particulate matter with an aerodynamic diameter smaller than 2.5 μm ($PM_{2.5}$) and mortality (long-term exposure),[10] as well as on-air pollutants and hospital admissions,[11] have been published, and an extensive review of the methodology has been produced.[12] A report on morbidity effects funded by the Swedish Environmental Protection Agency (EPA) is also available.[13] The recent impact assessment study[14] to support the revision of the EU Ambient Air Quality Directives (AAQD) included several morbidity CRFs associated with long-term exposure to $PM_{2.5}$ (such as stroke, nonfatal myocardial infarction, asthma in children, chronic obstructive pulmonary disease [COPD], type 2 diabetes, and lung cancer). Finally, the European Environmental Agency (EEA) has also considered morbidity effects in a recent report.[15] All these studies produced or used CRFs that need to be integrated into a single set of associations following a systematic and coherent approach.

The current work aimed to compile, from the available systematic reviews and meta-analyses, the CRFs for the incidence of several diseases that could be applied in HRAs. To achieve this goal, we developed a comprehensive strategy for selecting suitable CRFs from recent systematic reviews. This strategy involves a methodical appraisal of the existing systematic reviews and meta-analyses that examine the relationship between long-term exposure to fine particulates ($PM_{2.5}$), nitrogen dioxide (NO_2), and ozone (O_3) and various morbidity outcomes. Additionally, we formulated suggestions for applying these associations based on specific characteristics extracted from the studies that provided the CRFs. These suggestions aim to guide researchers and authorities in utilizing the appropriate CRFs for their HRA studies, which have further economic implications. This work is a part of a broader WHO project, "Estimation of Morbidity from Air Pollution and its Economic Costs" (EMAPEC), https://www.who.int/activities/estimating-the-morbidity-from-air-pollution-and-its-economic-costs.

Methods

Rationale

The epidemiological evidence indicates that the effects of long-term exposure to air pollution on mortality have a larger impact on public health than the effects of short-term exposure.[16–18] The associations between air pollutants and mortality reported in cohort studies usually represent the effects of long-term exposure to pollutants (incidence and progression of a disease) and short-term exposure to increased concentrations (exacerbation of pre-existing medical conditions). In other words, long-term studies on mortality encompass cumulative long- and short-term exposure burdens. In morbidity studies focused on disease incidence, long-term exposure contributes to progressive health deterioration and, consequently, to the development of new illnesses over time. Furthermore, short-term exposure also results in the manifestation of acute health outcomes and contributes to complications among persons already afflicted by underlying medical conditions. In this work, we focused on the CRFs for the incidence of morbidity outcomes due to long-term exposure to ambient air pollution. Our approach was based on the following considerations.

(1) Only well-defined clinical conditions should be considered. Potential physiopathological markers of the effects, such as lung function, intima-media thickness, and cognitive decline, although important to understand the air pollution mechanism of action, should be excluded.

expressed in this article, and they do not necessarily represent the decisions, policy, or views of the World Health Organization.

Data for the work has been taken from the published literature with methods described in the text.

SDC *Supplemental digital content is available through direct URL citations in the HTML and PDF versions of this article (www.environepidem.com).*

**Corresponding Author. Address: School of Public Health, Imperial College, Michael Uren Biomedical Engineering Hub. White City Campus, Wood Lane, London W12 OBZ, UK. E-mail: Fran.forastiere@gmail.com (F. Forastiere).*

Environmental Epidemiology (2024) 8:e314

Received 8 September, 2023; Accepted 14 May, 2024

Published online 25 June 2024

DOI: 10.1097/EE9.0000000000000314

(2) The potential health outcomes should be supported by robust and persuasive evidence of a causal relationship with the pollutant of concern. Therefore, scientific findings should be well-founded and rely on human evidence, namely, the integration of epidemiological, toxicological and clinical studies, and mechanistic explanations. Meeting this requirement necessitates a convincing causality assessment, which requires a thorough multidisciplinary evaluation that only a few authoritative organizations can provide.
(3) The human evidence should be adequately quantified in well-characterized multiple observational studies conducted across different locations of the world. These studies should indicate a statistically significant association with a relatively narrow margin of uncertainty (high precision). Overall, any considerable influence of chance, bias, and confounding should be minimized.

The choice of the appropriate CRFs is based on a consensus among experts followed by a peer review process, as was the case for the WHO HRAPIE 2013 project.[5] Further, a clear and transparent justification for the choices is needed. The selection is based on the available synthesis of the evidence from published systematic reviews and meta-analyses. First, the reliability of the systematic reviews and meta-analyses should be evaluated, as sometimes problems in methodological rigor have been reported.[18] Considering that the most recent systematic reviews are able to capture the newest studies, we prioritized such reviews in our search for CRFs. Next, for those systematic reviews of adequate reliability, the confidence in the findings of the selected systematic reviews and meta-analyses should be assessed.

Principles and PECOS

In our review of systematic reviews, we have followed the published recommendations for conducting an umbrella review.[19] It should be noted, however, that the scope of our search was not to evaluate the strength of the association from all the available evidence, as in the case of hazard identification, but rather to identify the most appropriate and reliable pooled effect estimates for quantitative risk assessment. We identified systematic reviews of studies on the association of morbidity due to selected clinical conditions (cardiovascular, respiratory, neurological, metabolic diseases, and lung cancer) with long-term exposure to air pollutants. Since we aimed to identify CRFs, the eligible systematic reviews must contain a meta-analysis of the included studies. We have considered the three pollutants, $PM_{2.5}$, NO_2, and O_3, because the observational evidence of their adverse health effects is the best developed. Exposure to these three pollutants at levels above the WHO Air Quality Guidelines (AQGs) 2021[20] is common in most settings at global and European levels.[21]

We formulated the following PECOS question: in a population (P), including subgroups of susceptible individuals, what is the increase in the health risk of selected clinical conditions (O) per 10 μg/m^3 increase (C) in long-term exposure (meaning an exposure window lasting several months to multiple years) to ambient air concentrations of $PM_{2.5}$, NO_2, or O_3 (E) that are reported in a meta-analysis of cohort or case–control studies on incidence (S)?

To select the CRFs for long-term exposure, we followed the steps below.

(1) Evaluate the causality determination;
(2) Produce a provisional list of pollutant/outcome pairs;
(3) Search the literature for available systematic reviews and meta-analyses;
(4) Appraise the systematic review and meta-analyses;
(5) Validate the data of the systematic reviews providing the CRFs; and if necessary, revise the meta-analysis and the forest plots;
(6) Evaluate the confidence in the epidemiological evidence supporting the CRFs and suggest their application in an HRA.

The review protocol was registered in Prospero on 15 February 2023 (CRD42023397145, https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=397145), (with an addendum submitted on 11 April 2023). In the late stages of our work, we decided to make some deviations from the protocol: (1) we did not evaluate low birth weight to limit ourselves to clear clinical outcomes; (2) we decided not to score the results of the AMSTAR2-EH evaluation (see the following sections) but to present all the results for each reviewed systematic review (SR); (3) We avoided using the criteria for the strength of the evidence as a "grading" instrument, and we used the criteria as guidance for CRFs selection. The steps undertaken are described in greater detail below.

Causality determination

We evaluated the causality determination using the US EPA Integrated Science Assessment (ISA) reports ($PM_{2.5}$ 2019[22] and update 2022,[23] NO_2 2016,[24] and O_3 2020[25]) and established a list of potential pollutant/outcome pairs (step 2). The US EPA ISA was used because it accurately represents the scientific knowledge on the health effects of air pollutants by considering evidence from controlled human exposure, animal toxicology, and epidemiology studies. It should be noted that the US EPA ISA also considered other size fractions of particulate matter besides $PM_{2.5}$. However, the most substantial scientific evidence, including mechanistic studies, supporting the relationship between exposure to PM and health effects was for $PM_{2.5}$, so our review focused on this fraction.

