Comments on EPA's "Integrated Review Plan for the Review of the Ozone National Ambient Air Quality Standards, External Review Draft"

Docket ID: EPA-HQ-OAR-2018-0279

Julie E. Goodman, Ph.D., DABT, FACE, ATS
Gradient

November 21, 2018

GRADIENT

www.gradientcorp.com
20 University Road
Cambridge, MA 02138
617-395-5000

Table of Contents

Page

Executive Summary ES-1

1 Introduction 1

2 Science Assessment 2
- 2.1 Scope of the ISA (4.3.2) 3
- 2.2 Literature Search and Selection of Relevant Studies (4.3.3) 3
- 2.3 Discipline-specific Scoping, Searching, and Screening (4.3.4) 4
- 2.4 Identification of Policy-relevant Studies (4.3.5) 4
- 2.5 Evaluation of Individual Study Quality (4.3.6) 5
- 2.6 Integration of Evidence and Determination of Causality (4.3.7) 6
- 2.7 Causality Determination from the 2013 ISA (4.4.1) 9
- 2.8 Health Effects (4.4.4) 9
- 2.9 At-Risk Life Stages and Populations and Public Health Impact (4.4.5) 9

3 Quantitative Risk and Exposure Assessments 11

4 Recommendation for CASAC 15

References 17

Abbreviations

APEX	Air Pollutants Exposure Model
CASAC	Clean Air Scientific Advisory Committee
EFSA	European Food Safety Authority
FEV_1	Forced Expiratory Volume in One Second
IOM	Institute of Medicine
IRP	*Integrated Review Plan for the Ozone National Ambient Air Quality (External Review Draft)*
ISA	Integrated Science Assessment
MoA	Mode of Action
NAAQS	National Ambient Air Quality Standards
NTP	National Toxicology Program
OHAT	Office of Health Assessment and Translation
PECOS	Populations, Exposures, Comparators, Outcomes, Study Designs
REA	Health Risk and Exposure Assessment
SciRAP	Science in Risk Assessment and Policy
TCEQ	Texas Commission on Environmental Quality
TSCA	Toxic Substances Control Act
US EPA	United States Environmental Protection Agency
WoE	Weight-of-Evidence

Executive Summary

EPA released the *Integrated Review Plan for the Ozone National Ambient Air Quality (External Review Draft)* (herein referred to as the "draft IRP") in October 2018. The draft IRP reviews key findings from EPA's prior ozone assessment and outlines its approach for the forthcoming Integrated Science Assessment (ISA) and Health Risk and Exposure Assessment (REA). The draft IRP also refers to the Preamble to the ISAs, which outlines the general process used by EPA for developing the ISAs, including the overall framework for evaluating evidence and making causal conclusions (US EPA, 2015).

The draft IRP has a useful summary of several important overarching principles; however, the draft IRP and the Preamble do not provide sufficiently specific instructions for all phases of the analysis. As they currently stand, these documents do not have sufficient detail to ensure that studies are identified and reviewed in a systematic and consistent manner or integrated in a way that considers study quality and the coherence of results across studies, both within and across disciplines. Further, the current causal framework is biased in the direction of affirmative causal determinations.

The draft IRP should be revised to include the following:

- A detailed, objective, and transparent systematic review protocol for the ISA that includes details for each step of the review process;
- A more objective and robust causality evaluation system, ideally the Institute of Medicine (IOM) framework upon which the existing framework is based;
- Clear study selection criteria and a thorough, well-described study quality evaluation system;
- An *a priori* methodology for evidence integration that includes a detailed process for integrating evidence included in the 2013 ISA with new evidence, including a consideration of new evidence that may impact the previous causal determinations; and
- Methods for evaluating at-risk factors (*i.e.*, effect modifiers, such as pre-existing disease) with the same rigor as those for evaluating studies of ozone health effects.

With respect to the potential forthcoming REA, the draft IRP provides very little detail. It should:

- Explain how it will determine whether new information, tools, or methods address uncertainties discussed in the last review.
- Better characterize the uncertainty surrounding controlled exposure studies, and all aspects of uncertainty should be quantitatively addressed in risk estimates.
- Indicate that it will consider a threshold concentration-response function when modeling the risk of ozone-induced health effects, if data allow.

The Clean Air Scientific Advisory Committee (CASAC) should recommend that EPA provide a more detailed, objective, and transparent methodology in the ozone IRP, as laid out in the bullets above. This will limit potential biases in the ISA and REA, help ensure that the analyses can be reproduced by others, and increase confidence in the causal determinations regarding ozone-associated effects.

1 Introduction

EPA released the *Integrated Review Plan for the Ozone National Ambient Air Quality (External Review Draft)* (herein referred to as the "draft IRP") in October 2018 (US EPA, 2018). The draft IRP reviews key findings from EPA's prior assessment, outlines its approach for the forthcoming Integrated Science Assessment (ISA) and Health Risk and Exposure Assessment (REA), and briefly discusses plans for the Policy Assessment and Proposed and Final Decisions. The draft IRP has a useful summary of several important overarching principles; however, as discussed in more detail below, it should provide a more explicit methodology for the ISA to ensure that it is conducted in a robust, transparent, and reproducible manner. It should also better describe how EPA plans to evaluate and resolve any discrepancies in evidence published before *versus* after the previous ISA review. In addition, the IRP should provide more detailed information on specific approaches and methods for a potential REA.

I have prepared these comments on the draft IRP with funding from the American Petroleum Institute; however, the opinions expressed are my own. The sections that follow are arranged in the order of discussion in the draft IRP. Section 2 includes comments on the Science Assessment, as presented in draft IRP Chapter 4. In Section 3, we provide brief comments on the quantitative risk and exposure assessments, as presented in draft IRP Chapter 5. Finally, in Section 4, we provide specific recommendations for the Clean Air Scientific Advisory Committee (CASAC) to consider as it provides feedback to EPA on the draft IRP.

2 Science Assessment

The draft IRP (and ISA) need to provide a more detailed protocol.

The draft IRP presents the ISA development steps in Figure 4-1 and indicates the steps are described in greater detail in the Preamble to the ISAs. The Preamble, which is intended to serve as a companion document to ISAs, outlines the general process used by EPA for developing the ISAs, including the overall framework for evaluating evidence and making causal conclusions (US EPA, 2015). The Preamble discusses several important considerations for robust scientific review, including an analysis of strengths and limitations of the available evidence. However, both the Preamble and the draft IRP primarily discuss overarching principles; neither provides specific instructions with sufficient detail. The Preamble should be thought of as a standard operating procedure for ISAs, and the IRP as the pollutant-specific protocol for the ISAs and other assessments. In both cases, step-by-step instructions are necessary. As they currently stand, these documents do not have sufficient detail to ensure that studies are identified and reviewed in a systematic and consistent manner, or integrated in a way that considers study quality and the coherence of results across studies within and across disciplines. In addition, the National Ambient Air Quality Standards (NAAQS) causal framework tends to bias determinations in the direction of causality (see Section 2.7). Finally, the current framework is not detailed enough for replication by others.

We understand that these documents need to maintain some flexibility to allow for pollutant- and discipline-specific issues and an iterative process, but a certain degree of prescriptiveness is required to facilitate a transparent, objective assessment. The IRP should include a protocol that includes well-developed methods for the literature search strategy (including keywords and databases to be used[1]); study inclusion/exclusion criteria[2]; the process for data extraction and quality control; specific, prescriptive criteria for evaluating study quality; potential analyses (*e.g.*, pooling results across studies, as in a meta-analysis); and ozone-specific methods for evidence integration and causality determinations (including plans for assessing data gaps, limitations, and uncertainties in the evidence and the overall systematic review). While the draft IRP contains information for some of these steps, others are undeveloped or lack critical details. Given that this assessment is informed by the 2013 ISA, many of the challenges and critical issues surrounding ozone-related health effects have already been identified (*i.e.*, problem formulation is well developed), so providing a detailed protocol in the IRP should be achievable. Developing a detailed *a priori* protocol for the systematic review will limit potential biases in the review and help ensure that its results can be reproduced by others.

Below we provide several examples for considerations that should be included in both the IRP and the ISA, based on a review of several systematic review/weight-of-evidence approaches, including another EPA office (*e.g.*, EFSA, 2017; TCEQ, 2017; US EPA, 2018; Goodman *et al.*, 2013, 2015; Rhomberg *et al.*, 2013). In addition, we have attached a recent analysis by Zu *et al.* (2018) on long-term exposure to ozone and asthma that shows how these aspects should be laid out in the IRP and applied in the ISA and REA.

[1] These are provided in the draft IRP, but keywords are far too general ("ozone") and should be further tailored.

[2] The draft IRP contains only exclusion criteria regarding study type. The IRP should also include other exclusion criteria. For example, animal toxicity studies of mixed pollutant exposures that do not conduct statistical analyses on the independent effects of ozone should be excluded.

2.1 Scope of the ISA (4.3.2)

All existing and new studies should be evaluated and weighed based on their relative quality.

The draft IRP states that the focus of the forthcoming ISA will be studies published since 2011, and new evidence will be synthesized and integrated in the context of the conclusions from the previous review. EPA indicates it will more fully evaluate any effects in which conclusions were less certain (*i.e.*, "likely to be causal," "suggestive," or "inadequate," as described below), and new information has become available. For those health effects for which the 2013 ISA concluded that evidence was sufficient to infer a causal relationship, EPA "will integrate and synthesize the new evidence, placing emphasis on policy-relevant considerations, such as the concentrations at which effects are observed, and characterizing the extent to which new studies address key uncertainties and limitations identified in the previous review or provide insight on new issues."

There are a few issues with this approach. Perhaps most importantly, study quality was not sufficiently considered in the 2013 ISA, which resulted in conclusions that were not based on the true weight of evidence. It is not clear how new evidence and that previously considered will be integrated if study quality is not considered in a consistent manner across all studies. In addition, the draft IRP does not explain what will be done if new evidence adds uncertainty (*i.e.*, is contradictory) to existing conclusions. New evidence will need to be assessed for quality, and all studies (existing and new) should be weighed according to their relative study quality to determine whether the 2013 ISA causal conclusions should be revised.

2.2 Literature Search and Selection of Relevant Studies (4.3.3)

The IRP should provide more detail regarding its literature search strategy.

The draft IRP describes a literature search strategy, but it does not provide specific search terms other than "ozone" and, as noted before, provides limited specific study inclusion and exclusion criteria. The IRP should be updated to include more refined search terms and additional clarity with regard to some of the automated literature search and selection tools EPA plans to use, such as the SWIFT-Active Screener software. All of this should be determined *a priori*, with any changes documented and discussed in the ISA. EPA should consider closely following the Cochrane Review process, which serves as a basis for several weight-of-evidence (WoE) frameworks, such as the Office of Health Assessment and Translation within the National Toxicology Program (The Cochrane Collaboration, 2011; NTP OHAT, 2015). EPA should also consider the methods of EPA/OPPT outlined in the new Toxic Substances Control Act (TSCA) systematic review guidance (US EPA, 2018). For example, the TSCA framework recommends conducting a pilot study to test the search screening and inclusion criteria for full text screening and tagging, and suggests stipulations regarding conflict resolution strategies. This allows a place to stop and revise processes that may not be working as well in practice as anticipated. As long as process modifications are documented, these iterative search and selection methods will ensure the most relevant and useful data are identified in a consistent manner for inclusion in the ISA.

The draft IRP notes that study inclusion and exclusion decisions will be documented in the HERO database but does not state what these criteria are. This is a critical step, and it should be stated in the IRP. A specific example of a list of criteria is provided in Zu *et al*. (2018):

> We included peer-reviewed observational studies of humans that evaluated the associations between long-term ozone exposure (defined as approximately 30 d or longer induration;

> US EPA, 2013a) and incident asthma (i.e. incidence, incidence rate or lifetime prevalence of asthma). We excluded studies that met any of our exclusion criteria: laboratory animal or *in vitro* studies; experimental studies; reviews; editorials; commentaries; correspondence/communications; letters to the editor; studies that were not published in English; studies that did not evaluate long-term exposure to ambient ozone (e.g. studies that evaluated short-term exposure or indoor exposure); studies that did not evaluate incident asthma (e.g. studies that evaluated prevalence of asthma or asthma exacerbation); and studies that did not report specific results on the associations between long-term ozone exposure and incident asthma.

It is also important that the process of study screening and selection for a systematic review is clearly presented. An example summary of this information is shown in Figure 1 of Zu *et al*. (2018). This is expected of all systematic reviews to ensure that no relevant study is ignored, and so that the literature searches and study selection can be reproduced by others.

Finally, it is standard systematic review practice to have two independent reviewers conduct the literature search and study selection and have a third person resolve disagreements (see, for example, discussion of this process in the TSCA framework [US EPA, 2018] and Goodman *et al*., 2015). The IRP should indicate whether this is the case for the ISA.

2.3 Discipline-specific Scoping, Searching, and Screening (4.3.4)

Statistical significance and biological significance should be considered.

The draft IRP outlines the scope of the ISA using PECOS (*i.e*., populations, exposures, comparators, outcomes, study designs) statements for each realm of evidence. The main components of PECOS statements are generally clearly written. However, included at the end of each PECOS statement is the following note: "Also of interest is the lowest concentration that produces a measureable change in risk." It is unclear what is meant by "measureable." The IRP should more fully describe how EPA will evaluate this critical issue (which is essentially the point of departure for each identified endpoint). EPA should go beyond simply determining whether an effect was measureable at all and take into account statistical significance and biological relevance when evaluating effects. More specifically, the IRP should articulate how statistical significance is considered when evaluating studies of different designs and in different disciplines, and it should specify the criteria that define adversity, *i.e*., when a measurable change reflects a true increase in disease risk, rather than a homeostatic response. This is because a biological effect may be upstream of an adverse, apical effect, but the occurrence of such an effect does not necessarily mean that apical effect will occur (*e.g*., see Goodman and Lynch, 2017).

2.4 Identification of Policy-relevant Studies (4.3.5)

Study quality should not be a criterion for study inclusion.

The draft IRP indicates that study quality should not be the sole criterion for study inclusion. Study quality is extremely important when evaluating study results and integrating evidence, but it should not have any role in determining study inclusion. Study inclusion should be based on relevance alone, so that all studies that contain relevant information are included. Study quality should be evaluated in a later step in the process.

2.5 Evaluation of Individual Study Quality (4.3.6)

The IRP and Preamble should provide explicit details regarding study quality metrics.

On page 4-17, the draft IRP states:

> After selecting studies for inclusion, individual study quality is evaluated by considering the design, methods, conduct, and documentation of each study, but not the study results. In the ISA for the current review, conclusions about the strength of inference from study results will be made by independently evaluating the overall quality of each study.

This is extremely important for a systematic review—an evaluation of study quality should be independent of the results and funding source of that study; it should be based purely on the reported methods, in a consistent manner across studies. Further, studies with more robust methods should receive more weight in causal determinations. Currently, the draft IRP provides no explicit rationale for why certain studies are considered key evidence while others of similar quality are not.

