March 5, 2014

Dr. H. Christopher Frey
Chair, Clean Air Scientific Advisory Committee
Science Advisory Board
US Environmental Protection Agency
1200 Pennsylvania Avenue NW
Washington, DC 20460

via email to Aaron Yeow, Designated Federal Officer, at yeow.aaron@epa.gov

Re: Comments on the NO_x Integrated Review Plan and Integrated Science Assessment

Dear Dr. Frey:

Attached to this letter are three files for the Clean Air Scientific Advisory Committee (CASAC) to consider in advance of the Oxides of Nitrogen Primary NAAQS Review Panel meeting next week. Attachment A is a brief set of comments regarding the NO_x IRP. Attachment B is a set of PowerPoint slides I presented to EPA in February summarizing Gradient's evaluation of the first draft ISA for NO_x. Although I summarize our key points in these slides, I am also providing lengthy written comments on the NO_x ISA as Attachment C in case CASAC members would like more detail beyond this presentation.

Thank you for your consideration. I look forward to testifying to CASAC next week.

Sincerely,

GRADIENT

Julie E. Goodman, Ph.D., DABT
Principal

G:\Projects\212112_APINOx\TextProc\l030514b.docx

Attachment A

Comments on the NO_x Integrated Review Plan

Comments on EPA's "Integrated Review Plan for the Primary National Ambient Air Quality Standards for Nitrogen Dioxide, External Review Draft"

Prepared by Julie E. Goodman, Ph.D., DABT
on behalf of the
American Petroleum Institute
1220 L St NW
Washington, DC 20005

March 5, 2014

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

Executive Summary

EPA released the *Integrated Review Plan for the Primary National Ambient Air Quality Standards for Nitrogen Dioxide (External Review Draft)* (IRP) in February 2014 (US EPA, 2014). The draft IRP reviews key findings from EPA's prior assessment and outlines its approach for the current Integrated Science Assessment (ISA) and Health Risk and Exposure Assessment (REA). Overall, the draft IRP presents a reasonable approach for conducting the ISA and REA. However, we identified several instances where the IRP could be more explicit and additional issues that should be (or should have been) considered or evaluated in the ISA and REA. Also, athough the IRP should have been completed prior to release of the first draft ISA, this did not occurr. Thus, we were able to identify instances where the first draft ISA (US EPA, 2013) did not strictly follow the proposed IRP approach. We discuss some of the issues we identified in further detail in our ISA comments (Gradient, 2014).

- Instances where the IRP should be more explicit: 1) the literature search strategy; 2) how studies evaluated in the 2008 ISA for nitrogen oxides will be integrated with those reviewed for the current ISA; 3) how study quality will be determined in a consistent manner across all studies; 4) how accuracy of summarized information will be ensured; 5) how evidence suggesting a lack of effect of nitrogen dioxide (NO_2) will be considered; and 6) the evaluation of whether observed health effects are due to NO_2 *per se* or traffic-related pollutants in general.
- Additional issues that should be considered or evaluated in the ISA and REA: 1) the ability to identify whether observed effects are due to NO_2 *per se* or traffic-related pollution should be considered when evaluating study quality; 2) evidence from all realms (*i.e.*, epidemiology, controlled exposure, animal toxicology) should be considered when interpreting results for individual studies; 3) uncertainties inherent in extrapolating potential health effects from high, experimental concentrations to low ambient exposure concentrations; 4) biological plausibility and coherence of proposed modes of action at ambient exposure levels; and 5) uncertainty in the risk estimates.
- Instances where the first draft ISA did not follow the approach outlined in the draft IRP: 1) evaluation of key considerations for determining causality; 2) consideration of whether studies conducted at high exposure concentrations are informative regarding biological plausibility and modes of action; and 3) evaluation of whether NO_2 is a surrogate for traffic-related pollutants.

In addition to cases were the IRP could be more explicit, or were where additional issues should be considered, the EPA's causal framework should be revised to more fully evaluate Bradford Hill's "aspects of association."

EPA released the *Integrated Review Plan for the Primary National Ambient Air Quality Standards for Nitrogen Dioxide (External Review Draft)* (IRP) in February 2014 (US EPA, 2014). The draft IRP reviews key findings from EPA's prior assessment and outlines its approach for the current Integrated Science Assessment (ISA) and Health Risk and Exposure Assessment (REA). These comments address instances where the IRP should be more explicit and identify additional issues that should be considered (or should have been considered) in the ISA or REA. Because the draft IRP was released after the first draft ISA (US EPA, 2013), these comments also identify instances where the ISA was inconsistent with the approach described in the draft IRP.

Regarding the General Approach for the Current Review (Section 3.2)

The extent to which studies provide evidence that effects are due to NO_2 *per se* rather than traffic-related pollution in general should be included as a "key policy-relevant" issue

Consideration of whether nitrogen dioxide (NO_2) serves as a surrogate for other traffic-related pollutants is listed as a "specific policy-relevant" issue in Section 4.4 of the draft IRP. However, given the importance of this issue, it should also be included as a "key policy-relevant" issue in Section 3.2.