Table 1 shows the US EPA causality assessments for $PM_{2.5}$, NO_2, and O_3. Following the NO_2 US EPA 2016 determination, we focused on:

(1) Causal, "... the pollutant has been shown to result in health effects in studies in which chance, confounding, and other biases could be ruled out with reasonable confidence" or
(2) Likely to be a causal relationship, "... the pollutant has been shown to result in health effects in studies where results are not explained by chance, confounding, and other biases, but uncertainties remain in the evidence overall."

Determinations are not formulated for individual outcomes but rather for the overall effect on a specific system (e.g., respiratory, cardiovascular, metabolic, nervous, and reproductive), cancer, morbidity, and mortality. In selecting individual outcomes for the current assessment, we considered the text supporting the US EPA determinations. The determination "suggestive of,

Table 1.

Causality determinations of long-term exposure according to the US EPA Integrated Science Assessments

Category	$PM_{2.5}$[22,23]	NO_2[24]	Ozone[25]
Reproductive and developmental effects	Suggestive	Suggestive	Suggestive
Respiratory effects	Likely	Likely	Likely
Cardiovascular diseases	Causal	Suggestive	Suggestive
Metabolic effects	Suggestive	Suggestive	Suggestive
Neurological effects	Likely		Suggestive
Cancer (lung cancer)	Likely[a]	Suggestive	Inadequate
Mortality	Causal	Suggestive	Suggestive

[a]IARC 2013: $PM_{2.5}$ and PM_{10} exposure is carcinogenic to humans (group 1).

but not sufficient to infer, a causal relationship" was not considered, except for diabetes, as discussed later. However, several health outcomes in this category may deserve additional attention in the future (see Discussion).

List of pollutant/outcome pairs

We selected a list of pollutant/outcome pairs based on causality assessment (Table 1). Regarding long-term exposure to air pollutants, causal or likely causal findings were determined for $PM_{2.5}$, including diseases of the respiratory, cardiovascular, and neurological systems, and lung cancer. For NO_2 and O_3, we considered respiratory diseases.

Based on the epidemiological literature, the following specific outcomes were defined a priori for long-term exposure: asthma in children and adults, acute lower respiratory infections (ALRI) in children and adults, COPD for nonmalignant respiratory diseases; ischemic heart disease (IHD) events (i.e., acute myocardial infarction), stroke, heart failure, hypertension, and atrial fibrillation for cardiovascular diseases; autism spectrum disorders (ASD), Parkinson's disease, and dementia for neurological diseases; and lung cancer.

Regarding type 2 diabetes (metabolic outcomes), the US EPA causality assessment is "suggestive" (Table 1). However, in this study, we have included diabetes because recently published studies[26,27] on mechanisms of action have increased the credibility of a causal relationship with $PM_{2.5}$.

Search the literature for available systematic reviews and meta-analyses

The available systematic reviews and meta-analyses for each pollutant/outcome pair were identified using two databases (PubMed for Medline and Web of Science) for the years 2013–2022 (closing date 30 November 2022). The basic search terms were: "$PM_{2.5}$" or "fine particle" or "NO_2" or "nitrogen dioxide" or "O_3" or "air pollution," and the specific outcome (e.g., stroke), as well as "systematic review" or "meta-analysis" in the title or abstract. No restriction was imposed on the publication language. The full search strategy is reported in the Prospero protocol (CRD42023397145). It is important to acknowledge that using PubMed we conducted a search for systematic reviews published after our closing date (30 November 2022), up to 10 February 2024. The results of the latter search were documented in eAppendix 2; http://links.lww.com/EE/A280, specifically in the outcome-specific descriptions. However, it is noted that the newer systematic reviews were not considered in our evaluation.

We also searched the reference lists of the retrieved articles for full-text reading to identify additional studies. A helpful document to check the completeness of our search was the "umbrella review,"[28] with evaluations of systematic reviews of $PM_{2.5}$ health effects with a search up to August 2021.

The steps in selecting the systematic reviews were identification, screening of title and abstract assessment, eligibility after full-text reading, and final inclusion. The results of the selection processes were reported in tables documenting the reason for an article's exclusion. We included systematic reviews and meta-analyses focused on the general population. We excluded studies on sulfur dioxide, carbon monoxide, indoor air pollution, occupational exposures, and transboundary haze/desert storms.

Systematic reviews and meta-analyses were included if they met all the following conditions:

(1) The original studies included in the meta-analysis for long-term exposure were cohort or case–control studies.
(2) The studies provided pooled and individual effect sizes of the association between long-term exposure and health outcomes.
(3) The number of subjects and cases in the study group were provided.
(4) The systematic review was not restricted to regional populations (but systematic reviews restricted to North America and Europe were allowed).

Studies were excluded if any of the following conditions were noted: (1) they were not formally published peer-reviewed journal articles, such as conference abstracts or editorials; (2) the number of included original studies for the same outcome was fewer than three; and (3) there were obvious errors in the data reported.

A two-step process was adopted in the search and study selection based on title and abstract. (1) Two researchers (I.C. and J.Z.) independently searched the databases using the keywords and saved the identified articles. A panel discussion involving two additional researchers (Z.J.A. and C.A.) clarified and eliminated any discrepancies. (2) Next, I.C. and J.Z. independently screened the titles and abstracts of the selected papers to remove those that did not meet the selection criteria, while Z.J.A. and C.A. clarified and eliminated any discrepancies. The selected papers were delivered to two other investigators (F.F. and M.K.) for full-text review. These two investigators independently read the papers and excluded irrelevant articles. The discrepancies were clarified and eliminated by discussion between these two investigators.

Appraisal of the systematic review/meta-analysis, and data extraction

We followed the approach for assessing multiple systematic reviews (AMSTAR 2) proposed by Shea et al.[29] We evaluated the characteristics of the selected meta-analyses, including the literature search, literature inclusion, data extraction, statistical analysis, and bias evaluation. We modified and augmented AMSTAR 2 tool to include additional criteria not foreseen in AMSTAR 2 but relevant for air pollution studies (see eAppendix 1; http://links.lww.com/EE/A279). The modified instrument (AMSTAR2-EH) includes 21 criteria, of which we considered five "critical" for the reliability of the systematic review (see Box 2 in eAppendix 1; http://links.lww.com/EE/A279). We propose that all "critical" criteria should be met in an SR used as a credible source of CRF for HRA. In addition, such "acceptable" SRs should not miss more than four other criteria, noting that the credibility of the SR is greater if fewer (noncritical) criteria are missed.

Two authors (F.F. and M.K.) evaluated the systematic reviews using the adapted tool separately, and the evaluation results were reported in a table designed in advance. Any discrepancies were resolved by discussion.

For each health outcome, we attempted to identify at least three systematic reviews/meta-analyses in the literature search. Next, we reviewed the most recent ones. If we did not find an acceptable study among those published in recent years, we kept searching back in time, until an acceptable SR was found or until 2013. If two or more studies published in the same year were acceptable, we selected the one with the largest number of studies in the meta-analysis as the source of recommended CRF. The AMSTAR2-EH criteria not met by each of the appraised SRs, the number of studies included in the meta-analysis, the pooled estimate of the relative risk (RR) and its 95% confidence interval (CI), heterogeneity statistic (I^2), and, optionally, comments regarding the analysis were reported in a table (see Table A1 in eAppendix 2; http://links.lww.com/EE/A280).

For each acceptable systematic review/meta-analysis, the basic data were extracted and reported in a table designed in advance. A forest plot was also extracted. Two authors performed the data extraction separately (F.F. and M.K.), and any discrepancies were resolved by discussion.