The draft IRP discusses several metrics for evaluating study quality and states that additional considerations are discussed "in detail in the Preamble to the ISAs." However, these issues are not actually discussed in detail in either document. Rather, a set of seven general considerations are provided. Notably, criteria are not specific to each type of evidence (*e.g.*, animal, epidemiology, controlled human exposure). The IRP and the Preamble should both specify criteria for assessing these metrics and ensure they are applied in a consistent manner across studies. Further, these quality criteria should be specific to each discipline and study design. For example, the Integrated Science Assessment for Oxides of Nitrogen – Health Criteria (US EPA, 2016) includes specific criteria for different disciplines (*i.e.*, epidemiology, controlled human exposure, animal toxicity). Although these criteria are not entirely comprehensive, sufficiently detailed, or specific to different study designs, they can be used as a starting point for the ozone ISA. For example, the IRP should discuss all of the ways in which ozone exposure can be measured, the strengths and limitations of each method, the potential for exposure measurement error, and which methods carry the most weight. There should also be a discussion of statistical methods used among all studies evaluated and which specific methods are more robust and why (*e.g.*, whether multiple comparisons have been addressed or whether the assumption for the Cox proportional hazard model has been tested and satisfied). The IRP should address specific confounders (*e.g.*, co-pollutants, socioeconomic status, age, weather) in terms of how they are handled in different studies and their likely impact on results. Other factors the IRP should consider in detail include measurement bias, measurement precision, replicability of observations, data reliability, outliers, selective outcome reporting, and fraudulent studies.

It is critical that the ISA evaluates these quality measurements in the same way across studies. For example, if EPA considers a particular statistical model a limitation for one study, it must consider it a limitation in all studies that use the same model (unless there is a reason to conclude otherwise; if this is the case, the ISA should state the reason clearly).

There are numerous study quality evaluation systems, some of which include all lines of evidence (see, for example, the NTP OHAT and ScriRAP systems [NTP OHAT, 2015; Beronius *et al*., 2018). Several good examples of peer-reviewed study quality criteria and evaluations can be found in LaKind *et al*. (2018) Goodman *et al*. (2018) (in Sections S1.2 in each of the three supplements) and Zu *et al*. (2018) (Study Quality Evaluation in Methods and Results and Table 2). These systematic reviews explicitly state the metrics used to determine study quality for each discipline, and what constitutes higher and lower quality is also explicitly stated. Quality considerations for each study are tabulated, with studies in rows and study

quality aspects in columns. All low-quality metrics are highlighted or color-coded, so that the quality of the literature is clear.

It should be evident from these papers that several study quality metrics are specific to the realm of evidence (*i.e.*, epidemiology, controlled exposure studies, animal toxicity studies), individual study design, and also to the exposure and outcome. Quality criteria should not (and cannot) be the same for everything. The IRP should specifiy what these criteria are for the studies the ISA will review, so that the public and CASAC can review the criteria and make recommendations.

Study quality criteria need to be applied to every study reviewed in the ISA, including animal toxicity and *in vitro* studies and studies evaluating "at-risk" factors. By precisely defining and applying study quality criteria to all studies, the ISA will be more transparent and balanced.

In addition, the past ISAs did not sufficiently address study quality when evaluating exposure-response data. While the lack of a thorough, systematic study quality evaluation is an issue for determining causation, it is even more problematic in the context of concentration-response relationships. For causal determinations, studies need to establish the presence of an effect following an exposure, but for concentration-response relationships, studies need to not only establish the presence of an effect, but also the magnitude of an effect in relation to the level of the exposure. The draft IRP should articulate detailed and more stringent quality considerations for the concentration-response data, because they will have the most impact on the level of the ozone NAAQS.

Finally, a related, but separate, issue to study quality is study relevance. This can be an issue for epidemiology studies (*e.g.*, generalizability, relevance to the US population), but it is a critically important consideration for toxicity and mechanistic studies. The IRP should explicitly state criteria that must be met for study results to be considered relevant.

2.6 Integration of Evidence and Determination of Causality (4.3.7)

Evidence integration and causal determination methods should be more robust and explicit.

The draft IRP indicates that EPA will use a structured framework to integrate and classify the weight of available evidence, as described in the Preamble to the ISAs (US EPA, 2015). Specifically, EPA will classify the weight of available evidence for health and welfare effects according to the NAAQS causal framework, based on a five-level hierarchy: (1) causal relationship; (2) likely to be a causal relationship; (3) suggestive of, but not sufficient to infer, a causal relationship; (4) inadequate to infer the presence or absence of a causal relationship; and (5) not likely to be a causal relationship. This hierarchy is similar to the Institute of Medicine (IOM) report *Improving the Presumptive Disability Decision-making Process for Veterans* (IOM, 2008), except that the IOM recommended four levels of evidence categories for causation. Unlike the IOM determinations, however, the exact process whereby causal conclusions are reached in the ISA are not entirely clear in either the IRP or the Preamble. Both documents provide only high-level guidance on the methodology for the ISAs.

The draft IRP notes that the key issues considered in evaluating causality include "consistency of findings for an endpoint across studies, coherence of the evidence across disciplines and across related endpoints, and biological plausibility." As indicated in the Preamble, this is based on EPA's modified Bradford Hill aspects. Both the original and modified Bradford Hill aspects (*i.e.*, strength of association, consistency and coherence, biological plausibility, biological gradient or exposure-response, specificity, temporality of effect, and adversity) are useful tools for evaluating causation. This is because it may be difficult to ascribe

observations to causation if these aspects are not met, whereas it may be difficult to ascribe observations to anything other than causation if they are met. However, as discussed in Goodman *et al*. (2013), EPA's application of the causal framework for NAAQS reviews is not congruent with the judgments based on the original or modified Bradford Hill aspects. For example, the NAAQS framework claims to rely heavily on the aspect of consistency across studies in its categorization scheme, but, in practice, it does not always fully evaluate consistency or fully incorporate aspects such as coherence, biological plausibility, biological gradient, and strength of association.

The draft IRP indicates that "studies in which chance, confounding, and other biases could be ruled out with reasonable confidence are sufficient to infer a causal relationship." However, "reasonable confidence" is not explicitly defined; it is unclear how one can determine what is reasonable, particularly considering all studies have at least some methodological limitations and some uncertainties in results (*e.g.,* the possibility for residual confounding and other issues). Any WoE evaluation, by definition, involves a consideration of all lines of evidence in a consistent manner. It is not about resolving all uncertainty but, rather, determining whether the evidence as a whole supports causation more than it supports a lack of effect. If co-pollutants cannot be addressed or results are inconsistent, the WoE may indicate a lack of causality or inadequate evidence to assess causation. If positive effects in well-conducted, high-dose animal studies cannot be extrapolated to humans, this does not constitute suggestive evidence; instead, these effects are essentially uninformative regarding causation in humans. Not every study evaluating a criteria pollutant is informative for evaluating human health risk, and EPA should not place undue weight on studies that are not relevant or informative.

With regard to the process of evidence integration, the draft IRP indicates that all types of evidence may not be given equal consideration in the ISA. For example, the draft IRP states, "Other information including mechanistic evidence, toxicokinetics, and exposure assessment may be drawn upon if relevant to the evaluation of health effects and if of sufficient importance to affect the overall evaluation." Mechanistic information, in particular, should be considered concurrently with the other bodies of evidence during the evidence integration process. Well-conducted, relevant mode-of-action studies, if their results are relevant to humans, should be given equal weight to well-conducted controlled human exposure studies, for example.

It is notable that the NAAQS causal framework requires only one high-quality study for evidence of causal relationship to be deemed *suggestive.* Using this definition, high-quality studies that are inconsistent with evidence of an association may exist but—as long as one high-quality study demonstrates an effect—there would still be enough evidence to constitute a suggestive relationship. Instead, all studies (including positive, null, and negative evidence) should be reviewed using the same criteria. Then, the evidence should be evaluated as a whole—not only based on how much evidence supports (or counters) the hypothesized causal effect, but how separate lines of evidence support (or contradict) one another. That is, it is critical to determine the most likely explanation for discrepancies across studies by evaluating all of the evidence and not selectively considering evidence that supports or counters a given hypothesis. In situations where there are multiple, but inconsistent, high-quality studies, the appropriate conclusion is that the evidence is inadequate to determine causality.

Many of the issues noted above could be resolved if EPA revises the NAAQS framework to make categories for causal determination more similar to the IOM framework on which it was based originally (See Table 2.1, below). As noted above, the IOM framework does not have a "suggestive" category; the presence of suggestive animal studies but weak or inconsistent human evidence, for example, warrants a "below equipoise" determination. This translates to a conclusion that "the evidence is not sufficient to conclude that a causal relationship is at least as likely as not, or is not sufficient to make a scientifically informed judgment." Using the NAAQS framework in this same example would likely lead to a "suggestive of a

causal relationship" determination. This demonstrates how the NAAQS causal framework biases conclusions toward causality.

Fully adopting the IOM approach would allow EPA to evaluate all evidence in a consistent manner using well-specified criteria and determine whether, as a whole, it constitutes evidence for causation or is more likely indicative of an alternative hypothesis. EPA should proceed with a risk assessment on a particular health effect only if the evidence is clearly supportive of causation (*i.e.*, equipoise and above in the IOM framework). With regard to evidence integration methodology, EPA should consider reviewing and drawing from the recent European Food Safety Authority (EFSA) *Guidance on the Use of the Weight of Evidence Approach* in scientific assessments (EFSA, 2017).

Regardless of whether the overarching framework is revised for the purposes of the forthcoming ISA, the ozone IRP should be revised to provide a detailed description of how EPA will evaluate causality in a consistent manner, considering these critical issues. As discussed above, while it is necessary for the IRP to maintain some flexibility to allow for chemical- and study-specific issues and an iterative process, a certain degree of prescriptiveness is required to facilitate a clear, objective assessment. Descriptions of the evidence integration methods and the process of causal determinations require the highest level of detail, so that the basis for EPA's conclusions is clear.

Table 2.1 IOM's Recommended Categories for the Level of Evidence for Causation

Causal Determination	Evidence
Sufficient	The evidence is sufficient to conclude that a causal relationship exists. For example: (a) replicated and consistent evidence of an association from several high-quality epidemiologic studies that cannot be explained by plausible noncausal alternatives (*e.g.*, chance, bias, or confounding); or (b) evidence of causation from animal studies and mechanistic knowledge; or (c) compelling evidence from animal studies and strong mechanistic evidence from studies in exposed humans, consistent with (*i.e.*, not contradicted by) the epidemiologic evidence.
Equipoise and Above	The evidence is sufficient to conclude that a causal relationship is at least as likely as not, but not sufficient to conclude that a causal relationship exists. For example: (a) evidence of an association from the preponderance of several high-quality epidemiologic studies that cannot be explained by plausible noncausal alternatives (*e.g.*, chance, bias, or confounding) as well as animal evidence and biological knowledge consistent with a causal relationship; or (b) strong evidence from animal studies or mechanistic evidence that is not contradicted by human or other evidence.
Below Equipoise	The evidence is not sufficient to conclude that a causal relationship is at least as likely as not, or is not sufficient to make a scientifically informed judgment. For example: (a) consistent human evidence of an association that is limited by the inability to rule out chance, bias, or confounding with confidence, and weak animal or mechanistic evidence; or (b) animal evidence suggestive of a causal relationship, but weak or inconsistent human and mechanistic evidence; or (c) mechanistic evidence suggestive of a causal relationship, but weak or inconsistent animal and human evidence; or (d) the evidence base is very thin.
Against	The evidence suggests the lack of a causal relationship. For example: (a) consistent human evidence of no causal association from multiple studies covering the full range of exposures encountered by humans; or (b) animal or mechanistic evidence supportive of a lack of a causal relationship.

Notes:
IOM = Institute of Medicine.
Source: IOM (2008).

2.7 Causality Determination from the 2013 ISA (4.4.1)

The IRP should indicate the ISA will evaluate evidence in an unbiased manner.

The draft IRP indicates in this section that it plans to address this question: "Does the evidence base from recent studies contain new information to support or call into question the causality determinations made for relationships between O_3 exposure and various health and welfare effects in the 2013 ISA?"

It is notable that the question is phrased such that it is not biased towards information strengthening a causal conclusion but rather objectively looking at the state of the science as a whole. This phrasing should be used throughout the IRP. For example, two questions below the previous question, the draft IRP indicates the ISA will address: "Does new evidence confirm or extend biological plausibility of O_3-related health effects?", which implicitly omits the possibility that new evidence may refute biological plausibility of ozone-related health effects. Instead, the ISA should evaluate whether new evidence supports or calls into question the biological plausibility of ozone-related health effects.

2.8 Health Effects (4.4.4)

The IRP should indicate the ISA should determine whether experimental studies support or call into question causal relationships.

In Section 4.4.4, the draft IRP indicates the ISA will address whether recent controlled human exposure and toxicity studies "continue to provide support for biologically plausible relationships" between short-term and long-term ozone exposures and respiratory health effects and whether recent studies report ozone-attributable effects at lower ozone exposure concentrations or for different durations or patterns of exposure than indicated by studies available in the last review.

These are very important issues to address. However, as discussed above, as phrased, this implies a bias toward finding effects. EPA should also consider whether scientific evidence supports the occurrence of health effects from ozone exposure only at *higher* concentrations than previously observed. For example, it is possible high-quality studies have emerged that support a threshold for ozone-related health effects that is higher than previously thought.

The IRP should describe how the ISA will determine when evidence calls into question a causal association. For example, it is often the case that evidence indicates a lack of causation to be as likely, or even more likely, than causation (*e.g.*, if confounders cannot be totally accounted for or if exposure misclassification causes false positive results). In other cases, mechanistic data should be considered when actual adverse effects have been established. As discussed above, a biological effect may be upstream of an adverse, apical effect, but the occurrence of such an effect does not necessarily mean that apical effect will occur (*e.g.*, see Goodman and Lynch, 2017). When true adverse effects do not occur, it is not appropriate to conclude that mechanistic data support causality because of the *possibility* for causation. Instead, one should conclude the mechanistic evidence alone is non-informative.

2.9 At-Risk Life Stages and Populations and Public Health Impact (4.4.5)

EPA should evaluate "at-risk" factors with the same rigor as it evaluates ozone exposures.

The draft IRP indicates that the evaluation of recent evidence will build on the conclusions from the 2013 ISA, where application of the at-risk framework to classify evidence demonstrated that these populations are at increased risk of ozone-related health effects. In the footnote on page 4-29, the draft IRP states

> In recent reviews, the term "at-risk" has been used to define populations and lifestages potentially at increased risk of an air pollutant-related health effect (e.g., see 2013 O_3 ISA and 2016 NO_X ISA; U.S. EPA, 2013; U.S. EPA, 2016). At-risk populations can include those with intrinsic factors that make them more susceptible to pollutant-related effects (e.g., pre-existing disease, genetic characteristics) or that increase pollutant dose (e.g., breathing patterns), and extrinsic factors that could increase pollutant exposures (e.g., activity patterns; proximity to sources) (U.S. EPA, 2016, pp. 1xiii to 1xiv).