Regarding Literature Search and Selection of Relevant Studies (Section 4.3.2)

The IRP should provide more detail regarding its literature search strategy

The draft IRP (US EPA, 2014) describes the literature search strategy but does not indicate which databases EPA will search, the specific search terms it will use, or specific study inclusion and exclusion criteria. All of this should be determined *a priori*, with any changes 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 one being developed by the Office of Health Assessment and Translation within the National Toxicology Program (Higgins and Green, 2011; NTP, 2012; Thayer, 2012; Boyles, 2012). A Cochrane Review is a systematic review of original studies of health care data (predominantly randomized and clinical controlled trials, but also observational studies) that uses a predefined, rigorous, and explicit methodology. The key goals of a Cochrane Review are 1) to collate all evidence that fits pre-specified eligibility criteria to address a specific research question; 2) to minimize bias by using explicit, systematic methods; and 3) to prepare, maintain, and promote systematic reviews to inform health care decisions. Specifically, the results of the primary investigations are synthesized by using strategies that limit bias and random error. These strategies include a comprehensive search of all potentially relevant studies and the use of explicit, reproducible criteria in the selection of studies for review. Primary research designs and study characteristics are appraised, data are synthesized, and results are interpreted. The *Cochrane Handbook for Systematic Reviews of Interventions*[1] is the official document that describes in detail the process of preparing and maintaining Cochrane Reviews.

The IRP should clarify how it will integrate studies evaluated in the 2008 NO_X ISA with those reviewed for the current ISA

The draft IRP (US EPA, 2014, p. 4-5) states:

[1] http://www.cochrane-handbook.org/

> In keeping with the purpose to accurately reflect the latest scientific knowledge, the ISA will emphasize studies published since the 2008 NO_X ISA. However, evidence from previous studies will be included to integrate with results from recent studies, and in some cases, characterize the key policy-relevant evidence in a particular subject area.

EPA should carefully consider and clearly articulate how evidence from older studies will be incorporated in the overall causality assessment. By incorporating older studies after the evaluation of recent studies, older studies may receive less weight in the analysis because of publication date. EPA should consider reanalyzing the key older studies so that they are considered in the same manner as the newer studies.

Regarding Evaluation of Individual Study Quality (Section 4.3.3)

The IRP should provide more detail regarding how it will determine study quality, and how it will do so in a consistent manner, for all studies

The draft IRP (US EPA, 2014) notes that study quality will be evaluated by assessing such things as the representativeness of the exposure assessment, adequacy of the study population, appropriateness of statistical analyses, control of potential confounders, and validity and reliability of health endpoints Throughout the plan, the draft IRP lists several important questions it will address in the ISA. All of these are crucial to consider when evaluating study quality and relevance.

The IRP should specify criteria for assessing these metrics to ensure they are applied in a consistent manner across studies. For example, there should be a discussion of statistical methods used among the studies evaluated and which specific methods are more robust and why (*e.g.*, have multiple comparisons been addressed? Are assumptions in Cox proportional hazard model appropriate?). 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 that the IRP should consider in detail include bias, measurement precision, replicability of observations, data reliability, outliers, selective outcome reporting, and fraudulent studies. It is crucial that these quality measurements be evaluated in the same way across studies. If a particular statistical model is considered a limitation in one study, for example, it must be considered a limitation in all studies that use it, unless there is a reason to conclude otherwise (and, if this is the case, the reason should be made clear). Importantly, an evaluation of study quality should be independent of the results of and funding source for that study. The evaluation should be based solely on the methodology, and studies with more robust methods should receive more weight in the overall analysis.

It is expected that, during the review, EPA will find study quality criteria not determined *a priori*. If so, these criteria should be applied to earlier studies to ensure that the assessment is consistent across studies. By more precisely defining and applying these criteria, the ISA will be more transparent and balanced.

The ability to identify whether effects are due to NO_2 *per se* or to traffic-related pollutants in general should be considered when evaluating study quality

Given the importance of this potential source of uncertainty, as discussed at the Workshop on science/policy issues (February 29-March 1, 2012), the extent to which studies can discern whether effects are due to NO_2 *per se* or traffic-related pollutants should be considered when evaluating individual studies. For example, the ISA should consider other pollutants associated with vehicle emissions, such as

aldehydes, that could be causally related to observed health effects. Studies that evaluate potential confounding by these other pollutants should be given greater weight than studies that do not.

The ISA should evaluate *in vitro* studies that both support and call into question results demonstrated *in vivo*

The draft IRP states, "*In vitro* studies may be included if they provide mechanistic insight or support results demonstrated *in vivo*." Notwithstanding their limitations, *in vitro* studies should be included not only if they support *in vivo* results, but also if they call *in vivo* results into question. In particular, concordance of results from *in vitro* studies with results from epidemiology studies should be evaluated, as findings from epidemiology studies are most likely to suffer from issues related to chance, bias, and/or confounding. Although results from *in vitro* studies by themselves should not be used to refute findings from epidemiology studies, results from *in vitro* studies should nevertheless be considered along with other lines of evidence, regardless of whether they support or refute findings from epidemiology, controlled human exposure, or animal studies.