Validation of the data in the systematic reviews/ meta-analyses

Before we accepted the CRF from the selected SR, we validated the information from the individual studies included in the SR and its meta-analysis (MA). The validation process was aimed at (1) verifying the correct extraction and use in the meta-analysis of the effect estimates from the original studies, (2) checking the presence of duplicate studies or overlaps, (3) checking whether the effect measures were not a mix of incidence and prevalence data, and (4) reviewing additional information on the individual studies for the generalizability of the results. If any of the issues mentioned in points 1–3 were identified, a revised meta-analysis was performed with the updated input data to produce a revised forest plot and a pooled effect estimate (with 95% CI and I^2). A single (rotating) investigator extracted data from the individual studies that formed the systematic review. M.K. checked the results, and agreement was achieved through discussion. The new random effect meta-analyses (using the restricted maximum-likelihood random-effects model) and forest plots were produced using the command "metaan" in STATA (https://www.stata.com/) by F.F. We have not included any effect estimates from more recent studies in the revised MA because we have not conducted our own systematic search to ascertain the completeness of the most recent evidence. However, eAppendix 2; http://links.lww.com/EE/A280 contains information on studies published after the closing date of our search.

Finally, additional information was used to support the review's conclusions, namely, if any of the individual studies analyzed the shape of CRFs and provided information on whether the CRF was linear or curvilinear, and whether there was any evidence for an effect threshold. This process also allowed the collection of three important pieces of information from the single studies that contributed to the relevant systematic reviews and the meta-analyses: the age range of the population studied, the International Classification of Diseases (ICD) codes (either the 9th or the 10th revision), and the range of pollutant concentrations. The data were collected in a table and formed the basis for the suggestions for the application of the CRFs.

Confidence in the epidemiological evidence and suggestions of the concentration–response functions

To evaluate the confidence in the findings of the selected systematic reviews and meta-analyses as the sources of the recommended CRFs, we considered several aspects concerning the robustness of the results. We chose not to use the grading of recommendations, assessment, development and evaluation approach[30] as this would require an entirely new assessment of the certainty of the evidence, including an in-depth analysis of each original study included in each SR. Instead, we relied on the information presented in the selected SR papers and used the following criteria as a guide in our narrative assessment. The criteria for our assessment were:

(1) Key criteria:
   (i) A sufficient number of studies are included in the meta-analysis (e.g., five or more studies). The exact number is somehow arbitrary, but it was seen as a guide to be considered along with other characteristics of the meta-analysis, such as the number of subjects in the various studies, the geographical distribution of the studies, and the heterogeneity of the results. Hence, a meta-analysis with fewer than five studies but with many subjects (i.e., greater than 100,000 participants) from different populations/countries, large size of the effect, and low heterogeneity across studies might still be acceptable.
   (ii) Statistical significance ($P < 0.05$) of the meta-analysis result. This criterion applies to systematic reviews reporting an association between the pollutant and the outcome.

(2) Important criteria:
   (i) Studies in the MA should be distributed across various continents/countries.
   (ii) The meta-analytic weights should be distributed across the studies to avoid for only a few, say, one-third of the studies or less, contributing over two-thirds of the overall sum of weights.
   (iii) The effect estimate should be sufficiently precise, i.e., the width of the 95% CI as a proportion of the central effect estimate should be less than 100%.
   (iv) Heterogeneity (I^2) should be <75% or the lower bound of the 80% prediction interval should not include the null hypothesis (RR or hazard ratio [HR] = 1). A good systematic review is interested in heterogeneity that is not explained by known factors such as population differences, exposure assessment, study design, outliers, or other variables. Prediction interval is rarely reported and is relevant mainly for clinical trials.
   (v) Results of recently published individual studies (not included in the systematic review) should be consistent with the systematic review/meta-analysis results. Divergent results should be discussed.

Based on the consideration of the above criteria, for the pollutant/outcome pairs from outcome categories determined as "causally" or "likely to be causally" associated with the exposure, we suggested the classification of the CRF according to the following three lists:

(1) List A (a reliable quantification of health effects is possible in an HRA);
(2) List B+ (an HRA is possible, but there is greater uncertainty around the reliability of the CRF compared to pollutant/outcome pairs on list A);
(3) List B– (CRF is not recommended for use in an HRA due to the substantial uncertainty of the CRF).

For pollutant/outcome pairs with less certain causality determination (as the "suggestive" causality for diabetes), the classification was downgraded. Not included on the lists (i.e., no CRFs recommendation) were all the outcomes for which a meta-analysis was unavailable, an acceptable (according to AMSTAR2-EH appraisal) meta-analysis could not be identified, or only a few criteria related to the confidence in the evidence were met.

Note that we evaluated the epidemiological evidence included in the meta-analysis for the purposes of quantifying the effects of HRA. This is not in itself a causality assessment because, of course, no toxicological evidence was included in the evaluation, as we relied on the US EPA ISA for that. Note also that it is possible to have reasonable evidence for an association while being more cautious about quantifying that effect, e.g., due to variability in the effect size. The formulation of suggestions on CRFs for HRA followed these considerations.

Results

Selection of the systematic reviews

A PubMed and Web of Science search identified 2096 relevant records, including 664 duplicate records (Figure 1). Duplicate studies were excluded, leaving 1432 records. The number of papers removed based on the assessment of the title and the abstract was 1330 (1327 listed in File 1; http://links.lww.com/EE/A283). In addition, three papers presenting umbrella reviews on the same topic as this study were excluded.[31–33]. We added the recently published HEI traffic review report,[34] resulting in 103 papers for the full-text assessment (File 2; http://links.lww.com/EE/A284). Based on the full-text assessment, 28 articles were excluded for $PM_{2.5}$ and 94 for NO_2, leaving 75

systematic reviews addressing long-term $PM_{2.5}$ exposure and various health outcomes (note that the numbers of papers on $PM_{2.5}$ include those addressing low birth weight [14 out 103 of identified for full-text assessment, out of which seven met inclusion criteria]), while 10 addressed long-term NO_2 exposure and respiratory outcomes. The reasons for exclusions are indicated in Figure 1. The HEI traffic review (2022) dealt with traffic-related air pollutants (including NO_2 and $PM_{2.5}$ from traffic) and nonmalignant respiratory conditions (asthma, ALRI, and COPD). We used it to assess the NO_2 effects but excluded it from the evaluation of $PM_{2.5}$-related effects because we considered $PM_{2.5}$ from all sources in our review. In all, we considered 15 outcomes for $PM_{2.5}$ and 5 for NO_2. No systematic review on long-term exposure to ozone and morbidity was found, as all systematic reviews about this pollutant focus on mortality or short-term effects.

Appraisal of the systematic reviews and suggestion of concentration–response functions

The full description of the selection of the systematic reviews and meta-analyses, the results of the AMSTAR2-EH appraisal, the description of the selected systematic reviews, and the evaluation of the confidence in their results as the source of CRFs for HRA are presented in eAppendix 2; http://links.lww.com/EE/A280.

A total of 44 systematic reviews (of a total of 75) were evaluated in duplicate with the AMSTAR2-EH tool; 22 of them were found to be credible sources of CRFs for HRA (i.e., all "critical" AMSTAR2-EH criteria were met in these SRs, and they missed not more than four "other criteria"). Apart from the HEI (2022) review,[34] all five SRs considering exposure to NO_2 included studies on $PM_{2.5}$, so the AMSTAR2-EH evaluation applied to both exposures. The agreement between the two raters was excellent (crude agreement of the classification was 92.3%, $\kappa = 0.86$ [0.72–1.00]). Due to a conflict of interest (F.F. was an author of the study), the HEI traffic report was only evaluated by a single rater (M.K.), who found that only two AMSTAR2-EH "other criteria" were not met, so it was a credible source of CRFs.