An "at-risk" factor is more accurately described as an effect modifier, which is a technical term defined in epidemiology as a variable that modifies the observed effect of an exposure (in this case, ozone) on disease risk. To the extent that "at-risk" factors are evaluated in the ISA based on epidemiology studies, it may be more appropriate for EPA to refer to them as effect modifiers, as we do herein.

Setting that aside, as noted in Section 4.3.7 of the draft IRP, EPA's frameworks for both causal determination and classification of evidence for potential effect modifiers are based IOM framework, but the former has five categories and the latter four, and the criteria for classification are not the same. It appears that, although defined differently, the four categories in the IOM framework for effect modifiers are roughly equivalent to *causal relationship, suggestive of a causal relationship, inadequate to infer a causal relationship, and not likely to be a causal relationship* in the NAAQS causal framework; there does not appear to be an equivalent to the *likely to be a causal relationship* category. Although one is an assessment of direct causation and the other an assessment of factors that can contribute to (or prevent) causation, in both cases, the goal is to critically, systematically, and transparently review the weight of scientific evidence, paying particular attention to study quality and relevance, and the consistency and coherence of effects across studies. Ideally, the same rules should be applied for both types of analysis; if not, there needs to be justification for using different rules to conduct the same type of analysis. As discussed above, the IRP and ISA should adopt the IOM-recommended categories for the level of evidence for causation, which consider whether the WoE is above or below equipoise (IOM, 2008).

We also note that issues with the WoE for causal determination apply to the effect modifier classifications as well (see Goodman *et al*., 2013). For example, the ISA defines evidence for an effect modifier to be suggestive if it "is limited due to some inconsistency within a discipline or, where applicable, a lack of coherence across disciplines." If an inconsistency or a lack of coherence is large enough, it should lead to a conclusion that the WoE is below equipoise. In these circumstances, the evidence is inadequate to make a determination as to whether a factor is an effect modifier for an ozone-health endpoint association.

3 Quantitative Risk and Exposure Assessments

EPA should clarify the REA methods.

The Quantitative Risk and Exposure Assessments Section outlines overarching principles of the risk and exposure assessment and defines the populations and pollutant concentrations of most interest to public health for the purpose of policy decision-making. Generally, EPA is most concerned with the nature and magnitude of exposure and potential risk associated with air quality conditions that just meet the current standard, particularly in at-risk populations.

With regard to the specific plans outlined for the REA, the draft IRP provides very little detail. EPA indicates that it plans to focus on key uncertainties and limitations of the analyses conducted for the last ozone review, considering whether and to what extent new information, tools, or methods could address these issues. However, the specific methods for identifying and evaluating the value of these new resources is not discussed.

Further, there is very little discussion on how EPA will decide on the specific analyses most appropriate for the forthcoming REA. For example, on page 5-18, the draft IRP states, "Given the rapid timeline for this review, we would expect to focus on a streamlined set of study areas and air quality scenarios compared to the expansive set assessed in the last review." It is not clear what type of study areas will be assessed, and thus the IRP should be revised to include this information. It appears as though only urban areas have been studied in the past, and EPA should consider exploring rural and suburban areas, as well.

The IRP should properly characterize the uncertainty surrounding controlled human exposure studies.

The draft IRP indicates that for quantitative purposes, EPA will continue to rely upon exposure-based risk analyses, particularly the analyses used in the last three ozone reviews, which involve comparisons of estimated population-based ozone exposures while at elevated exertion to benchmark concentrations in the controlled human exposure studies.

The draft IRP states that the decision to focus on exposure-based analyses is largely based on the many issues and uncertainties with the epidemiology-based risk analysis, and EPA concludes that it is unlikely sufficient new information would be available to improve upon issues with the epidemiology risk-based approach. Overall, this focus on the exposure-based risk analyses is well supported, but the draft IRP should stipulate that the ISA provide a discussion of how any new epidemiology studies are or are not sufficient to use such an approach. Further, in general, the description of the way in which the IRP will conduct the exposure-based risk approach is too general (although there is some indication that the focus will remain on benchmark concentrations). With regard to the controlled human exposure studies, EPA simply re-emphasizes conclusions from previous reviews, including "effects reported in controlled human exposure studies are due solely to O_3 exposures, and interpretation of study results is not complicated by the presence of co-occurring pollutants or pollutant mixtures (as is the case in epidemiologic studies)."

There are numerous controlled human exposure studies of ozone and respiratory outcomes. These studies are unique in that they involve controlled, experimental conditions with specific ozone exposure levels. Participants typically also serve as their own controls. As such, we agree that these studies are superior to

many epidemiology studies on ozone-associated respiratory effects. However, the draft IRP does not discuss the implications of the limitations of controlled human exposure studies—most notably, that these studies by design can only induce temporary, reversible effects that may or may not be clinically significant under real-world conditions, and that the small, homogenous participant populations in these studies limit the extrapolation of results to the US population. Interestingly, the draft IRP notes the latter issue in Section 4.3.5 but does not mention it in Chapter 5. Therefore, it is unclear how EPA plans to consider issues with extrapolation of these small studies. The IRP should be revised to include a focus on detailing the specific methods and tools for addressing these very critical issues when relying on controlled human exposure studies.

Finally, with regard to the risk assessments, as noted above, the draft IRP suggests that it will again assess the response to exposure at "elevated exertion." In the previous analysis, ozone exposure and risk were overestimated because EPA focused on children with "moderate or greater exertion level at the time of exposure" (and thus, very high ventilation rates) who do not exhibit averting behavior (*i.e.*, staying indoors when there are high ozone concentrations) (US EPA, 2014, p. 5-2). In addition, in the previous REA, APEX-modeled FEV_1 decrements greater than 10%, greater than 15%, and greater than 20% showed a large degree of both statistical and model uncertainty. The IRP should provide more detail with regard to EPA's plans to address some of these critical issues from the prior ozone review.

The REA should consider exploring the use of a threshold concentration-response function.

The word "threshold" is not mentioned in Section 5 of the draft IRP. Given that the proposed MoAs for ozone-induced health effects are likely to operate only above a threshold, the IRP (and the REA) should consider using a threshold concentration-response function, if data allow.

Uncertainty should be addressed in the risk estimates.

The IRP should include plans to address uncertainties in the risk estimates. To reduce uncertainty overall, quantitative risk estimates should focus on endpoints for which there is strong evidence of a causal association, with robust data for both air quality evaluations and concentration-response functions. To the extent possible, quantitative estimates of uncertainty should be incorporated into the confidence bounds around risk estimates provided in the REA. If it is not possible to quantify certain aspects of uncertainty, then the REA should indicate whether the uncertainty is likely to over- or underestimate risks and provide a qualitative indication of the magnitude of the uncertainty (*e.g.*, high, medium, or low). The IRP should indicate that, if an aspect of uncertainty could produce both outcomes, the REA should provide examples of when and how the uncertainty would lead to an under- and overestimate of risks.

Table 3.1 IOM Recommended Categories for the Level of Evidence for Causation

Causal Determination	Evidence
Sufficient	The evidence is sufficient to conclude that a causal relationship exists. For example: a) replicated and consistent evidence of an association from several high-quality epidemiologic studies that cannot be explained by plausible noncausal alternatives (*e.g.*, chance, bias, or confounding); or b) evidence of causation from animal studies and mechanistic knowledge; or c) compelling evidence from animal studies and strong mechanistic evidence from studies in exposed humans, consistent with (*i.e.*, not contradicted by) the epidemiologic evidence.
Equipoise and Above	The evidence is sufficient to conclude that a causal relationship is at least as likely as not, but not sufficient to conclude that a causal relationship exists. For example: a) evidence of an association from the preponderance of several high-quality epidemiologic studies that cannot be explained by plausible noncausal alternatives (*e.g.*, chance, bias, or confounding) as well as animal evidence and biological knowledge consistent with a causal relationship; or b) strong evidence from animal studies or mechanistic evidence that is not contradicted by human or other evidence.
Below Equipoise	The evidence is not sufficient to conclude that a causal relationship is at least as likely as not, or is not sufficient to make a scientifically informed judgment. For example: a) consistent human evidence of an association that is limited by the inability to rule out chance, bias, or confounding with confidence, and weak animal or mechanistic evidence; or b) animal evidence suggestive of a causal relationship, but weak or inconsistent human and mechanistic evidence; or c) mechanistic evidence suggestive of a causal relationship, but weak or inconsistent animal and human evidence; or d) the evidence base is very thin.
Against	The evidence suggests the lack of a causal relationship. For example: a) consistent human evidence of no causal association from multiple studies covering the full range of exposures encountered by humans; or b) animal or mechanistic evidence supportive of a lack of a causal relationship.

Source: IOM (2008).

Table 3.2 EPA's Weight of Evidence for Causal Determination

Causal Determination	Health Effects
Causal relationship	Evidence is sufficient to conclude that there is a causal relationship with relevant pollutant exposures (*i.e.*, doses or exposures generally within one to two orders of magnitude of current levels). That is, the pollutant has been shown to result in health effects in studies in which chance, bias, and confounding could be ruled out with reasonable confidence. For example: a) controlled human exposure studies that demonstrate consistent effects; or b) observational studies that cannot be explained by plausible alternatives or are supported by other lines of evidence (*e.g.*, animal studies or mode of action information). Evidence includes multiple high-quality studies.
Likely to be a causal relationship	Evidence is sufficient to conclude that a causal relationship is likely to exist with relevant pollutant exposures, but important uncertainties remain. That is, the pollutant has been shown to result in health effects in studies in which chance and bias can be ruled out with reasonable confidence but potential issues remain. For example: a) observational studies show an association, but co-pollutant exposures are difficult to address and/or other lines of evidence (controlled human exposure, animal, or mode of action information) are limited or inconsistent; or b) animal toxicological evidence from multiple studies from different laboratories that demonstrate effects, but limited or no human data are available. Evidence generally includes multiple high-quality studies.
Suggestive of a causal relationship	Evidence is suggestive of a causal relationship with relevant pollutant exposures, but is limited. For example, (a) at least one high-quality epidemiologic study shows an association with a given health outcome but the results of other studies are inconsistent; or (b) a well-conducted toxicological study, such as those conducted in the National Toxicology Program (NTP), shows effects in animal species.
Inadequate to infer a causal relationship	Evidence is inadequate to determine that a causal relationship exists with relevant pollutant exposures. The available studies are of insufficient quantity, quality, consistency, or statistical power to permit a conclusion regarding the presence or absence of an effect.
Not likely to be a causal relationship	Evidence is suggestive of no causal relationship with relevant pollutant exposures. Several adequate studies, covering the full range of levels of exposure that human beings are known to encounter and considering at-risk populations, are mutually consistent in not showing an effect at any level of exposure.

Source: US EPA (2015, Table II).

4 Recommendation for CASAC

As it currently stands, the IRP does not have sufficient detail to ensure that studies are identified and reviewed in a systematic and consistent manner or integrated in a way that considers study quality and the coherence of results across studies, both within and across disciplines. This may result in causality determinations that are not based on the weight of the scientific evidence. To address this, CASAC should recommend that the draft IRP be revised to include the following:

- **A detailed, objective, and transparent systematic review protocol for the ISA that includes details for each step of the review process.** The IRP must include more detail regarding the specific methodology EPA will employ for the literature search, literature selection, study quality evaluation, evidence integration, and causality determinations (see Section 2).
- **A more objective and robust causality evaluation system, such as the Institute of Medicine (IOM) framework upon which the existing framework is based.** The draft IRP uses the NAAQS causality framework, which is biased toward affirmative causality conclusions (see Section 2.7). Reliance on a framework that more closely resembles the original IOM framework will lead to more objective causal conclusions.
- **Clear study selection criteria and a thorough, well-described study quality evaluation system.** The draft IRP describes only general quality considerations, but not specific criteria (see Sections 2.2-2.5). Specific criteria, and what is required to meet them, are necessary for a systematic, consistent review.
- **An *a priori* methodology for evidence integration that includes a detailed process for combining evidence included in the 2013 ISA with new evidence.** This should include a description of how inconsistencies between new and existing evidence will be evaluated and the potential implications for revised causal determinations (see Sections 2.6 and 2.7).
- **Methods for evaluating at-risk factors (*i.e.*, effect modifiers, such as pre-existing disease) with the same rigor as those for evaluating studies of ozone health effects.** Study quality should be considered in evaluations of effect modifiers, and conclusions regarding whether something is an effect modifier should be made using the IOM framework (see Section 2.9).
- **More detail regarding methods for conducting a potential REA.** This includes additional discussion of new information, tools, or methods that will be used to address uncertainties in the REA discussed in the last review (see Section 3).
- **Better characterize uncertainty.** This includes uncertainty surrounding controlled exposure studies. Also, all aspects of uncertainty should be quantitatively addressed in risk estimates (see Section 3).
- **Explore the use of a threshold concentration-response function.** The IRP should specify that EPA explore the possibility of a threshold concentration-response function when modeling the risk of ozone-induced health effects, if data allow. The available mode-of-action evidence suggests that ozone-associated health effects have a threshold concentration under which they are not likely to occur (see Section 3).

Overall, CASAC should recommend that EPA provide a more detailed, objective, and transparent methodology in the IRP, as laid out in the bullets above. This will limit potential biases in the ISA and potential REA, help ensure that the analyses can be reproduced by others, and increase confidence in the causal determinations regarding ozone-associated effects.

References

Beronius, A; Molander, L; Zilliacus, J; Ruden, C; Hanberg, A. 2018. "Testing and refining the Science in Risk Assessment and Policy (SciRAP) web-based platform for evaluating the reliability and relevance of in vivo toxicity studies." *J. Appl. Toxicol.* doi: 10.1002/jat.3648.

European Food Safety Authority (EFSA) Scientific Committee. 2017. "Guidance on the use of the weight of evidence approach in scientific assessments." *EFSA J.* 15(8):4971. doi: 10.2903/j.efsa.2017.4971.

Goodman, J; Lynch, H. 2017. "Improving the International Agency for Research on Cancer's consideration of mechanistic evidence." *Toxicol. Appl. Pharmacol.* 319:39-46. doi: 10.1016/j.taap.2017.01.020.

Goodman, JE; Prueitt, RL; Sax, SN; Bailey, LA; Rhomberg, LR. 2013. "Evaluation of the causal framework used for setting National Ambient Air Quality Standards." *Crit. Rev. Toxicol.* 43(10):829-849. doi: 10.3109/10408444.2013.837864.

Goodman, JE; Prueitt, RL; Sax, SN; Pizzurro, DM; Lynch, HN; Zu, K; Venditti, FJ. 2015. "Ozone exposure and systemic biomarkers: Evaluation of evidence for adverse cardiovascular health impacts." *Crit. Rev. Toxicol.* 45(5):412-452. doi: 10.3109/10408444.2015.1031371.

Institute of Medicine (IOM), Committee on Evaluation of the Presumptive Disability Decision-Making Process for Veterans, Board on Military and Veterans Health. 2008. "Improving the Presumptive Disability Decision-Making Process for Veterans." (Eds.: Samet, JM; Bodurow, CC), National Academies Press, Washington, DC, 781p. Accessed at http://books.nap.edu/openbook.php?record_id=11908&.

LaKind, JS; Lehmann, GM; Davis, MH; Hines, EP; Marchitti, SA; Alcala, C; Lorber, M. 2018. "Infant dietary exposures to environmental chemicals and infant/child health: A critical assessment of the literature." *Environ. Health Perspect.* 126(9):96002. doi: 10.1289/EHP1954.

National Toxicology Program (NTP). 2015. "Handbook for Conducting a Literature-Based Health Assessment Using OHAT Approach for Systematic Review and Evidence Integration." Office of Health Assessment and Translation (OHAT). 98p., January 9.

Rhomberg, LR; Goodman, JE; Bailey, LA; Prueitt, RL; Beck, NB; Bevan, C; Honeycutt, M; Kaminski, NE; Paoli, G; Pottenger, LH; Scherer, RW; Wise, KC; Becker, RA. 2013. "A survey of frameworks for best practices in weight-of-evidence analyses." *Crit. Rev. Toxicol.* 43(9):753-784. doi: 10.3109/10408444.2013.832727.

Texas Commission on Environmental Quality (TCEQ). 2017. "TCEQ Guidelines for Systematic Review and Evidence Integration." Toxicology Division. 53p., December 20. Accessed at https://www.tceq.texas.gov/assets/public/implementation/tox/dsd/whitepaper/srguidelines.pdf.

The Cochrane Collaboration. 2011. "Cochrane Handbook for Systematic Reviews of Interventions (Version 5.1.0)." (Eds.: Higgins, JPT; Green, S). March. Accessed at http://training.cochrane-handbook.org

US EPA. 2014. "Integrated Review Plan for the Primary National Ambient Air Quality Standards for Nitrogen Dioxide (External Review Draft)." National Center for Environmental Assessment (NCEA), Office of Air Quality Planning and Standards. 92p., February.

US EPA. 2015. "Preamble to the Integrated Science Assessments." EPA/600/R-15/067. November. Accessed at http://www.epa.gov/isa.

US EPA. 2016. "Integrated Science Assessment for Oxides of Nitrogen – Health Criteria (Final)." National Center for Environmental Assessment (NCEA). EPA/600/R-15/068, 1148p., January.

US EPA. 2018. "Application of Systematic Review in TSCA Risk Evaluations (Final)." Office of Chemical Safety and Pollution Prevention, Office of Pollution Prevention and Toxics. EPA Document # 740-P1-8001. 248p., May. Accessed at https://www.epa.gov/sites/production/files/2018-06/documents/final_application_of_sr_in_tsca_05-31-18.pdf.

Zu, K; Shi, L; Prueitt, RL; Liu, X; Goodman, JE. 2018. "Critical review of long-term ozone exposure and asthma development." *Inhal. Toxicol.* 30(3):99-113. doi: 10.1080/08958378.2018.1455772.

Inhalation Toxicology
International Forum for Respiratory Research

ISSN: 0895-8378 (Print) 1091-7691 (Online) Journal homepage: http://www.tandfonline.com/loi/iiht20

Critical review of long-term ozone exposure and asthma development

Ke Zu, Liuhua Shi, Robyn L. Prueitt, Xiaobin Liu & Julie E. Goodman

To cite this article: Ke Zu, Liuhua Shi, Robyn L. Prueitt, Xiaobin Liu & Julie E. Goodman (2018) Critical review of long-term ozone exposure and asthma development, Inhalation Toxicology, 30:3, 99-113, DOI: 10.1080/08958378.2018.1455772

To link to this article: https://doi.org/10.1080/08958378.2018.1455772

© 2018 Gradient. Informa UK Limited, trading as Taylor & Francis Group

Published online: 05 Jun 2018.

Submit your article to this journal

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
http://www.tandfonline.com/action/journalInformation?journalCode=iiht20

INHALATION TOXICOLOGY
2018, VOL. 30, NO. 3, 99–113
https://doi.org/10.1080/08958378.2018.1455772

Taylor & Francis
Taylor & Francis Group

REVIEW ARTICLE

OPEN ACCESS Check for updates

Critical review of long-term ozone exposure and asthma development

Ke Zu[a], Liuhua Shi[a], Robyn L. Prueitt[b], Xiaobin Liu[a] and Julie E. Goodman[a]

[a]Gradient, Cambridge, MA, USA; [b]Gradient, Seattle, WA, USA

ABSTRACT

Asthma, a chronic respiratory disorder with complex etiology and various phenotypes, is a considerable public health concern in the USA and worldwide. While there is evidence suggesting ambient ozone exposure may exacerbate asthma, information regarding the potential role of ozone in asthma development is more limited. Thus, we conducted a critical review of observational epidemiology studies to determine whether long-term ambient ozone exposure is a risk factor for asthma development. We identified 14 relevant studies; 11 evaluated asthma development in children, while three studies, based on a single cohort, assessed this outcome in adults. Studies of childhood asthma and long-term ozone exposure – including exposure *in utero*, during the first year of life and during early childhood – reported inconsistent findings, which were further weakened by critical methodological limitations in statistical analyses and in exposure and outcome assessments, such as exposure measurement error and a lack of adjustment for key confounders. For adult-onset asthma, long-term ozone exposure was associated with an increased risk in men but not women. In addition to considerable uncertainties due to potential exposure measurement error and a lack of adjustment for key confounders, this study has limited generalizability to the US general population. While experimental evidence indicates that it may be biologically plausible that long-term ozone exposure could contribute to asthma development, it does not provide insight regarding an established mode of action. Future research is needed to address the uncertainties regarding the role of long-term ambient ozone exposure in asthma development.

ARTICLE HISTORY
Received 28 December 2017
Revised 9 March 2018
Accepted 19 March 2018

KEYWORDS
Ozone; asthma development; childhood asthma; adult-onset asthma; long-term exposure; review

Introduction

Asthma is a common chronic disease and a considerable public health concern in the USA, with an estimated prevalence of 7.7% in adults and 9.5% in children between 2008 and 2010 (Moorman et al., 2012). Globally, the prevalence of asthma is increasing in many countries, particularly among children (GINA, 2017). Asthma is a multifactorial, heterogeneous disease with many clinical phenotypes, each of which involves chronic airway inflammation, reversible airway obstruction and airway hyperresponsiveness (AHR) to various triggers, as well as airway remodeling in later stages of the disease (Bates et al., 2009; Currie & Baker, 2012; Grainge & Davies, 2013; Myers & Tomasio, 2011; Noutsios & Floros, 2014; Shin et al., 2009).

The etiology of asthma is complex, and a specific cause has yet to be identified. For childhood asthma, genetics (e.g. parental asthma), respiratory infections, prenatal exposure to environmental tobacco smoke (ETS) and premature birth are well-established risk factors (Bates et al., 2009; Myers & Tomasio, 2011; Noutsios & Floros, 2014; Subbarao et al., 2009); other putative risk factors include allergens, traffic-related air pollution, breastfeeding, family size, and structure and sex (Subbarao et al., 2009). In contrast, the risk of true adult-onset asthma (i.e. not a relapse of childhood asthma in adulthood) may be increased by allergic conditions, obesity, family history of asthma, and exposures to environmental allergens and irritants (Ilmarinen et al., 2015; Subbarao et al., 2009).

Many observational and experimental studies have investigated the potential association between short-term exposure to ozone and outcomes related to asthma severity, such as emergency department visits and hospital admissions for asthma, as well as respiratory symptoms and lung function changes in asthmatics (reviewed by Goodman et al., 2018). The evidence from these studies is suggestive of an association between short-term exposure to ambient concentrations of ozone and asthma severity (Goodman et al., 2018). By contrast, the potential impact of long-term ozone exposure on the development of asthma is unclear. The 2013 Integrated Science Assessment for Ozone and Related Photochemical Oxidants (ISA; US EPA, 2013a) reviewed observational epidemiology studies published through 2011 and concluded that the evidence supported long-term ozone exposure as a risk factor for asthma development in adults and as an effect modifier in children. Here, we conduct a critical review that includes more recently published observational studies that evaluated the association between long-term ambient ozone exposure and incident asthma.

CONTACT Julie E. Goodman jgoodman@gradientcorp.com Gradient, 20 University Road, Cambridge, MA 02138, USA

© 2018 Gradient. Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

Methods

Literature search and study selection

We searched PubMed and Scopus for studies published through October 2017 that evaluated long-term ozone exposure and asthma development using the following terms: ("ozone" OR "air pollution") AND ("asthma development" OR "development of asthma" OR "asthma onset" OR "onset of asthma" OR "newly onset of asthma" OR "incidental asthma" OR "incidence of asthma" OR "asthma incidence" OR "airway hyperresponsiveness" OR "bronchial hyperresponsiveness" OR "wheeze" OR "wheezing") AND ("prospective" OR "cohort" OR "longitudinal" OR "epidemiology" OR "epidemiologic" OR "epidemiological" OR "risk factor" OR "risk factors"). In addition, we cross-referenced the ozone ISA (US EPA, 2013a) and the bibliographies of review articles to identify any studies that were not found in the literature searches.

We included peer-reviewed observational studies of humans that evaluated the associations between long-term ozone exposure (defined as approximately 30 d or longer in duration; US EPA, 2013a) and incident asthma (i.e. incidence, incidence rate or lifetime prevalence of asthma). We excluded studies that met any of our exclusion criteria: laboratory animal or *in vitro* studies; experimental studies; reviews; editorials; commentaries; correspondence/communications; letters to the editor; studies that were not published in English; studies that did not evaluate long-term exposure to ambient ozone (e.g. studies that evaluated short-term exposure or indoor exposure); studies that did not evaluate incident asthma (e.g. studies that evaluated prevalence of asthma or asthma exacerbation); and studies that did not report specific results on the associations between long-term ozone exposure and incident asthma.

Data extraction

From each selected study, we extracted and tabulated information on general and quality characteristics, such as study design, location, study population/cohort, study period, sample size, various features of the exposure assessment (e.g. metric, averaging time, estimation method), outcome ascertainment and statistical approach (e.g. regression model, confounders adjusted for, sensitivity analyses). We also extracted and tabulated detailed study results and graphically presented the study-specific results with the risk estimates standardized to a 1 part per billion (ppb) increment in ozone concentrations.

Study quality evaluation

We developed detailed study quality criteria to evaluate the internal validity of selected studies. These criteria were tailored specifically for long-term ozone exposure and asthma development, and were based on several frameworks, including the Integrated Risk Information System (IRIS) Risk of Bias (RoB) evaluation (US EPA, 2013b, 2014), the National Toxicology Program's (NTP) Office of Health Assessment and Translation (OHAT) approach (NTP, 2015a,b), the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) system (von Elm et al., 2007a–e), the Navigation Guide (Koustas et al., 2013, 2014) and the recent ISAs for oxides of nitrogen (NO_x) and sulfur oxides (SO_x) (US EPA, 2016a,b). We did not conduct a quantitative RoB evaluation (e.g. by categorizing studies into different quality tiers like in the OHAT approach) because it was extremely difficult, if not impossible, to delineate the magnitude and direction of the biases in the available observational studies. We also considered the external validity of the studies in relation to the US general population.

The quality criteria include various methodological attributes, which we applied to individual studies to assess their quality and risk of bias. We first considered the study design: cohort or nested case-control studies had a lower risk of selection bias than case-control studies.

Next, we considered various aspects of the exposure assessment, including spatial and temporal variability, residential mobility, exposure validation, and temporality between exposure and outcome. Studies that employed inverse distance weighting (IDW) or modeling of monitoring data better accounted for spatial variability of the ozone exposure than the studies that only relied on monitoring data. Studies that used time-varying ozone exposure estimates or assessed multiple exposure windows in their analyses better accounted for temporal variability than the studies that only used time-invariant ozone exposure and assessed a single exposure window. Studies that considered residential mobility in the exposure estimates or validated the exposure estimates likely had less exposure measurement error than studies that did neither. Also, studies that clearly established the exposure-outcome temporality (i.e. the exposure occurred before the outcome) had a lower risk of bias than the studies that did not.

We considered two aspects of assessing asthma development: the diagnostic accuracy in the study population and the ascertainment of asthma diagnosis. Because it is difficult to accurately diagnose asthma in small children (e.g. younger than 6 years of age; ATS, 2007; GINA, 2017), studies with participants aged 6 years or older had less potential for outcome misclassification than studies that included younger children. Studies that relied on medical records, hospital discharge records, insurance claims, or adults' reports of physician diagnoses had less potential for outcome misclassification than the studies that used first asthma hospitalization as the surrogate for incident asthma. Asthma hospitalization, a serious exacerbation event, is a poor surrogate for incident asthma because only a small proportion of individuals with asthma are hospitalized (Akinbami et al., 2012). Also, ozone exposure may be associated with asthma severity (Goodman et al., 2018; US EPA, 2013a); therefore, the use of first asthma hospitalization, a severe exacerbation event, may bias the results regarding ozone and asthma development.

We also considered the robustness of the statistical analyses in each study. Studies that employed at least one of the following measures likely yielded more robust results than the studies that did not: used appropriate statistical models

(e.g. logistical regression models for incidence/lifetime prevalence, Cox regression models for incidence rate) with the model assumptions tested (e.g. the proportional hazard assumption for Cox regression models); adjusted for potential confounders identified from the literature including age, sex, race, socioeconomic status (e.g. income, education attainment), maternal age, gestational duration, family history of asthma or allergy, smoking, indoor environment (e.g. mold, dampness, pet ownership), traffic-related pollution (e.g. NO_2, $PM_{2.5}$, distance to major roadway) and other environmental covariates (e.g. surrounding greenness); evaluated non-linear dose-response relationships (e.g. quantiles, splines); and conducted sensitivity analyses.

Last, we evaluated the external validity of the study relative to the US general population. We determined that the results of childhood asthma from western populations (i.e. North American and European cohorts) are more generalizable to the USA, while results from Asian children populations are less so. For studies of adult-onset asthma, if the characteristics of the study population were similar to those of the US general population, we concluded that the findings are generalizable to the USA.

Results

Literature search and study selection

The literature searches resulted in a total of 1279 publications, and we identified 28 potentially relevant epidemiology studies by reviewing titles and abstracts. We did not identify any additional publications from cross-referencing the ozone ISA and review articles. After full-text review, we excluded 14 studies, resulting in 14 studies for review (Figure 1).

Table 1 presents the general characteristics of the selected studies. The majority of the studies were conducted in the USA and Canada in the past two decades. Out of the 14 studies, 11 evaluated asthma development in children, while three studies based on a single cohort, assessed this outcome in adults.