Figure 2 shows noncompliance percentages with the 21 AMSTAR2-EH criteria in the appraised SRs. The most frequent critical problems were mixing effect estimates regarding incidence and prevalence or incidence and mortality in the same meta-analysis (item A1, 21% of SRs); no appropriate assessment of the risk of bias from individual studies included in the review (item 8, 18%), lack of justification for studies exclusion from the SRs (item 6, 16%), and effect estimates from individual studies included in MA not transformed into a common metric (item A4, 14%). Among the noncritical criteria, the review protocol registration was most often missed (item A9, 71% of considered SRs). Many systematic reviews ignored results from the studies not included in the MA (item A8, 43%). For each appraised SR, the criteria missed are summarized in Table A1 in eAppendix 2; http://links.lww.com/EE/A280. In addition, the table lists the number of original studies included in each SR, the value of the pooled HR and its 95% CI (and, in several cases, the revised meta-analytic values estimated in our review for the reasons mentioned in eAppendix 2; http://links.lww.com/EE/A280), and a heterogeneity indicator (I^2), as well as comments related to potential limitations of the scope of the MA or other aspects of the review.

As documented in Table A1; http://links.lww.com/EE/A280, for a given SR and health outcome, the number of studies included in the MA could be different, and often, newer SRs included fewer original studies than earlier ones. For example, the SR on lung cancer by Pyo et al[35] included six studies (all on cancer incidence), while the review by Yu et al[36] included 21 studies, of which seven were on incidence and the remaining 14 on mortality. For certain health outcomes, the scope and the inclusion criteria varied between SRs (e.g., studies on childhood asthma or ASD considered various exposure windows [prenatal or postnatal], some studies on IHD and stroke included mortality together with incidence, and one of the SRs on stroke included short-term studies). However, despite these differences in completeness, inclusion criteria, and design across SRs, for any particular pollutant–outcome pair, the CI of the pooled risk estimates clearly overlapped for the reviews of similar scope. This increases the confidence in the risk estimate recommended for the HRA according to our selected SR.

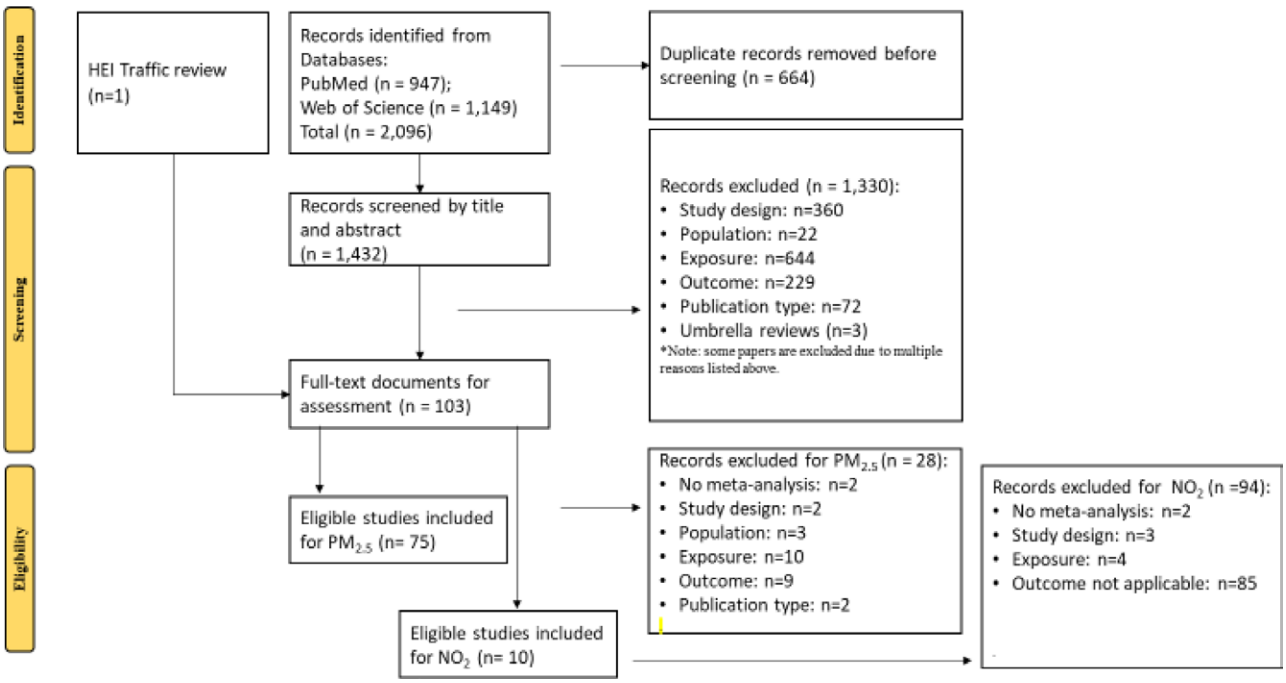

Figure 1. Selection process of the systematic reviews and meta-analyses.

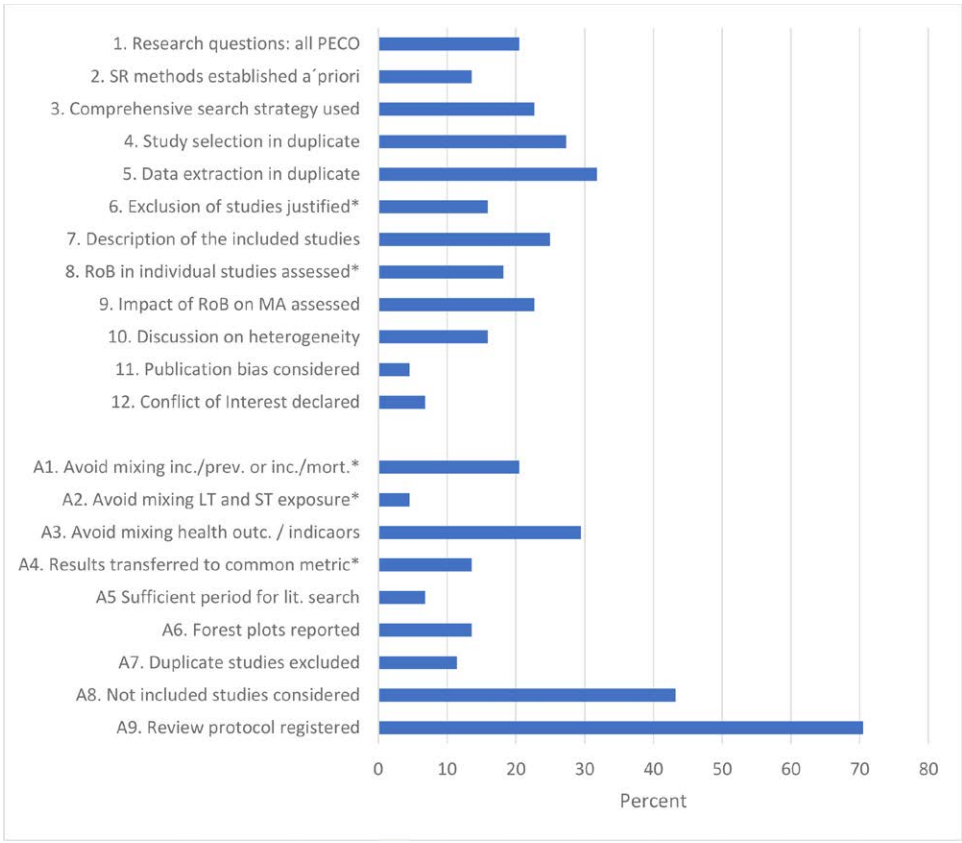

Figure 2. Percentages of the 44 systematic reviews noncompliant with the AMSTAR2-EH criteria. *Items that were considered "critical."

For five pollutant/outcome pairs, no systematic review was found (asthma and ALRI in adults for $PM_{2.5}$) or the identified systematic reviews were not a credible source for CRF according to AMSTAR2-EH appraisal (ALRI in children and heart failure for $PM_{2.5}$ and ALRI in adults for NO_2).

Following the evaluation of the confidence in the findings of the remaining selected SRs as the sources of CRFs for HRA, we classified six outcomes related to $PM_{2.5}$ in list A (asthma in children, COPD, IHD events, stroke, hypertension, and lung cancer) and three outcomes (diabetes, dementia, and ASDs) in list B+ (Table 2). For two outcomes (atrial fibrillation and Parkinson's disease), the evidence was considered insufficient for risk assessment, and the outcomes were assigned to list B–. For NO_2, three outcomes were placed in list A (asthma in children and adults, ALRI in children), and one (COPD) was assigned to list B–.