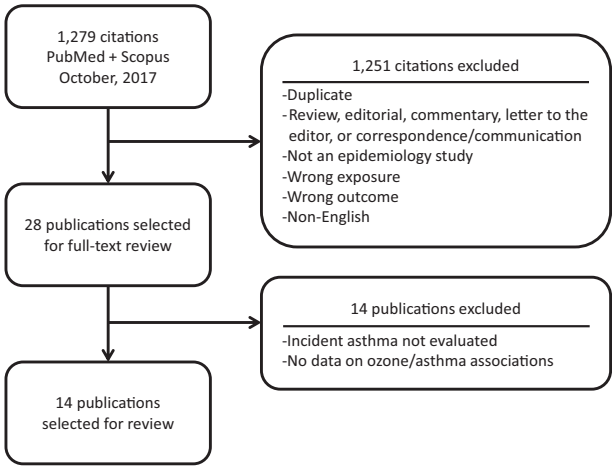

Figure 1. Selection of epidemiology studies for long-term ozone exposure and asthma development.

Study quality evaluation

We evaluated the internal and external validity of individual studies based on criteria we developed (Table 2). Thirteen out of the fourteen studies selected were cohort studies or nested case-control studies, which had a lower risk of selection bias than the one population-based case control study (Nishimura et al., 2013).

Regarding exposure assessment, there were considerable uncertainties in all of the selected studies. Several studies did not account for spatial variability or residential mobility when estimating ozone exposures. The majority of the studies did not account for the temporal variability in ozone exposure or validate the ozone exposure estimates. In addition, a source of considerable uncertainty in most of the studies with retrospective exposure assessment was a lack of temporality between the ozone exposure and the development of asthma (i.e. exposure windows covered time periods after the development of asthma).

Most studies of childhood asthma included children younger than 6 years old, in whom asthma could not be diagnosed accurately. However, these studies relied on medical records, insurance claims and parental reports of physician diagnoses to ascertain a diagnosis of asthma, which reduced the potential for outcome misclassification. Notably, Lin et al. (2008) relied solely on hospital discharge records to identify a child's first hospitalization for asthma as a surrogate for incident asthma, which, as discussed, is a poor proxy for asthma development. All of the studies of asthma development in adults ascertained incident asthma through self-reported physician diagnosis and symptoms, and thus had a low risk for outcome misclassification.

With regard to statistical analyses, most studies employed suitable statistical models and tested model assumptions when appropriate, except for those conducted by McConnell et al. (2010) and Islam et al. (2008), who did not test the proportional hazard assumption. Half of the 14 studies assessed potentially non-linear dose-response relationships by using quantiles or splines of ozone exposure in their analyses, and most conducted some form of sensitivity analysis. We also found that, while the vast majority of the 14 studies adjusted for confounders, such as age, sex, socioeconomic status and smoking, several studies did not account for race, maternal age, gestational duration, or family history of asthma or allergy, and few studies considered the potential confounding effects of indoor environment, traffic-related pollution or other environmental factors (e.g. surrounding greenness).

Regarding external validity, the majority of the studies were conducted in the USA, Canada and Europe, so their findings are more generalizable to the US population, compared to the three studies in Asia (Dong et al., 2013; Kim et al., 2013; Wang et al., 2016).

Overall, while findings from most of the selected studies are generalizable to the US population, limitations of ozone exposure assessment and the lack of adjustment for several key environmental confounders may impact the interpretation of study results.

Table 1. General characteristics of epidemiology studies of long-term ozone exposure and asthma development.

			Study population[a]						Ozone				
Study	Location	Study period	Cohort	Population	Sample size	# of cases	# of non-cases	Age (years, mean ± SD)	Exposure metric	Exposure window	Averaging period	Spatial scale	Estimation method
Childhood asthma													
Sbihi et al. (2016)	Canada	1999–2009	Vancouver birth cohort	Children	51,857	8659	43,198	Preschool: 2.6 ± 1.4 School age: 7.0 ± 1.2	NR	*In utero*	Months	Postal code level	Monitoring data
Tetreault et al. (2016)	Canada	1996–2011	Quebec birth cohort	Children	829,277	NR	NR	60% of follow-up occurred before the age of 6 years	NR	Early life	Months (June–August)	Postal code level	Monitoring data and land-use mixed-effects model
Wang et al. (2016)	Taipei	2010	CEAS cohort	Children	2661	336	2325	5.5 ± 1.1	8 h average, hourly	Early life	Years	District level	Monitoring data and distance to monitoring station
MacIntyre et al. (2014)	Canada, Europe	Birth to 8 years old	CAPPS, LISA, GINI, PIAMA	Children	2671	534	2137	NR	NR	First year of life	Annual	Continuous spatial surface (residence level)	Monitoring data, IDW, and APMoSPHERE modeling
Dong et al. (2013)	China	2008–2009	SNEC cohort	Children	30,056	1963	28,093	8.4 ± 2.7 Range: 2–14	8 h average (daytime)	Early life	Years	District level	Monitoring data
Kim et al. (2013)	Korea	2005–2008	CHEER cohort	Children	369	NR	NR	At enrollment: 6.83 ± 0.52	NR	Early life	Years	Continuous spatial surface (residence level)	Monitoring data and ArcGIS Modeling
Nishimura et al. (2013)	USA	2006–2011	GALA II and SAGE II	Children	4320	2291	2029	Range: 8–21	8 h max, 1 h max	Early life	Years	Continuous spatial surface (residence level)	Monitoring data and IDW
Clark et al. (2010)	Canada	1999–2003	British Columbia birth cohort	Children	37,401	3482	33,919	At end of follow-up: 4 ± 0.6 range: 3–4.9	24 h average	*In utero* and first year of life	Months	Residential six-digit postal code	Monitoring data and IDW
McConnell et al. (2010)	USA	2002–2006	CHS cohort	Children	2497	120	2377	At enrollment: range: 4.8–9.0	8 h average (10 am–6 pm)	Early life	Years	Community level	Monitoring data
Islam et al. (2008)	USA	1993–2004	CHS cohort	Children	1701	160	1541	At enrollment: ≥ 7	8 h average (10 am–6 pm)	Early life	Years	Community level	Monitoring data
Lin et al. (2008)	USA	1995–2000	NYS birth cohort	Children	1,204,396	10,429	1,193,967	Range: 1–6	1 h max (10 am–6 pm)	Early life	Years	Region level	Monitoring data
Adult-onset asthma													
McDonnell et al. (1999)	USA	1977–1992	SDA cohort	Adults	2758	111	2647	Range: 27–87	8 h average (9 am–5 pm), hourly	Adult life	Years	ZIP code level	Monitoring data and interpolation
Greer et al. (1993)	USA	1977–1987	SDA cohort	Adults	3577	78	3499	Range: 26–95	Hourly	Adult life	Years	ZIP code level	Monitoring data and interpolation
Abbey et al. (1991)	USA	1977–1987	SDA cohort	Adults	6058	80	5978	NR	Hourly	Adult life	Years	ZIP code level	Monitoring data and interpolation

APMoSPHERE: Air Pollution Modeling for Support to Policy on Health and Environmental Risks in Europe; IDW: inverse distance weighting; NR: not reported; SD: standard deviation
[a]The characteristics of the study population are specific to the ozone-asthma analyses.

Table 2. Quality and generalizability of epidemiology studies of long-term ozone exposure and asthma development.

		Exposure assessment					Outcome assessment		Statistical analysis														
										Confounders considered													
Study	Study design	Spatial variability considered	Temporal variability considered	Residential mobility considered	Exposure estimates validated	Exposure-outcome temporality established	Physician diagnostic accuracy	Ascertainment	Appropriate statistical model employed	Age	Sex	Race	SES	Maternal age	Gestational duration	Family history	Smoking	Indoor environment	Traffic-related pollution	Other environmental covariates	Non-linear dose-response relationship assessed	Sensitivity analysis conducted	Generalizability to US population
Childhood asthma																							
Sbihi et al. (2016)	Nested case-control	Y	N	Y	N	Y	High[a]	Physician billing, hospital discharge records	Y	Y	Y	N	Y	Y	Y	N	N	N	N	Y	Y	Y	High
Tetreault et al. (2016)	Cohort	Y	Y	Y	N	N	Low	Hospital discharge, two physician claims	Y	Y	Y	N	Y	N	N	N	Y	N	N	N	Y	Y	High
Wang et al. (2016)	Cohort	Y	N	N	N	N	Low	Parental report of physician diagnosis	Y	Y	Y	Y	Y	N	N	Y	Y	Y	N	Y	Y	N	Low[b]
MacIntyre et al. (2014)	Cohort	Y	N	N	N	N	Low	Parental report of physician diagnosis	Y	N	Y	N	N	Y	N	Y	Y	N	N	N	N	N	High
Dong et al. (2013)	Cohort	N	N	N	N	N	Low	Parental report of physician diagnosis	Y	Y	Y	N	Y	N	N	Y	Y	N	N	N	N	N	Low[b]
Kim et al. (2013)	Cohort	Y	N	Y	Y	N	Low	Parental report of physician diagnosis	Y	Y	Y	N	Y	N	N	Y	Y	N	Y	N	N	N	Low[b]
Nishimura et al. (2013)	Population-based case-control	Y	Y	Y	N	Y	High[a]	Physician diagnosis	Y	Y	Y	Y	Y	N	N	Y	Y	N	N	N	N	Y	High
Clark et al. (2010)	Nested case-control	Y	Y	Y	N	N	Low	Outpatient and hospital records	Y	Y	Y	N	Y	Y	Y	N	Y	N	N	N	N	Y	High
McConnell et al. (2010)	Cohort	N	N	N	N	N	Low	Parental report of physician diagnosis	N[c]	Y	Y	Y	Y	N	N	Y	Y	Y	Y	N	Y	Y	High
Islam et al. (2008)	Cohort	N	N	Y	N	N	High[a]	Parental report of physician diagnosis	N[c]	Y	Y	Y	Y	N	N	Y	Y	Y	N	N	N	N	High
Lin et al. (2008)	Cohort	N	Y	Y	N	Y	Low	Hospital discharge records[d]	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	High
Adult-onset asthma																							
McDonnell et al. (1999)	Cohort	Y	N	Y	Y	N	High[a]	Self-reported physician diagnosis	Y	Y	Y	N	Y	NA	NA	N	Y	Y	Y	N	Y	Y	Low[e]
Greer et al. (1993)	Cohort	Y	N	Y	Y	N	High[a]	Self-reported physician diagnosis and symptoms, medical records	Y	Y	Y	N	Y	NA	NA	N	Y	N	N	N	N	Y	Low[e]
Abbey et al. (1991)	Cohort	Y	N	Y	Y	Y	High[a]	Self-reported physician diagnosis and symptoms	Y	Y	Y	N	Y	NA	NA	N	Y	N	N	N	N	N	Low[e]

SES: socioeconomic status

Grey shading indicates a higher potential for error and/or bias.

[a]The study population did not include children younger than 6 years old, for whom physicians' diagnoses of asthma may be inaccurate.

[b]The study population is Asian, so the study results may not be generalizable to the US population.

[c]The proportional hazard assumption was not tested for the Cox regression model used.

[d]Asthma hospitalization was ascertained, but this is a poor surrogate for incident asthma.

[e]The Seventh-day Adventists have many lifestyle characteristics that differ from the general US population.

Table 3. Detailed results of epidemiology studies of long-term ozone exposure and asthma development.

Study	Stratum		Ozone			Asthma	Statistical models	Results				
			Metric	Distribution	Categories			Effect measure	Effect estimate	95% CI	*P* value for point estimate	*P* value for interaction
Childhood Asthma												
Sbihi et al. (2016)	Preschool	All study subjects	*In utero* average	Mean ± SD (IQR): cases: 27.7 ± 5.8 (8.6) μg/m^3 controls: 27.9 ± 5.8 (8.5) μg/m^3	Continuous	Incidence	Conditional logistic regression models	OR per IQR (8.5 μg/m^3) increase	**0.92**	**(0.87,0.97)**	*P* < .05	NA
					Quartile 1			OR	1		NR	
					Quartile 2				**0.88**	**(0.81,0.95)**		
					Quartile 3				0.95	(0.87,1.03)		
					Quartile 4				**0.85**	**(0.77,0.94)**		
		Gestational duration ≥ 37 weeks			Continuous			OR per IQR (8.5 μg/m^3) increase	**0.91**	**(0.86,0.96)**	NR	NR
		Gestational duration < 37 weeks							1.01	(0.60,1.68)		
		Birth weight ≥ 2,500 g			Continuous			OR per IQR (8.5 μg/m^3) increase	**0.9**	**(0.86,0.95)**	NR	NR
		Birth weight < 2,500 g							0.59	(0.28,1.19)		
	School age	All study subjects	*In utero* average	cases: 27.9 ± 6.0 (8.7) μg/m^3 controls: 28.3 ± 5.9 (8.6) μg/m^3	Continuous	Incidence	Conditional logistic regression models	OR per IQR (8.5 μg/m^3) increase	**1.18**	**(1.07,1.31)**	*P* < .05	NA
					Quartile 1			OR	1		NR	
					Quartile 2				1.02	(0.87,1.20)		
					Quartile 3				**1.19**	**(1.01,1.41)**		
					Quartile 4				**1.22**	**(1.01,1.48)**		
Tetreault et al. (2016)	Using ozone estimates at the birth address		Annual average	Mean: 32.07 ppb Range: 12.19 – 43.12 ppb IQR: 3.22 ppb	Continuous	Incidence rate	Cox proportional hazard models	HR per IQR (3.22 ppb) increase	**1.06**	**(1.05,1.07)**	*P* < .001	NA
	Using ozone estimates for full residential history		Annual average	Mean: 31.97 ppb Range: 12.19 – 43.39 ppb IQR: 3.26 ppb		Incidence rate	Cox proportional hazard models	HR per IQR (3.26 ppb) increase	**1.07**	**(1.06, 1.08)**	*P* < .001	NA
Wang et al. (2016)	Full cohort		Long-term average: Hourly	Mean ± SD: 27.50 ± 0.61 ppb Median: 27.62 ppb	Low	Lifetime prevalence	Logistic regression models	OR	1		NR	NA
					High				**0.68**	**(0.51, 0.92)**	*P* < .05	
			Long-term average: 8h mean		Low				1		NR	
					High				0.79	(0.59, 1.07)	*P* > .05	
MacIntyre et al. (2014)	Full cohort		Annual average	Mean ± SD: 36.8 ± 6.5 μg/m^3 IQR: 9.1 μg/m^3 Range: 13.3 – 55.9 μg/m^3	Continuous	Lifetime prevalence	Logistic regression models	OR per 10 μg/m^3 increase	0.86	(0.66, 1.13)	*P* = .277	NA
	GSTP1 rs1138272 TT/TC								1.13	(0.57, 2.21)	NR	NR
	GSTP1 rs1138272 CC								0.84	(0.62, 1.12)	NR	
	GSTP1 rs1695 GG/GA								0.89	(0.62, 1.28)	NR	
	GSTP1 rs1695 AA								0.73	(0.48, 1.12)	NR	
	TNF rs1800629 AA/AG								0.77	(0.49, 1.19)	NR	NR
	TNF rs1800629 GG								0.91	(0.64, 1.29)	NR	
Dong et al. (2013)	Normal weight		3-year average	Mean: 54.8 μg/m3 Range: 34 – 89 μg/m^3	Continuous	Lifetime prevalence	Logistic regression models	OR per IQR (23 μg/m^3) increase	**1.25**	**(1.12, 1.40)**	NR	*p* = .167
	Overweight								**1.53**	**(1.26, 1.86)**	NR	
	Obesity								**1.31**	**(1.08, 1.58)**	NR	
Kim et al. (2013)	Full cohort		5-year average	Mean: 10.32 ppb Range: 3.74 – 29.66 ppb	Continuous	Lifetime prevalence	Logistic regression models	OR per 5 ppb increase	1.04	(0.83, 1.30)	NR	NA
						Incidence			0.94	(0.61, 1.46)		

(*continued*)

Table 3. Continued

Study	Stratum	Ozone				Asthma	Statistical models	Results				
		Metric		Distribution	Categories			Effect measure	Effect estimate	95% CI	*P* value for point estimate	*P* value for interaction
Nishimura et al. (2013)	All subjects	1-year average	1h max	Mean ± SD: 1-h max average in 1996 34.3 ± 7.7 ppb 8-h max average in 1996 27.6 ± 6.6 ppb	Continuous	Lifetime prevalence	Two stage analysis: Logistic regression models and random-effects models	OR per 5 ppb increase	0.98	(0.89, 1.08)	NR	NA
			8h max						0.96	(0.87, 1.06)		
		3-year average	1h max						0.98	(0.84, 1.14)	NR	
			8h max						0.95	(0.81, 1.12)		
	Boys	1-year average	1h max		Continuous	Lifetime prevalence	Two stage analysis: Logistic regression models and random-effects models	OR per 5 ppb increase	0.98	(0.87, 1.10)	NR	NR
			8h max						0.96	(0.85, 1.09)		
		3-year average	1h max						0.96	(0.78, 1.19)	NR	
			8h max						0.93	(0.76, 1.13)		
	Girls	1-year average	1h max		Continuous	Lifetime prevalence	Two stage analysis: Logistic regression models and random-effects models	OR per 5 ppb increase	0.95	(0.86, 1.05)	NR	
			8h max						0.93	(0.82, 1.04)		
		3-year average	1h max						0.96	(0.79, 1.18)	NR	
			8h max						0.94	(0.74, 1.19)		
	Total IgE > 200	1-year average	1h max		Continuous	Lifetime prevalence	Two stage analysis: Logistic regression models and random-effects models	OR per 5 ppb increase	0.9	(0.77, 1.06)	NR	NR
			8h max						0.9	(0.72, 1.12)		
		3-year average	1h max						0.87	(0.75, 1.02)	NR	
			8h max						0.85	(0.69, 1.04)		
	Total IgE ≤ 200	1-year average	1h max		Continuous	Lifetime prevalence	Two stage analysis: Logistic regression models and random-effects models	OR per 5 ppb increase	1.03	(0.92, 1.15)	NR	
			8h max						1.01	(0.90, 1.15)		
		3-year average	1h max						1.04	(0.82, 1.32)	NR	
			8h max						1.01	(0.77, 1.32)		
	Family history	1-year average	1h max		Continuous	Lifetime prevalence	Two stage analysis: Logistic regression models and random-effects models	OR per 5 ppb increase	0.96	(0.76, 1.21)	NR	NR
			8h max						0.97	(0.73, 1.29)		
		3-year average	1h max						0.92	(0.69, 1.21)	NR	
			8h max						0.9	(0.69, 1.18)		
	No family history	1-year average	1h max		Continuous	Lifetime prevalence	Two stage analysis: Logistic regression models and random-effects models	OR per 5 ppb increase	0.96	(0.82, 1.11)	NR	
			8h max						0.94	(0.79, 1.11)		
		3-year average	1h max						0.95	(0.76, 1.19)	NR	
			8h max						0.92	(0.73, 1.17)		
Clark et al. (2010)	Full cohort	*In utero* average		Mean ± SD: Asthma cases in utero 30.05 ± 6.39 ppb Controls *in utero* 30.48 ± 6.32 ppb Asthma cases in 1st year 27.64 ± 4.94 ppb Controls in 1st year 28.06 ± 4.86 ppb	Continuous	Incidence	Conditional logistic regression models	OR per 10 μg/m^3 increase	**0.83**	**(0.77, 0.89)**	NR	NA
		1-year average							**0.81**	**(0.74, 0.87)**		
	Boys	*In utero* average			Continuous	Incidence	Conditional logistic regression models	OR per 10 μg/m^3 increase	**0.86**	**(0.78, 0.94)**	NR	NR
		1-year average							**0.84**	**(0.76, 0.94)**		
	Girls	*In utero* average			Continuous	Incidence	Conditional logistic regression models	OR per 10 μg/m^3 increase	**0.79**	**(0.70, 0.89)**	NR	
		1-year average							**0.74**	**(0.64, 0.84)**		
McConnell et al. (2010)	Full cohort	2-year average		Mean: 44.6 ppb	Continuous	Incidence rate	Cox proportional hazard models	HR across the range of ozone (30.3 ppb)	1.01	(0.49, 2.11)	NR	NA

(*continued*)

Table 3. Continued

Study	Stratum	Ozone			Asthma	Statistical models	Results				
		Metric	Distribution	Categories			Effect measure	Effect estimate	95% CI	*P* value for point estimate	*P* value for interaction
Islam et al. (2008)	No short-allele	Long-term average	Mean (range): Higher ozone communities: 55.2 (46.5, 64.9) ppb Lower ozone communities: 38.4 (28.6, 45.5) ppb	Low	Incidence rate	Cox proportional hazard models	HR	1			$p = .003$
				High				0.94	(0.36, 2.43)	NR	
	Short-allele			Low	Incidence rate	Cox proportional hazard models	HR	1			
				High				2	(0.40, 10.17)	NR	
Lin et al. (2008)	Full cohort	Long-term average	Mean: 41.06 ppb	Continuous	Incidence	Logistic regression models	OR per 1 ppb increase	**1.16**	**(1.15, 1.17)**	NR	NA
		Long-term ozone-season average	Mean: 50.62 ppb				OR per 1 ppb increase	**1.22**	**(1.21, 1.23)**	NR	
		Exceedance proportion (%) > 70 ppb	Mean: 9.72% Range: 1.66–26.27%				OR per IQR (2.51%) increase	**1.68**	**(1.64, 1.73)**	NR	
	New York City	Long-term average	Mean: 37.51 ppb	Low	Incidence	Logistic regression models	OR	1		NR	NR
				Medium				**1.43**	**(1.29, 1.58)**	NR	
				High				**1.69**	**(1.52, 1.80)**	NR	
	Other NYS cities	Long-term average	NR	Low	Incidence	Logistic regression models	OR	1		NR	
				Medium				**1.64**	**(1.48, 1.82)**	NR	
				High				**2.06**	**(1.87, 2.27)**	NR	
Adult-onset Asthma											
McConnell et al. (1999)	Men	Long-term average, 8h mean	Mean: 46.5 ppb Range: 0–74.9 ppb	Tertile 1	Incidence	Logistic regression models	OR	1		NR	NR
				Tertile 2				**4.44**	**(1.32, 13.81)**	NR	
				Tertile 3				**4.01**	**(1.15, 13.00)**	NR	
				Continuous			OR per IQR (27 ppb) increase	**2.09**	**(1.03, 4.16)**	$P < .05$	
		Long-term average, Hourly		Continuous			OR per IQR (11.7 ppb) increase	**2.46**	**(1.28, 4.67)**	$P < .05$	
		Exceedance frequency (# of hours) > 60 ppb					OR per IQR (907 h) increase	**1.90**	**(0.99, 3.64)**	$P < .05$	
		Exceedance frequency (# of hours) > 80, 100, 120, or 150 ppb					OR per IQR increase	NR	NR	$P > .05$	
	Women	Long-term average, 8h mean		Continuous	Incidence	Logistic regression models	OR per IQR (27 ppb) increase	0.86	(0.58, 1.26)	$P > .05$	
		Long-term average, Hourly					OR per IQR (11.7 ppb) increase	NR	NR	$P > .05$	
		Exceedance proportion (%) > 60, 80, 100, 120, or 150 ppb					OR per IQR increase	NR	NR	$P > .05$	
Greer et al. (1993)	Full cohort	Long-term average	NR	Continuous	Incidence	Logistic regression models	OR per 10 ppb increase	1.31	(0.96, 1.78)	NR	NA
	Men							**3.12**	**(1.61, 5.85)**	NR	NR
	Women							0.94	(0.65, 1.34)	NR	
Abbey et al. (1991)	Full cohort	Exceedance frequency (# of hours) > 100 ppb	NR	Continuous	Incidence	Logistic regression models	OR per 500 hours increase	1.40	(0.90, 2.34)	0.056	NA

CI: confidence interval; GSTP: glutathione S-transferase P; HR: hazard ratio; IgE: immunoglobulin E; IQR: interquartile range; NA: not applicable; NR: not reported; OR: odds ratio; ppb: parts per billion; SD: standard deviation; TNF: tumor necrosis factor

Bolded fonts indicate statistically significant results.

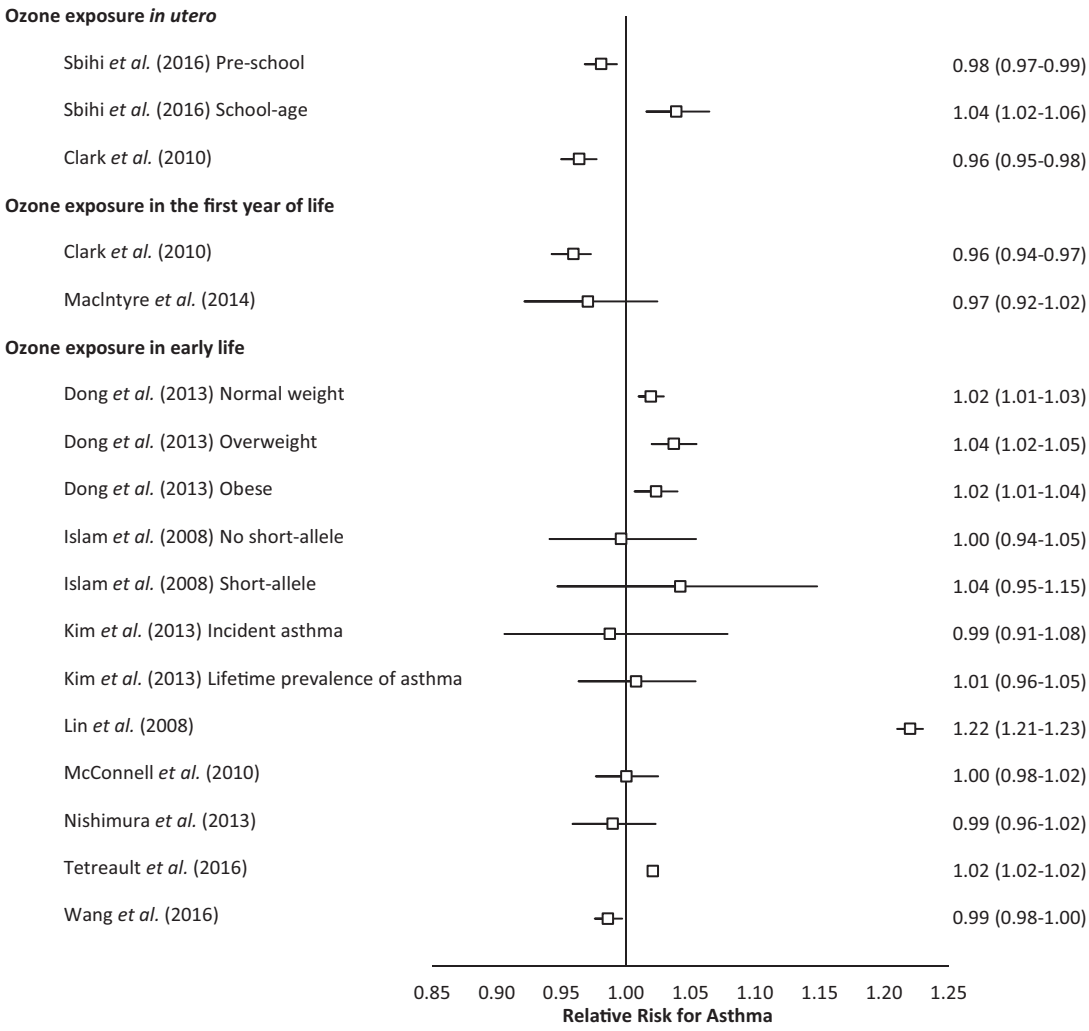

Figure 2. Study-specific results of long-term ozone exposure and childhood asthma development, by exposure window – the relative risks and their 95% confidence intervals have been standardized to a 1 ppb increment in ozone exposure.

Study results evaluation

We identified 11 studies from 10 non-overlapping populations that evaluated childhood asthma development, and three studies from one cohort that assessed asthma development in adults. As discussed above, all of these studies had considerable uncertainties and potential for bias with regard to exposure and outcome assessments and confounder adjustment. In light of the major methodological limitations in these studies, although several studies conducted some form of sensitivity analysis, we mainly focused on the study findings from the primary and stratified analyses with a focus on the within- and between-study consistency. Detailed study results are presented in Table 3.

Childhood asthma

To aid the evaluation of study results regarding long-term ozone exposure and childhood asthma development, we graphically present the main findings from these studies in Figure 2, with the risk estimates standardized to a 1 ppb increment in ozone exposure and grouped by exposure window.

***Ozone exposure* in utero.** Two studies, both of which were birth cohorts in Canada, evaluated the association between *in utero* ozone exposure and childhood asthma (Clark et al., 2010; Sbihi et al., 2016). Sbihi et al. (2016) reported that higher *in utero* ozone exposure was statistically significantly associated with decreased incident asthma in preschool-age children (OR = 0.92, 95% CI: 0.87–0.97). When ozone exposure was assessed as quartiles, the top three quartiles were all associated with reduced asthma risks (quartile 2: OR = 0.88, 95% CI: 0.81–0.95; quartile 3: OR = 0.95, 95% CI: 0.87–1.03; quartile 4: OR = 0.85, 95% CI: 0.77–0.94), compared to the bottom quartile. This statistically significant inverse association persisted among children with a gestational age of 37 weeks or longer (OR = 0.91, 95% CI: 0.86–0.96) or with a birth weight of 2500 g or more (OR = 0.9, 95% CI: 0.86–0.95) but was null for children with a gestational age less than 37 weeks or with a birth weight less than 2500 g. Clark et al. (2010), who assessed incident asthma only in children aged younger than 5 years, also reported a statistically significant decrease in asthma development in association with increases in *in utero* ozone exposure (OR = 0.83, 95% CI: 0.77–0.89). The observed statistically significant inverse association persisted when the

study population was stratified by sex (boys: OR = 0.86, 95% CI: 0.78–0.94; girls: OR = 0.79, 95% CI: 0.70–0.89).