Most (8 of the 12) effect estimates allocated to lists A or B+ were based on SRs published within 2 years before the search closing date (i.e., in 2021 or 2022). The oldest review[37] focused on asthma in children from postnatal exposure. The other recent review[45] on childhood asthma included studies only on maternal exposure during pregnancy, with the effect estimate for the entire pregnancy close to the null, and we have decided not to recommend it for HRA.

We verified the inputs to the meta-analyses based on the individual papers included in the selected systematic reviews. We modified these inputs in many cases (6 of 12 systematic reviews), and then revised the meta-analysis using the correct input data. The details of these revisions are reported in eAppendix 2; http://links.lww.com/EE/A280.

For the pollution/outcome pairs classified in lists A and B+, we have collected basic information on the studies included in the systematic reviews (Table SR; http://links.lww.com/EE/A281). If a revised meta-analysis was necessary, we also included the new number of studies along with the revised pooled effect estimate and 95% CI, as well as the revised heterogeneity test (I^2).

Suggestions for concentration–response functions and application in health risk assessment

Table 2 summarizes the main results of the review of systematic reviews described in detail in eAppendix 2; http://links.lww.com/EE/A280 and Table SR; http://links.lww.com/EE/A281, including the RRs to be applied in an HRA, the ICD-10 codes, the age range of the susceptible population, and the applicable concentration range for the recommended RRs.

The precision of the pooled risk estimates produced by the selected studies was good (the width of the 95% CI as a proportion of the central effect estimate was less than 100%) for associations of COPD, stroke, and lung cancer with $PM_{2.5}$. The remaining estimates were less precise, with the largest uncertainty (the width of the 95% CI twice as large as the central effect estimate) noted for the association of asthma incidence with NO_2 exposure in adults.

Besides the reliability of the selected SR as the source of CRFs, established by the AMSTAR2-EH appraisal, the confidence in the recommended risk coefficient was based on the number of studies (all the selected CRFs were based on meta-analyses that included between 5 and 21 individual studies). For most outcomes, the number of subjects included in the cohorts considered in the MAs was 1.1–5.7 million (asthma in children and adults, COPD, IHD, stroke, diabetes, and lung cancer). More than 40.7 million participants were included in the MA of dementia. The smallest number of subjects (107,000 in 11 studies) was included in the studies on the incidence of ALRI and NO_2 exposure (complete details may be found in Table SR; http://links.lww.com/EE/A281).

Table 2.

Relative risk estimates for incidence of diseases from selected systematic reviews recommended for health risk assessment of $PM_{2.5}$ and NO_2

Outcome (incidence)	ICD-10 codes	Age (yrs)	List	RR (95% CI) per 10 μg/m³	Mean exposure range (μg/m³)	SR reference
Long-term exposure to $PM_{2.5}$						
Asthma in children	J45	0–18	A	1.34 (1.10, 1.63)	5–38	Khreis et al[37]
COPD	J41–J44	30+	A	1.18 (1.13, 1.23)	5–26	Park et al[38]
IHD events[a]	I21–I22	30+	A	1.13 (1.05, 1.22)[b]	5–65	Zhu et al[39]
Stroke	I60–I64	30+	A	1.16 (1.12, 1.20)[b]	5–36	Yuan et al[40]
Hypertension	I10–I11	30+	A	1.17 (1.05, 1.30)[b]	5–77	Qin et al[41]
Diabetes (type 2)	E11–E14	30+	B+	1.10 (1.03, 1.18)[b]	5–79	Yang et al[42]
Dementia	F00–F03, G30	60+	B+	1.46 (1.20, 1.78)[b]	5–25[c]	Cheng et al[43]
ASD	F84.0, F84.1, F84.5, F84.8, F84.9	2–12	B+	1.66 (1.23, 2.25)[b]	5–30[c]	Lin et al[44]
Lung cancer	C34	30+	A	1.16 (1.10, 1.23)	5–44	Yu et al[36]
Long-term exposure to NO_2						
Asthma in children	J45	0–18	A	1.10 (1.05, 1.18)	10–39	Khreis et al[37]
Asthma in adults	J45	19+	A	1.10 (1.01, 1.21)	10–40	HEI[34]
ALRI in children	J12–J18, J20–J22	0–12	A	1.09 (1.03, 1.16)	10–56	HEI[34]

[a]Acute myocardial infarction (AMI).

[b]Relative risk estimates from revised meta-analysis.

[c]Restrict applicability of the CRFs of these conditions to exposure differences not larger than 10 μg/m³ within the indicated concentration ranges (see discussion in eAppendix 2; http://links.lww.com/EE/A280).

For some outcomes (diabetes, asthma in adults, and ALRI), the recommended outcome definition (ICD-10 codes) may slightly differ from those considered in some epidemiological studies providing CRFs. We recommended ICD codes based on the knowledge of coding practices of the outcomes with a focus on the codes most often used in studies. Age ranges, which may differ from some of the source studies, are based on epidemiological knowledge of the age distribution of disease incidence and are aimed at standardizing age groups across outcomes in a HRA.

The range of mean concentrations in source studies indicates the range of mean exposures for which the uncertainty of the risk assessment is minimized. Note that for the sake of simplicity, the lowest mean concentration for each pollutant was set to be equal to the 2021 WHO AQG level (5 μg/m³ for $PM_{2.5}$ and 10 μg/m³ for NO_2). The original concentration ranges can be found in Table SR; http://links.lww.com/EE/A281. In most cases, the lowest concentration values reported in the original studies closely approximated the AQG levels, and a single value for all the outcomes was provided to simplify the HRA. On the other hand, considerable variability exists across all outcomes for the highest concentration values, so the recommendation of a single upper-bound concentration is not possible. If an HRA extends to exposures outside our recommended range, the HRA analyst should be aware of the potential for greater uncertainty in the results. A conservative approach would entail conducting a sensitivity analysis for the relevant outcome, assuming no further increase of the RR above the upper bound of the concentration range as documented in the epidemiological studies from which the upper bound of the mean $PM_{2.5}$ is obtained.

A special problem is applicability in HRA of the two largest CRFs that we found, i.e., the associations of $PM_{2.5}$ with dementia (RR = 1.46; 95% CI = 1.20, 1.78, list B+) and ASD (RR = 1.66; 95% CI= 1.23, 2.25, list B+), which have the highest RRs of all the outcomes we considered in this work. Such high RR values could produce unrealistically high ambient air pollution attributable fractions in both HRA and burden assessments. When we checked the original studies in the systematic reviews for dementia and ASD, we observed that the $PM_{2.5}$ contrast intervals for both diseases did not exceed 10 μg/m³. Therefore, we propose to apply these RRs within the applicable concentration range specified for each health outcome (Table 2) but for changes in $PM_{2.5}$ concentration less than 10 μg/m³.

Except for one SR on ASD,[46] the reviews do not provide evidence on the shape of the CRFs, and the relevant information available in some of the individual studies is inconsistent (see Table SR; http://links.lww.com/EE/A281). Therefore, a linear increase in risk can be assumed across the range of the observed concentrations.

Discussion

We used a systematic approach to determine the appropriate CRFs for the incidence of several diseases to be applied in an HRA of air pollution. In our final assessment, reliable quantification of health effects is possible (list A) for six outcomes due to long-term exposure to $PM_{2.5}$ (asthma in children, COPD, IHD events, stroke, hypertension, and lung cancer) and three outcomes for NO_2 (asthma in children and adults, ALRI in children). For three additional outcomes for $PM_{2.5}$ (diabetes, dementia, and ASD), HRA is also possible (list B+), but there is more uncertainty about the reliability of the CRF than pollutant/outcome pairs included in list A. The results can be used worldwide to cover a wide range of practical applications. By comprehensively evaluating the available literature and developing tailored advice, we ensure that the selection of CRFs for morbidity outcomes in HRA is based on sound scientific evidence and relevant study characteristics. This approach enhances the accuracy and reliability of HRA findings and contributes to informed decision-making in public health and environmental policy. Furthermore, as future empirical evidence becomes available, our methodological approach can be easily applied to update the recommended CRFs.

Several aspects regarding the methods and the results should be discussed in view of the possible limitations of our work and the need for further refinement.

We considered a causality assessment regarding the association between long-term exposure and a health outcome to be a fundamental prerequisite for the selection of the CRF. We relied on the causality assessments made using the US EPA ISA approach. In fact, to establish a causal relationship between any pollutant and a particular health effect, it is crucial to rely on robust epidemiologic studies supported by various other aspects of research. This includes providing convincing evidence on the relationship between exposure to a certain pollutant and the occurrence of adverse health outcomes in vulnerable human populations. Moreover, toxicological and clinical studies constitute a vital part of understanding the underlying mechanisms and biological pathways through which pollutants may influence human health. By leveraging collective expertise, the US

EPA ISAs comprehensively evaluate the available evidence and provide a confident assessment of the cause–effect link between a pollutant and specific health outcomes.

However, relying only on a few organizations capable of executing a fully integrated assessment (such as US EPA or IARC) may entail a significant delay before the assessment is completed, peer-reviewed, and published. For example, the evaluation of NO_2 is already more than 7 years old.[24] The next evaluation has just started, but it will take some time before an updated assessment will be available. To partially overcome the long time passed from the most recent comprehensive assessment, we have decided to consider an additional outcome related to $PM_{2.5}$ (diabetes), which was considered "suggestive" in the US EPA classification (Table 1). We also considered ASD, for which the evidence was poorer than for other neurological conditions in the US EPA evaluation. In the result of our assessment of epidemiological evidence, both diabetes and autism were included on list B+, together with dementia. Emerging research on further outcomes related to long-term exposure to $PM_{2.5}$ might have been considered (e.g., depression[47]), but a formal causality assessment is lacking.

Since the release of the US EPA causality assessment for NO_2 in 2016, evidence of additional outcomes, other than respiratory effects, has been growing (e.g., myocardial infarction, stroke, and diabetes outcomes have been implicated).[14,15] Although the case for NO_2 as a causal agent contributing to both mortality and morbidity is constantly evolving,[48] at this time, a cautious approach is warranted until further research can establish a clear causal determination between NO_2 exposure and nonrespiratory outcomes. In fact, a rigorous assessment of causality, considering all sources of evidence, is one of the guiding principles of this work. We recognize that this approach represents a conservative point of view, and consequently, the health burden of NO_2 may likely be underestimated.

According to our review, the evidence on the incidence of asthma in children is the only outcome related both to $PM_{2.5}$ and NO_2 exposure; for both pollutants, the CRFs are classified in list A. However, neither the systematic review, which is the source of the recommended CRF,[37] nor the two more recent studies[49,50] analyzing the relation of asthma in children with both pollutants provide clear insight to formulate expert advice on how to handle copollutant exposures in an HRA (see eAppendix 2; http://links.lww.com/EE/A280 for details). Since both $PM_{2.5}$ and NO_2 share common sources, the emissions are often highly correlated. Under these circumstances, it would be prudent not to combine the effects of both pollutants, as this would likely lead to a double counting of health impacts. Instead, we suggest choosing the largest impact of the individual pollutants and interpreting the result as the combined effect of exposure to both pollutants. In reference to the systematic review on $PM_{2.5}$ and asthma in children, it is important to highlight that we relied on a relatively dated paper.[37] Consequently, a new evaluation of the evidence, incorporating more recent studies, is deemed necessary.

A significant number of the evaluated systematic reviews and meta-analyses were found to miss several AMSTAR2-EH criteria, which reduces their reliability as a source of CRFs for HRA and highlights the need for improved methodological rigor in future SRs. Improvements for systematic reviews and meta-analyses should include clearer and more comprehensive search strategies, rigorous study selection and effect estimate criteria (a thorough assessment of the risk of bias in the included studies), transparent reporting of data extraction and synthesis methods, appropriate statistical analysis techniques, and consideration of sources of heterogeneity.

For air pollution studies, several aspects were important determinants affecting the applicability of SR results in HRA, such as combining effect estimates on incidence and prevalence in the same meta-analysis, combining results from long- and short-term exposure, lack of transformation of the effect estimates into a common metric, extraction of the wrong effect estimate, or the lack of consideration of the potentially useful information from the excluded studies (e.g., studies may be excluded as the effect estimates are expressed according to categories of exposure rather than according to continuous exposure, but the results may be still relevant).[51] The issue of the quality of the systematic reviews has to be given greater attention in future reviews. In our validation effort, we found mistakes in data included in the MAs, in which case the meta-analysis had to be revised.

One important limitation of our approach is the inability to assess how the risk of bias evaluation was conducted in the single SRs. It is not sufficient to know the name of the tool or the domain of the biases investigated to learn about the quality of the work done without repeating the assessment of individual studies. This quality aspect is something difficult to capture, and the final responsibility for its reliability is with the authors of the (peer-reviewed and published) SR. On the other hand, the responsibility to conduct a proper risk of bias evaluation that considers the most important biases, as well as the direction and magnitude of these biases, remains for the future new SRs.

Past reviews of SRs and MAs have focused on statistical criteria to evaluate the strength of the evidence (e.g., *P* value, number of subjects, low heterogeneity [I^2 < 50%] or 95% prediction interval that excluded the null value, excess significance, and no evidence of small-study effects).[28,52] In this work, we preferred to rely on more stringent epidemiological criteria and the consistency of the findings rather than relying on strict statistical approaches. Issues dealing with the internal validity of the studies (e.g., risk of bias) and publication bias were addressed in the AMSTAR2-EH evaluation.

Systematic appraisal of several SRs for the same outcome allowed us to compare the results across meta-analyses (see Table A1 in eAppendix 2; http://links.lww.com/EE/A280). For some outcomes, the comparison was made difficult (or impossible) because of differences in the exposure timing considered in the various SRs (e.g., asthma in children, ASD) or mixing of various health indicators in one MA (e.g., mortality and incidence in some of the SRs of IHD). For several SRs, the exposure increment was not standardized to the same unit, precluding a comparison of meta-analytic effects across SRs. There was good agreement across studies when none of the issues mentioned above were encountered, such as is the case for the MAs considering stroke, diabetes, and lung cancer.

During the peer review process of this paper, a suggestion was made to consider an alternative strategy to the one we have proposed. Instead of relying solely on the AMSTAR2-EH evaluation to select an SR, it was recommended that all available SRs be utilized as the source of individual studies. This alternative approach aims to enhance the sensitivity of the evidence search. We have tested the effectiveness of such an approach for two of the outcomes (incidence of IHD events and dementia) related to long-term $PM_{2.5}$ exposure. The results are presented in eAppendix 4; http://links.lww.com/EE/A282. The rationale for selecting these specific outcomes is the following: multiple SRs were available for IHD events, each with a varying number of studies in its assessment, prompting concerns about the completeness of the searches; dementia exhibited different numbers of studies and diverse, often high, effect estimates across various SRs.