In contrast, Sbihi et al. (2016) found that higher *in utero* ozone exposure was associated with a statistically significant increase in incident asthma in school-age children (OR = 1.18, 95% CI: 1.07–1.31). The analyses with quartiles of ozone exposure showed that the risk of asthma was only increased in the top two quartiles, compared to the bottom quartile (quartile 2: OR = 1.02, 95% CI: 0.87–1.20; quartile 3: OR = 1.19, 95% CI: 1.01–1.41; quartile 4: OR = 1.22, 95% CI: 1.01–1.48).

Ozone exposure in the first year of life. Two studies assessed the association between ozone exposure in the first year of life and childhood asthma (Clark et al., 2010; MacIntyre et al., 2014). MacIntyre et al. (2014), who conducted a pooled analysis of four birth cohorts in Canada and Europe, reported a null association between ozone exposure in the first year of life and the risk of ever being diagnosed with asthma up to 8 years of age. MacIntyre et al. (2014) also evaluated the potential effect modification by several single-nucleotide polymorphisms (SNPs) in the glutathione S-transferase P1 (GSTP1) and tumor necrosis factor (TNF) genes and found that the null associations between ozone exposure and asthma did not vary by any of the SNPs examined. Similar to *in utero* ozone exposure, Clark et al. (2010) observed a statistically significant inverse association between ozone exposure in the first year of life and incident asthma (full cohort: OR = 0.81, 95% CI: 0.74–0.87; boys: OR = 0.84, 95% CI: 0.76–0.94; girls: OR = 0.74, 95% CI: 0.64–0.84).

Ozone exposure in early life. Eight studies investigated the relationship between ozone exposure in early life and risk of developing asthma, and there is considerable between-study heterogeneity in the direction of observed associations.

Four studies, from three non-overlapping populations in the US and Korea, observed a null association between long-term ozone exposure in early life and risk of developing childhood asthma (Islam et al., 2008; Kim et al., 2013; McConnell et al., 2010; Nishimura et al., 2013), and the null associations did not vary by sex, family history of asthma, total IgE level, functional polymorphisms tested, the metric and averaging time of ozone exposure or measures of asthma development (Islam et al., 2008; Kim et al., 2013; Nishimura et al., 2013).

Conversely, Wang et al. (2016), a retrospective cohort study in Taipei, reported a statistically significant decrease in risk of asthma associated with higher ozone exposure based on hourly measurements (OR = 0.68, 95% CI: 0.51–0.92). The association was no longer statistically significant when the 8 h mean ozone concentrations were used (OR = 0.79, 95% CI: 0.59, 1.07).

In addition, three studies reported statistically significant positive associations between the risk of childhood asthma and ozone exposure. Tetreault et al. (2016), in a birth cohort study in Canada, found that higher ozone exposure during summer months in early life was associated with increased risk of developing childhood asthma (HR = 1.07, 95% CI: 1.06–1.08). Tetreault et al. (2016) also evaluated the dose-response relationship using cubic splines of continuous ozone exposure but did not find any evidence for non-linearity. Dong et al. (2013), who conducted a population-based cohort study in China, observed statistically significant increases in asthma development risk associated with increases in ozone exposure, and this positive association was not modified by body weight (normal weight: OR = 1.25, 95% CI: 1.12–1.40; overweight: OR = 1.53, 95% CI: 1.26–1.86; obese: OR = 1.31, 95% CI: 1.08–1.58). Lin et al. (2008), in a cohort study in New York State, reported that higher ozone exposure in early life was associated with a statistically significantly increased risk of asthma hospitalization (OR = 1.16, 95% CI: 1.15–1.17). The results were consistent when only exposure in ozone season (April–October) was considered (OR = 1.22, 95% CI: 1.21–1.23) or when the annual percentage of hourly ozone measurements exceeding 70 ppb was considered (OR = 1.68, 95% CI: 1.64, 1.73). These increases persisted in the top two tertiles of ozone exposure when the study population was stratified by location (New York City: tertile 2: OR = 1.43, 95% CI: 1.29–1.58; tertile 3: OR = 1.69, 95% CI: 1.52–1.80 versus other New York State cities: tertile 2: OR = 1.64, 95% CI: 1.48–1.82; tertile 3: OR = 2.06, 95% CI: 1.87–2.27).

Adult-onset asthma

Three studies, all based on the Seventh-day Adventists (SDA) Cohort in California, assessed the risk of developing adult-onset asthma in association with long-term ozone exposure (Abbey et al., 1991; Greer et al., 1993; McDonnell et al., 1999).

For 6058 Seventh-day Adventists with a follow-up of 10 years, Abbey et al. (1991) reported that higher ozone exceedance frequency (i.e. annual average of greater than 500 hourly ozone measurements above 100 ppb) was associated with a borderline significant increase in asthma risk (OR = 1.40, 95% CI: 0.90–2.34). Greer et al. (1993) restricted their analyses to 3577 nonsmokers, and also reported a borderline significant increase in asthma risk associated with a 10 ppb increment in ozone exposure (OR = 1.31, 95% CI: 0.96–1.78). The positive association between ozone and asthma was more pronounced in men (OR = 3.12, 95% CI: 1.61–5.85) but was null in women (OR = 0.94, 95% CI: 0.65–1.34). McDonnell et al. (1999), who analyzed 2758 nonsmokers with a follow-up of 15 years, also reported a statistically significant association between 20-year cumulative ozone exposure (based on daily 8 h mean concentrations) and asthma in men (OR = 2.09, 95% CI: 1.03–4.16) but a null association in women (OR = 0.86, 95% CI: 0.58–1.26). The results did not change meaningfully when daily hourly ozone measurements were used to calculate the 20-year cumulative ozone exposure. In addition, McDonnell et al. (1999) assessed ozone exceedance frequencies above various thresholds and found that higher exceedance frequencies above 60 ppb were associated with statistically significant increases in asthma risk among men (OR = 1.90, 95% CI: 0.99–3.64) and that the statistical significance did

not persist with higher ozone thresholds (e.g. 80, 100, 120 and 150 ppb). Ozone exceedance frequencies were not associated with asthma risk in women. When tertiles of ozone exposure were considered, men in the top two tertiles had a statistically significantly increased risk of asthma, compared to men in the bottom tertile (tertile 2: OR = 4.44, 95% CI: 1.32–13.81; tertile 3: OR = 4.01, 95% CI: 1.15–13.00).

Discussion

Overall, there is substantial between-study heterogeneity in the direction of observed associations between long-term ozone exposure and asthma development. In addition, available epidemiology studies have considerable methodological limitations, including potential exposure measurement error and a lack of adjustment for several key confounders, which further complicates the interpretation of the findings.

Childhood asthma

The observed associations between ozone exposure and development of childhood asthma varied by exposure window assessed and the age of the study population.

Ozone exposure in utero

Consistent findings from two studies (Clark et al., 2010; Sbihi et al., 2016) suggested that higher *in utero* ozone exposure was associated with reduced risk of asthma in preschool-age children. However, the effect estimates for ozone, though statistically significant, were small in magnitude. The observed dose-response relationships appeared to be linear, without any statistical evidence for non-linearity. Notably, in these two studies, ozone exposure was inversely correlated with traffic-related pollutants, a risk factor for childhood asthma. The inverse correlation between ozone and traffic-related pollutants might contribute to, at least partially, the observed inverse association between ozone exposure and asthma in preschool-age children. In addition, the diagnostic accuracy in children aged 6 years or younger (i.e. preschool) is low, and neither of the studies relied on validated ozone exposure estimates or adjusted for key confounders such as a family history of asthma (Table 2).

In contrast, Sbihi et al. (2016) reported that higher *in utero* ozone exposure was linearly associated with a small, but statistically significant, increase in asthma risk among school-age children. The observed association in school-age children was the direct opposite of that in preschool-age children in the same study, which is unlikely to be entirely explained by the better diagnostic accuracy of asthma in school-age children. Moreover, the methodological limitations of this study discussed above also applied to the analyses of school-age children (Table 2).

Ozone exposure in the first year of life

Clark et al. (2010) reported decreased risks of asthma development associated with higher ozone exposure in the first year of life, while MacIntyre et al. (2014) observed a null association. The magnitude of the effect estimates, based on a linear dose-response relationship, was small. Both studies suffered from a high risk of bias from the exposure and outcome assessments, as well as from inadequate adjustment for confounders (Table 2).

Ozone exposure in early life

The most examined exposure window in available epidemiology studies was early life, but the results from these analyses were inconsistent with regard to the direction of observed associations. Most of the studies assumed a linear dose-response relationship between ozone and asthma development and reported ozone effect estimates that were small in magnitude. The magnitude of the ozone effect estimates observed in Lin et al. (2008) was considerably larger than that in other studies (Figure 2). As discussed above, Lin et al. (2008) assessed first asthma hospitalization, which indicates severe exacerbation and is a poor surrogate for incident asthma. Because ozone exposure may affect asthma severity (Goodman et al., 2018; US EPA, 2013a), the use of first asthma hospitalization in Lin et al. (2008) likely biased the results away from the null.

In addition, considerable uncertainty stemmed from the lack of ozone-asthma temporality in most of these studies, with the exception of Nishimura et al. (2013) and Lin et al. (2008). The ozone exposure periods evaluated overlapped with the follow-up periods for asthma; therefore, for most, if not all, of the children who developed asthma, part of their estimated ozone exposures occurred after the outcome occurred. Because of the continuing downward trend of ambient ozone concentration in the USA between 1998 and 2010 (US EPA, 2013a), this lack of temporality might have led to an overestimation of the association between early life ozone exposure and asthma development.

Similar to studies that evaluated ozone exposures *in utero* and in the first year of life, studies of early life ozone exposure were also limited by considerable measurement error and uncertainties in exposure and outcome assessments (Table 2). Several studies did not adjust for key confounders, such as a family history of asthma, and most did not control for preterm birth, exposure to traffic-related air pollutants or the indoor environment (Table 2). Findings from Asian children populations were less generalizable to the USA than those from the Western populations (Table 2).

Because the natural and anthropogenic sources of ozone precursors differ in urban and non-urban areas (US EPA, 2013a), it is conceivable that urbanicity could modify potential associations between long-term ambient ozone concentrations and asthma development. However, more than half of the studies of childhood asthma were conducted only in urban areas (Clark et al., 2010; Dong et al., 2013; Kim et al., 2013; Nishimura et al., 2013; Sbihi et al., 2016; Wang et al., 2016), while the rest of the studies included children living in both urban and non-urban areas (Islam et al., 2008; Lin et al., 2008; MacIntyre et al., 2014; McConnell et al., 2010; Tetreault et al., 2016). There were no apparent patterns in the observed associations between ozone and childhood asthma development with regard to whether the studies

were conducted only in urban areas or in urban and non-urban areas combined. Because there were no studies conducted only in non-urban areas, we were unable to examine the association between long-term ozone exposure and childhood asthma development in rural settings.

Adult-onset asthma

Adult-onset asthma was only evaluated in one population, the California Seventh-day Adventists cohort (Abbey et al., 1991; Greer et al., 1993; McDonnell et al., 1999). Long-term ozone exposure was associated with increased asthma risk in men but not women (Greer et al., 1993; McDonnell et al., 1999). The magnitude of the observed association was large, with the ORs greater than 2. The increased risk in men was persistent when ozone exposure was analyzed linearly, in tertiles, or with exceedance frequencies above various thresholds.

The ozone ISA relied on these studies to conclude that long-term ozone exposure is a risk factor for adult-onset asthma. However, there does not seem to be a biologically plausible reason for the different associations between men and women. Men, on average, had more severe asthma, and thus might have been more likely to get diagnosed with asthma than women in this cohort. Because ozone exposure may be associated with asthma severity, the observed association between ozone and asthma development in men could be partially attributable to the difference in asthma severity (i.e. men appear to have an increased incidence, but it is actually a reflection of men with more severe asthma being diagnosed). Also, in the analyses by Greer et al. (1993) and McDonnell et al. (1999), the exposure periods overlapped with the follow-up periods for asthma (Table 2); as discussed above, this might have biased the results away from the null. In addition, these studies did not consider temporal variability when estimating ozone exposure and did not adjust for key confounders, such as a family history of asthma and exposure to other environmental allergens and irritants (Table 2).

It is notable that while the Seventh-day Adventists cohort is in California, this population is characterized by markedly lower risks of certain chronic diseases than the US general population, which was hypothesized to be attributable to the lifestyle and diet in the Seventh-day Adventists. The observed association between ozone and asthma in Seventh-day Adventist men, and the effect modification by sex may not be entirely generalizable to the US general population.

Study quality evaluation

A study quality evaluation (i.e. assessing internal validity of individual studies) is a key component in the systematic review process (NRC, 2014). There are many study evaluation frameworks available for most realms of evidence, several of which formed the theoretical basis for our study quality criteria (Koustas et al., 2013, 2014; NTP, 2015a,b; US EPA, 2013b, 2014, 2016a,b; von Elm et al., 2007a–e;). However, while many of these frameworks provide guidance on key considerations in study quality evaluation, they do not stipulate detailed criteria specific to the exposure and outcome of interest. For example the OHAT RoB approach considers three key domains: exposure characterization, outcome assessment and adjustment for confounding, and several other RoB criteria for observational studies (NTP, 2015a,b). Studies with definitely or probably low risk of bias in all three key domains are categorized in the 1st study quality tier and those with definitely or probably high risk of bias in any of the three key domains are categorized in the 2nd or 3rd study quality tiers (NTP, 2015a,b). Studies are then evaluated in tiers, but the impact of specific aspects of study quality may not be considered when interpreting study results.

In contrast, our quality criteria address more detailed and distinct characteristics that are specific to long-term ambient ozone exposure and asthma development, resulting in a comprehensive and in-depth assessment of the quality of the overall evidence. We found that all of the available studies had substantial uncertainties in exposure and outcome assessments and in adjustments for confounders, which is consistent with categorizing these studies in 2nd and 3rd tiers using the OHAT approach. The overall low quality of the available studies reduces our confidence in the body of evidence.

Biological plausibility

Some of the reported effects of ozone exposure on the respiratory system are common features of asthma, providing biological plausibility for long-term ozone exposure as a contributor to asthma development. For example when ozone exposure concentrations are high enough and of sufficient duration to overwhelm antioxidant defenses in the respiratory tract, ozone reacts with macromolecules to form secondary oxidation products that can induce respiratory tract inflammation and epithelial cell injury with subsequent airway remodeling (Mudway & Kelly, 2000; US EPA, 2013a). Inflammation contributes to oxidative stress and reactive oxygen species (ROS) generation, which may cause increased bronchial reactivity (indicating AHR) observed as both an immediate and persistent response to ozone exposure in humans and experimental animals, respectively (US EPA, 2013a).