The results of this additional analysis indicate for IHD that the MA based on the combined 21 relevant papers included in the five SRs provides a pooled effect estimate (HR = 1.12 [1.07, 1.18] per 10 $\mu g/m^3$ $PM_{2.5}$), which is almost identical to that based on the seven effect estimates included in the revised Zhu et al[39] MA (HR = 1.13 [1.05, 1.22]), but with narrower 95% CI. The same was observed for dementia: the new MA (including 16 studies) had a slightly lower effect estimate (HR = 1.43 [1.20, 1.71] compared to the revised Cheng et al[43] MA (15 studies, HR = 1.46 [1.20, 1.78]), with a similar precision. The examples suggest that the combined analysis of all available (recent) SRs does not result in significantly different or more precise pooled

HR estimates than those from the SRs considered to be reliable based on the AMSTAR2-EH evaluation. The extra workload associated with conducting a detailed analysis of papers included in all SRs is substantial and exceptionally time-consuming. While it undoubtedly contributes to a more thorough assessment of available evidence from diverse SRs, it may not be deemed essential if the goal is to ensure the accuracy and reliability of the CRF to be recommended for HRA. However, the critical decision point lies in determining when our approach (select an SR + evaluation) is acceptable, when the alternative approach is preferable due to uncertainties in the enumeration and selection of studies, or when a completely new SR is deemed necessary.

Our results should be considered as an update to the evidence provided by HRAPIE[5,6] regarding long-term effects on morbidity. Our approach differs from the methods used in three recent reports: the Swedish EPA,[12] the impact assessment study to support the revision of the EU AAQD,[13] and the EEA.[14] The main differences are related to the fact that we relied on the causality determination of the association between the pollutant and the outcome, we selected CRFs based on a systematic appraisal of systematic reviews and did not depend on single studies (such as ELAPSE), and we performed a quality check of the systematic reviews, corrected any mistakes we found and updated the results of the meta-analysis. As a result, most of the CRFs for $PM_{2.5}$ differ from those used in these other analyses. Furthermore, we did not consider stroke, diabetes, and lung cancer as suitable outcomes for NO_2 because of the lack of a comprehensive causality assessment.

Table 2 provides CRFs, ICD codes, age range, and exposure levels for potential applications in risk assessment and burden calculation. The applications of the suggested CRFs should follow the general principles discussed in WHO.[53] Furthermore, caution should be exercised when applying the dementia and ASD CRFs to large changes in concentration that exceed those reported in the original epidemiological studies from which the CRFs were extracted, as these would produce extremely high population-attributable fractions.

In conclusion, the suggestions provided in this study can contribute to conducting reliable HRAs, allowing for evidence-based decision-making in public health and environmental policy. Future research must continue updating and refining these suggestions as new evidence becomes available and methodologies evolve to understand the relationship between air pollution and morbidity outcomes.

Conflicts of interest statement

H.W.'s post is partly funded by the National Institute for Health and Care Research (NIHR) Health Protection Research Unit in Environmental Exposures and Health, a partnership between the UK Health Security Agency and Imperial College London. The other authors have no conflicts to report.

Acknowledgments

This work was supported by the Climate and Clean Air Coalition through the grant provided to the World Health Organization for the "Estimation of Morbidity from Air Pollution and its Economic Costs" (EMAPEC) project. Grant number: Climate and Clean Air Coalition, award number 73943 (August 2022).

References

1. Chen J, Hoek G. Long-term exposure to PM and all-cause and cause-specific mortality: a systematic review and meta-analysis. *Environ Int*. 2020;143:105974.
2. Huangfu P, Atkinson R. Long-term exposure to NO2 and O3 and all-cause and respiratory mortality: a systematic review and meta-analysis. *Environ Int*. 2020;144:105998.
3. Hoffmann B, Brunekreef B, Andersen ZJ, Forastiere F, Boogaard H. Benefits of future clean air policies in Europe: proposed analyses of the mortality impacts of PM2.5 and NO2. *Environ Epidemiol*. 2022;6:e221.
4. Dechezleprêtre A, Rivers N, Stadler B. *The Economic Cost of Air Pollution: Evidence From Europe*. Economics Department Working Papers no.1584.OECD; 2020.
5. WHO Regional Office for Europe. *Health Risks of Air Pollution in Europe – HRAPIE Project: Recommendations for Concentration-response Functions for Cost–benefit Analysis of Particulate Matter, Ozone and Nitrogen Dioxide*. 2013. Available at: https://www.who.int/europe/publications/i/item/WHO-EURO-2013-6696-46462-67326. Accessed 31 May 2024.
6. Héroux ME, Anderson HR, Atkinson R, et al. Quantifying the health impacts of ambient air pollutants: recommendations of a WHO/Europe project. *Int J Public Health*. 2015;60:619–627.
7. WHO Regional Office for Europe. *Review of Evidence on Health Aspects of Air Pollution – REVIHAAP Project: Technical Report*. 2013. Available at: https://www.who.int/europe/publications/i/item/WHO-EURO-2013-4101-43860-61757. Accessed 31 May 2024.
8. Public Health England. *Estimation of Costs to the NHS and Social Care Due to the Health Impacts of Air Pollution*. PHE Publications; 2018. Available at: https://www.gov.uk/government/publications/air-pollution-a-tool-to-estimate-healthcare-costs. Accessed 31 May 2024.
9. Committee on the Medical Effects of Air Pollutants (COMEAP). *Summary of Recommendations for Quantifying the Health Effects of Air Pollutants:2023*. PHE publications; 2023. Available at: https://www.gov.uk/government/publications/air-pollutants-quantification-of-associated-health-effects. Accessed 31 May 2024.
10. Committee on the Medical Effects of Air Pollutants (COMEAP). *Statement on quantifying mortality associated with long-term exposure to $PM_{2.5}$*. UK Health Security Agency; 2022. Available at: https://www.gov.uk/government/publications/particulate-air-pollution-quantifying-effects-on-mortality. Accessed 31 May 2024.
11. Committee on the Medical Effects of Air Pollutants (COMEAP). *Statement on update of recommendations for quantifying hospital admissions associated with short-term exposures to air pollutants*. UK Health Security Agency; 2022. Available at: https://www.gov.uk/government/publications/air-pollution-quantifying-effects-on-hospital-admissions. Accessed 31 May 2024.
12. Dajnak D, Kitwiroon N, Assareh N, et al. *Pathway to WHO: Achieving Clean Air in the UK. Modelling Air Quality Costs and Benefits*. Environmental Reasearch Group Imperial College London; 2022. Available at: https://www.imperial.ac.uk/school-public-health/environmental-research-group/research/modelling/pathway-to-who/. Accessed 31 May 2024.
13. Oudin A, Flanagan E, Malmqvist E. *Evidence of Effect and Exposure-response Functions for PM2.5 and NO2 Linked to Morbidity*. Umea University; 2022. Available at: http://umu.diva-portal.org/smash/get/diva2:1658046/FULLTEXT01. Accessed 31 May 2024.
14. *European Commission, Directorate-General for Environment, Study to support the impact assessment for a revision of the EU Ambient Air Quality Directives: Final Report: Appendix*. Publications Office of the European Union; 2022. Available at: https://data.europa.eu/doi/10.2779/49843. Accessed 31 May 2024.
15. Kienzler S., Soares J, González Ortiz A, et al. *Estimating the Morbidity Related Environmental Burden of Disease Due to Exposure to PM2.5, NO2 and O3 in Outdoor Ambient Air (Eionet Report – ETC HE 2022/11)*. European Topic Centre on Human Health and the Environment; 2022. Available at: https://www.eionet.europa.eu/etcs/etc-he/products/etc-he-products/etc-he-reports/etc-he-report-2022-11-estimating-the-morbidity-related-environmental-burden-of-disease-due-to-exposure-to-pm2-5-no2-and-o3-in-outdoor-ambient-air. Accessed 31 May 2024.
16. Künzli N, Medina S, Kaiser R, Quénel P, Horak F, Studnicka M. Assessment of deaths attributable to air pollution: should we use risk estimates based on time series or on cohort studies? *Am J Epidemiol*. 2001;153:1050–1055.
17. Committee on the Medical Effects of Air Pollutants (COMEAP). *The Mortality Effects of Long-term Exposure to Particulate Air Pollution in the United Kingdom. A Report by the Committee on the Medical Effects of Air Pollutants*. PHE Publications; 2010. Available at: https://www.gov.uk/government/publications/comeap-mortality-effects-of-long-term-exposure-to-particulate-air-pollution-in-the-uk. Accessed 31 May 2024.
18. Menon JML, Struijs F, Whaley P. The methodological rigour of systematic reviews in environmental health. *Crit Rev Toxicol*. 2022;52:167–187.

19. Belbasis L, Bellou V, Ioannidis JPA. Conducting umbrella reviews. *BMJ Med*. 2022;1:e000071.
20. World Health Organization. *Air Quality Guidelines - Update 2021*. WHO Regional Office for Europe; 2021.
21. European Environmental Agency. *Report 5/2022. Air quality in Europe 2022*. Available at:https://www.eea.europa.eu/publications/status-of-air-quality-in-Europe-2022. Accessed 31 May 2024.
22. US EPA. *Integrated Science Assessment for Particulate Matter (Final Report, 2019)*. EPA/600/R-19/188. 2019. Available at: https://www.epa.gov/isa/integrated-science-assessment-isa-particulate-matter. Accessed 31 May 2024.
23. US EPA. *Supplement to the 2019 Integrated Science Assessment for Particulate Matter (Final Report, 2022))*. 2022. Available at: https://www.epa.gov/isa/integrated-science-assessment-isa-particulate-matter. Accessed 31 May 2024.
24. US EPA. *Integrated Science Assessment (ISA) for Oxides of Nitrogen – Health Criteria (Final Report, January 2016)*. 2016. Available at: https://www.epa.gov/isa/integrated-science-assessment-isa-oxides-nitrogen-health-criteria. Accessed 31 May 2024.
25. U.S. EPA. *Integrated Science Assessment (ISA) for Ozone and Related Photochemical Oxidants (Final Report, Apr 2020)*. EPA/600/R-20/012. 2020. Available at: https://cfpub.epa.gov/ncea/isa/recordisplay.cfm?deid=348522. Accessed 31 May 2024.
26. Zhao L, Fang J, Tang S, et al. PM2.5 and serum metabolome and insulin resistance, potential mediation by the gut microbiome: a population-based panel study of older adults in China. *Environ Health Perspect*. 2022;130:27007.
27. Liang D, Li Z, Vlaanderen J, et al. A state-of-the-science review on high-resolution metabolomics application in air pollution health research: current progress, analytical challenges, and recommendations for future direction. *Environ Health Perspect*. 2023;131:56002.
28. Zang ST, Wu QJ, Li XY, et al. Long-term PM2.5 exposure and various health outcomes: an umbrella review of systematic reviews and meta-analyses of observational studies. *Sci Total Environ*. 2022;812:152381.
29. Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ*. 2017;358:j4008.
30. Morgan RL, Thayer KA, Bero L, et al. GRADE: assessing the quality of evidence in environmental and occupational health. *Environ Int*. 2016;92-93:611–616.
31. de Bont J, Jaganathan S, Dahlquist M, Persson A, Stafoggia M, Ljungman P. Ambient air pollution and cardiovascular diseases: an umbrella review of systematic reviews and meta-analyses. *J Intern Med*. 2022;291:779–800.
32. Zang ST, Wu QJ, Li XY, et al. Long-term PM2.5 exposure and various health outcomes: an umbrella review of systematic reviews and meta-analyses of observational studies. *Sci Total Environ*. 2022;812:152381.
33. Markozannes G, Pantavou K, Rizos EC, et al. Outdoor air quality and human health: an overview of reviews of observational studies. *Environ Pollut*. 2022;306:119309.
34. HEI (Health Effects Institute) Panel. *Systematic Review and Meta-analysis of Selected Health Effects of Long-term Exposure to Traffic-related Air Pollution*. Special Report 23. 2022. Available at: https://www.healtheffects.org/publication/systematic-review-and-meta-analysis-selected-health-effects-long-term-exposure-traffic. Accessed 31 May 2024.
35. Pyo JS, Kim NY, Kang DW. Impacts of outdoor particulate matter exposure on the incidence of lung cancer and mortality. *Medicina (Kaunas)*. 2022;58:1159.
36. Yu P, Guo S, Xu R, et al. Cohort studies of long-term exposure to outdoor particulate matter and risks of cancer: a systematic review and meta-analysis. *Innovation (Camb)*. 2021;2:100143.
37. Khreis H, Kelly C, Tate J, Parslow R, Lucas K, Nieuwenhuijsen M. Exposure to traffic-related air pollution and risk of development of childhood asthma: a systematic review and meta-analysis. *Environ Int*. 2017;100:1–31.
38. Park J, Kim HJ, Lee CH, Lee CH, Lee HW. Impact of long-term exposure to ambient air pollution on the incidence of chronic obstructive pulmonary disease: a systematic review and meta-analysis. *Environ Res*. 2021;194:110703.
39. Zhu W, Cai J, Hu Y, et al. Long-term exposure to fine particulate matter relates with incident myocardial infarction (MI) risks and post-MI mortality: a meta-analysis. *Chemosphere*. 2021;267:128903.
40. Yuan S, Wang J, Jiang Q, et al. Long-term exposure to PM2.5 and stroke: a systematic review and meta-analysis of cohort studies. *Environ Res*. 2019;177:108587.
41. Qin P, Luo X, Zeng Y, et al. Long-term association of ambient air pollution and hypertension in adults and in children: a systematic review and meta-analysis. *Sci Total Environ*. 2021;796:148620.
42. Yang BY, Fan S, Thiering E, et al. Ambient air pollution and diabetes: a systematic review and meta-analysis. *Environ Res*. 2020;180:108817.
43. Cheng S, Jin Y, Dou Y, et al. Long-term particulate matter 2.5 exposure and dementia: a systematic review and meta-analysis. *Public Health*. 2022;212:33–41.
44. Lin LZ, Zhan XL, Jin CY, Liang J-H, Jing J, Dong G-H. The epidemiological evidence linking exposure to ambient particulate matter with neurodevelopmental disorders: a systematic review and meta-analysis. *Environ Res*. 2022;209:112876.
45. Hua L, Ju L, Xu H, et al. Outdoor air pollution exposure and the risk of asthma and wheezing in the offspring. *Environ Sci Pollut Res Int*. 2023;30:14165–14189.
46. Lin CK, Chang Y-T, Lee F-S, Chen S-T, Christiani D. Association between exposure to ambient particulate matters and risks of autism spectrum disorder in children: a systematic review and exposure-response meta-analysis. *Environ Res Lett*. 0630;16:063003.
47. Borroni E, Pesatori AC, Bollati V, Buoli M, Carugno M. Air pollution exposure and depression: a comprehensive updated systematic review and meta-analysis. *Environ Pollut*. 2022;292(Pt A):118245–118245.
48. Forastiere F, Peters A. Invited perspective: the NO2 and mortality dilemma solved? Almost there! *Environ Health Perspect*. 2021;129:121304.
49. Gehring U, Wijga AH, Koppelman GH, Vonk JM, Smit HA, Brunekreef B. Air pollution and the development of asthma from birth until young adulthood. *Eur Respir J*. 2020;56:2000147.
50. Tingskov Pedersen CE, Eliasen AU, Ketzel M, et al. Prenatal exposure to ambient air pollution is associated with early life immune perturbations. *J Allergy Clin Immunol*. 2023;151:212–221.
51. Savitz DA, Forastiere F. Do pooled estimates from meta-analyses of observational epidemiology studies contribute to causal inference? *Occup Environ Med*. 2021;78:621–622.
52. Belbasis L, Bellou V, Evangelou E, Ioannidis JPA, Tzoulaki I. Environmental risk factors and multiple sclerosis: an umbrella review of systematic reviews and meta-analyses. *Lancet Neurol*. 2015;14:263–273.
53. WHO Regional Office for Europe. *Health Risk Assessment of Air Pollution – General Principles*. 2016. Available at: https://www.who.int/publications/i/item/9789289051316. Accessed 31 May 2024.