Asthma development also involves alterations in the immune system, and T-cells play a central role in asthma pathogenesis (Wu et al., 2014). Development of the allergic asthma phenotype is associated with a skewing of T-cell responses toward a T helper 2 (Th2) immune response instead of a T helper 1 (Th1) response. Th2 responses include several features of asthma, such as increased production of inflammatory cytokines, eosinophils and immunoglobulin E (IgE) antibodies in the respiratory tract, which induces a state of allergic sensitization (Bates et al., 2009; Marino et al., 2015; Noutsios & Floros, 2014). Exposure of children to ozone has been shown to increase eosinophils and proinflammatory cytokines in the respiratory tract (Noutsios & Floros, 2014) and short-term ozone exposure

studies in rodents and primates have reported enhanced IgE production (US EPA, 2013a). Short-term ozone exposure has also been shown to enhance antigen presentation in humans, which contributes to exaggerated T-cell responses and promotes Th2-mediated inflammation and an allergic phenotype (US EPA, 2013a). The persistent nature of these short-term responses contributes to the biological plausibility of an association between long-term ozone exposure and development of allergic asthma (US EPA, 2013a). We note, however, that although ozone exposure has been linked with Th2-related responses in several studies, there are no direct experimental studies associating ozone exposure with the initial polarization of Th1–Th2 responses (Auten & Foster, 2011).

Studies in non-human primates exposed repeatedly to ozone, with or without exposure to house dust mite antigen (an inhaled allergen), provide a model of allergic asthma in humans, as these studies reported reversible impaired airflow, elevated IgE levels in serum and airways, increased eosinophils and other immune cells in the airways and development of AHR (US EPA, 2013a). Long-term exposure of infant rhesus monkeys to 0.5 parts per million (ppm) ozone and an antigen-induced greater than additive changes in airway resistance and AHR, and exposure to ozone alone increased eosinophil levels and induced effects related to airway obstruction and AHR, including decreased growth of the distal conducting airways that persisted at least six months after cessation of ozone exposure (Fanucchi et al., 2006; Plopper et al., 2007; Schelegle et al., 2003). This indicates that long-term exposure to ozone and allergens in early life can disrupt growth and differentiation processes in the airways and provides plausibility that ozone could contribute to asthma development in children.

Several genetic factors that confer susceptibility to asthma development have been identified (Gowers et al., 2012; US EPA, 2013a). Some of these factors may modify asthma risk by affecting the internal dose of ozone reaction products or their response in the respiratory tract, providing biological plausibility for ozone exposure having a role in asthma development. For example variation in genes for antioxidant or anti-inflammatory enzymes could increase susceptibility to ozone-induced oxidative stress, inflammation and subsequent airway remodeling (Gowers et al., 2012). Another hypothesis is that in those with a genetic predisposition to atopy (i.e. a tendency to develop allergic diseases, such as asthma), there is a dominant Th2 inflammatory phenotype, and continuous exposure to ozone and allergens could cause atopic bronchial hyperresponsiveness and lead to asthma (Gowers et al., 2012; Marino et al., 2015).

Variability of asthma phenotypes has also been hypothesized to be attributable to epigenetic influences from early life environmental exposures (such as to ozone). In *in vitro* studies, histone modification and DNA methylation of Th lymphocyte genes have been induced by perinatal exposure to environmental insults, such as oxidative stress, leading to polarization toward Th2 responses associated with development of allergic asthma (Noutsios & Floros, 2014). Thus, it is plausible that ozone exposure may induce epigenetic changes associated with development of certain asthma phenotypes, although there are no studies available to support this directly.

Overall, the experimental evidence discussed above indicates that ozone exposure can induce many persistent respiratory effects that are also involved in asthma pathogenesis, suggesting that long-term ozone exposure could plausibly contribute to asthma development. However, the underlying mechanisms of asthma development in humans are poorly understood, so it is unclear whether the ozone-induced effects are direct contributors to asthma onset.

Conclusion

Available epidemiology studies of long-term ozone exposure, including *in utero* exposure, and development of childhood asthma reported inconsistent findings, which were further weakened by considerable methodological limitations in exposure and outcome assessments, and in statistical analyses. Epidemiology findings regarding long-term ozone exposure and adult-onset asthma, from one unique population with limited generalizability to the general population, were also limited by considerable uncertainties due to exposure measurement error and a lack of adjustment for key confounders. While experimental evidence suggests that it may be biologically plausible that long-term ozone exposure could contribute to asthma development, it has not established a definitive mode of action. Future research is needed to address key uncertainties regarding the role of ambient ozone exposure in asthma development.

Acknowledgments

The authors gratefully acknowledge the editorial review provided by Ms. Carla Walker (Gradient).

Disclosure statement

The authors are employed by Gradient, a private environmental consulting firm.

Funding

The work reported in this article was conducted by the authors during the normal course of employment with financial support provided to Gradient by the Electric Power Research Institute (EPRI). This article was reviewed by members of EPRI while in preparation, but the authors retain sole responsibility for the writing, content and conclusions in this article.

References

Abbey DE, Mills PK, Petersen FF, Beeson WL. (1991). Long-term ambient concentrations of total suspended particulates and oxidants as related to incidence of chronic disease in California Seventh-day Adventists. Environ Health Perspect 94:43–8.

Akinbami LJ, Moorman JE, Bailey C, et al. (2012). Trends in asthma prevalence, health care use, and mortality in the United States, 2001–2010. NCHS Data Brief 94:1–8.

American Thoracic Society, Ad Hoc Committee on Issues in Screening for Asthma in Children (ATS). (2007). An official ATS workshop

report: issues in screening for asthma in children. Proc Am Thorac Soc 4:133–41.
Auten RL, Foster WM. (2011). Biochemical effects of ozone on asthma during postnatal development. Biochim Biophys Acta 1810:1114–9.
Bates JH, Rincon M, Irvin CG. (2009). Animal models of asthma. Am J Physiol Lung Cell Mol Physiol 297:L401–10.
Clark NA, Demers PA, Karr CJ, et al. (2010). Effect of early life exposure to air pollution on development of childhood asthma. Environ Health Perspect 118:284–90.
Currie GP, Baker JFW. (eds.). (2012). Asthma. 2nd ed. Oxford, UK: Oxford University Press.
Dong GH, Qian Z, Liu MM, et al. (2013). Obesity enhanced respiratory health effects of ambient air pollution in Chinese children: the seven northeastern cities study. Int J Obes 37:94–100.
Fanucchi MV, Plopper CG, Evans MJ, et al. (2006). Cyclic exposure to ozone alters distal airway development in infant rhesus monkeys. Am J Physiol Lung Cell Mol Physiol 291:L644–50.
Global Initiative for Asthma (GINA). (2017). Global strategy for asthma management and prevention. Available from: http://ginasthma.org/download/317/.
Goodman JE, Zu K, Loftus CT, et al. (2018). Short-term ozone exposure and asthma severity: weight-of-evidence analysis. Environ Res 160:391–7.
Gowers AM, Cullinan P, Ayres JG, et al. (2012). Does outdoor air pollution induce new cases of asthma? biological plausibility and evidence; a review. Respirology 17:887–98.
Grainge CL, Davies DE. (2013). Epithelial injury and repair in airways diseases. Chest 144:1906–12.
Greer JR, Abbey DE, Burchette RJ. (1993). Asthma related to occupational and ambient air pollutants in nonsmokers. J Occup Med 35:909–15.
Ilmarinen P, Tuomisto LE, Kankaanranta H. (2015). Phenotypes, risk factors, and mechanisms of adult-onset asthma. Mediators Inflamm 2015:514868.
Islam T, McConnell R, Gauderman WJ, et al. (2008). Ozone, oxidant defense genes, and risk of asthma during adolescence. Am J Respir Crit Care Med 177:388–95.
Kim BJ, Seo JH, Jung YH, et al. (2013). Air pollution interacts with past episodes of bronchiolitis in the development of asthma. Allergy 68:517–23.
Koustas E, Lam J, Sutton P, et al. (2013). Applying the Navigation Guide: Case Study #1 – The impact of developmental exposure to perfluorooctanoic acid (PFOA) on fetal growth, a systematic review of the non-human evidence (final protocol). Available from: http://coe.ucsf.edu/prhe/pdfs/PFOA%20NON-HUMAN%20PROTOCOL.pdf.
Koustas E, Lam J, Sutton P, et al. (2014). The Navigation Guide – evidence-based medicine meets environmental health: systematic review of nonhuman evidence for PFOA effects on fetal growth. Environ Health Perspect 122:1015–27.
Lin S, Liu X, Le LH, Hwang SA. (2008). Chronic exposure to ambient ozone and asthma hospital admissions among children. Environ Health Perspect 116:1725–30.
MacIntyre EA, Bauer M, Melen E, et al. (2014). GSTP1 and TNF Gene variants and associations between air pollution and incident childhood asthma: the traffic, asthma and genetics (TAG) study. Environ Health Perspect 122:418–24.
Marino E, Caruso M, Campagna D, Polosa R. (2015). Impact of air quality on lung health: myth or reality? Ther Adv Chronic Dis 6:286–98.
McConnell R, Islam T, Shankardass K, et al. (2010). Childhood incident asthma and traffic-related air pollution at home and school. Environ Health Perspect 118:1021–6.
McDonnell WF, Abbey DE, Nishino N, Lebowitz MD. (1999). Long-term ambient ozone concentration and the incidence of asthma in nonsmoking adults: the Ahsmog study. Environ Res 80:110–21.
Moorman JE, Akinbami LJ, Bailey CM, et al. (2012). National surveillance of asthma: United States, 2001–2010. Vital Health Stat 3:1–67.
Mudway IS, Kelly FJ. (2000). Ozone and the lung: a sensitive issue. Mol Aspects Med 21:1–48.
Myers TR, Tomasio L. (2011). Asthma: 2015 and beyond. Respir Care 56:1389–407.
National Toxicology Program (NTP). (2015a). Handbook for conducting a literature-based health assessment using OHAT approach for systematic review and evidence integration. Office of Health Assessment and Translation (OHAT). Available from: http://ntp.niehs.nih.gov/pubhealth/hat/noms/index-2.html.
National Toxicology Program (NTP). (2015b). OHAT risk of bias rating tool for human and animal studies. Office of Health Assessment and Translation (OHAT). Available from: http://ntp.niehs.nih.gov/pubhealth/hat/noms/index-2.html.
Nishimura KK, Galanter JM, Roth LA, et al. (2013). Early-life air pollution and asthma risk in minority children. The GALA II and SAGE II studies. Am J Respir Crit Care Med 188:309–18.
Noutsios GT, Floros J. (2014). Childhood asthma: causes, risks, and protective factors; a role of innate immunity. Swiss Med Wkly 144:
National Research Council (NRC). (2014). Review of EPA's Integrated Risk Information System (IRIS) process. Washington, DC: National Academies Press. Available from: http://www.nap.edu/catalog.php?record_id=18764. [Last accessed: 6 May 2014].
Plopper CG, Smiley-Jewell SM, Miller LA, et al. (2007). Asthma/allergic airways disease: does postnatal exposure to environmental toxicants promote airway pathobiology? Toxicol Pathol 35:97–110.
Sbihi H, Tamburic L, Koehoorn M, Brauer M. (2016). Perinatal air pollution exposure and development of asthma from birth to age 10 years. Eur Respir J 47:1062–71.
Schelegle ES, Miller LA, Gershwin LJ, et al. (2003). Repeated episodes of ozone inhalation amplifies the effects of allergen sensitization and inhalation on airway immune and structural development in Rhesus monkeys. Toxicol Appl Pharmacol 191:74–85.
Shin YS, Takeda K, Gelfand EW. (2009). Understanding asthma using animal models. Allergy Asthma Immunol Res 1:10–8.
Subbarao P, Mandhane PJ, Sears MR. (2009). Asthma: epidemiology, etiology and risk factors. CMAJ 181:E181–90.
Tetreault LF, Doucet M, Gamache P, et al. (2016). Childhood exposure to ambient air pollutants and the onset of asthma: an administrative cohort study in Québec. Environ Health Perspect 124:1276–82.
United States Environmental Protection Agency (US EPA). (2013a). Integrated science assessment for ozone and related photochemical oxidants (final). National Center for Environmental Assessment (NCEA). Washington, DC: US EPA, EPA/600/R–10/076F.
United States Environmental Protection Agency (US EPA). (2013b). Materials submitted to the national research council part I: Status of implementation of recommendations. Integrated Risk Information System Program. Washington, DC: US EPA.
United States Environmental Protection Agency (US EPA). (2014). Draft Development Materials for the Integrated Risk Information System (IRIS) Toxicological Review of Inorganic Arsenic [CASRN 7440-38-2], EPA/630/R-14/101. Washington, DC: US EPA.
United States Environmental Protection Agency (US EPA). (2016a). Integrated science assessment for oxides of nitrogen – Health criteria (final). National Center for Environmental Assessment (NCEA), EPA/600/R-15/068. Washington, DC: US EPA.
United States Environmental Protection Agency (US EPA). (2016b). Integrated science assessment for sulfur oxides - Health criteria (second external review draft). National Center for Environmental Assessment (NCEA), EPA/600/R-16/351. Washington, DC: US EPA.
von Elm E, Altman DG, Egger M, et al. (2007a). Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. BMJ 335:806–8.
von Elm E, Altman DG, Egger M, et al. (2007b). STROBE statement - checklist of items that should be included in reports of case-control studies. Available from: https://www.strobe-statement.org/fileadmin/Strobe/uploads/checklists/STROBE_checklist_v4_case-control.pdf.
von Elm E, Altman DG, Egger M, et al. (2007c). STROBE statement - items to be included when reporting observational studies in a conference abstract. Available from: https://strobe-statement.org/fileadmin/Strobe/uploads/checklists/STROBE_checklist_conference_abstract_DRAFT.pdf.

von Elm E, Altman DG, Egger M, et al. (2007d). STROBE statement - checklist of items that should be included in reports of cross-sectional studies. Available from: https://strobe-statement.org/fileadmin/Strobe/uploads/checklists/STROBE_checklist_v4_cross-sectional.pdf.

von Elm E, Altman DG, Egger M, et al. (2007e). STROBE statement - checklist of items that should be included in reports of observational studies. Available from: https://strobe-statement.org/fileadmin/Strobe/uploads/checklists/STROBE_checklist_v4_combined_PlosMedicine.pdf.

Wang IJ, Tung TH, Tang CS, Zhao ZH. (2016). Allergens, air pollutants, and childhood allergic diseases. Int J Hyg Environ Health 219:66–71.

Wu D, Tan W, Zhang Q, et al. (2014). Effects of ozone exposure mediated by BEAS-2B cells on T cells activation: a possible link between environment and asthma. Asian Pac J Allergy Immunol 32:25–33.