Regarding Integration of Evidence and Determination of Causality (Section 4.3.4)

Determinations of causality in the first draft ISA have not adequately addressed key considerations outlined in the draft IRP

The draft IRP (US EPA, 2014, p. 4-9) states:

> [K]ey considerations in drawing conclusions about causality include consistency of findings for an endpoint across studies, coherence of the evidence across disciplines and across related endpoints, and biological plausibility, including key events that inform modes of action.

In practice, these considerations do not seem to factor prominently in the causal determinations presented in the first draft ISA. For example, as discussed in our comments regarding the first draft ISA (Gradient, 2014), evidence regarding lung function decrements in school children, as well as evidence from controlled human exposure studies, is not coherent with evidence for respiratory-related hospital admissions and emergency department visits. In addition, findings from animal toxicological and *in vitro* studies do not provide support for a biologically plausible mode of action (MoA) for these short-term respiratory effects.

EPA's NAAQS causal framework should be revised

The National Ambient Air Quality Standards (NAAQS) causal framework incorporates language from sources across the federal government and scientific community, particularly the Institute of Medicine (IOM) report *Improving the Presumptive Disability Decision-making Process for Veterans* (IOM, 2008). Where IOM recommended four levels of evidence categories for causation (Table 1), the NAAQS causal framework has five (Table 2). Based on these relationships, EPA determines which health effects will be evaluated in quantitative risk assessments (US EPA, 2013).

EPA's causal classification is largely based on Bradford Hill's "aspects of association." As discussed in Goodman *et al*. (2013), EPA's guidance for causality categorization is not congruent with the judgments

based on the Hill aspects. For example, the EPA NAAQS causal framework claims to rely heavily on consistency across studies in its categorization scheme, but, in practice, it has not always fully evaluated consistency nor fully incorporated aspects such as coherence, biological plausibility, biological gradient, and strength of association.

The first draft ISA states that evidence is sufficient to conclude a causal relationship if "chance, confounding, and other biases [can] be ruled out with reasonable confidence" (US EPA, 2013), yet there is no guidance on what constitutes "reasonable confidence." Based on the current NAAQS causal framework, this determination cannot be made reliably without fully exploring chance, bias, and confounding in a consistent manner. The draft ISA suggests that "controlled human exposure studies that demonstrate consistent effects" constitute evidence for a causal relationship (US EPA, 2013), but it should indicate that this is only true if the results are coherent with other lines of evidence. The draft ISA also indicates that "observational studies that cannot be explained by plausible alternatives" constitute evidence for a causal relationship (US EPA, 2013), yet, in practice, EPA has not fully explored alternative explanations for study results. In prior analyses, EPA set forth a hypothesis (*i.e.*, a criteria pollutant causes a particular health effect) and determined whether the data support that hypothesis. The ISA should fully explore whether and to what degree the data support *other* hypotheses (*e.g.*, a plausible confounder, rather than the criteria pollutant, causes a particular health effect). It is only in this manner that alternative hypotheses can truly be explored.

The draft ISA states that evidence is sufficient to conclude a likely causal relationship if "copollutant 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 if "animal toxicological evidence from multiple studies from different laboratories demonstrate effects, but limited or no human data are available" (US EPA, 2013). The draft ISA states that evidence is suggestive of a causal relationship if "at least one high-quality epidemiologic study shows an association with a given health outcome although inconsistencies remain across other studies that are or are not of comparable quality" or if "a well-conducted toxicological study, such as those conducted in the National Toxicology Program (NTP), shows effects in animal species" (US EPA, 2013). 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 a lack of causation. If positive effects in high-dose animal studies cannot be related to humans (*e.g.*, if mechanistic evidence indicates a lack of plausibility in humans), this does not constitute suggestive evidence; instead, these effects are uninformative regarding causation in humans. Not every study evaluating a criteria pollutant is informative for evaluating human health risk, and the ISA should not place undue weight on studies that are not relevant in humans.

The draft ISA's NAAQS causal framework requires only one high-quality study for evidence of a causal relationship to be deemed "suggestive." By this definition, other high-quality studies that are inconsistent may exist, but, as long as one high-quality study demonstrates an effect, this constitutes enough evidence for a suggestive relationship. Instead, all studies should be reviewed using the same criteria and a suggestive causal association chosen only if a causal association is more likely than not based on all the data.

Many of the issues noted above could be resolved if the EPA revises the NAAQS framework to make categories for causal determination more similar to the IOM framework on which it was based originally. The ISA should evaluate all data in a consistent manner using well-specified criteria and determine whether they constitute evidence for causation or are 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.

The ISA should consider each realm of evidence when interpreting study results

The NAAQS causal framework looks separately at epidemiology, controlled exposure, and animal toxicology evidence, first coming to a synthesized judgment for each, and then integrating these separate judgments into an overall qualitative statement about causality (US EPA, 2013). As discussed by Goodman *et al*. (2013) the data evaluation should be integrated across all lines of evidence *before* coming to judgments based on each realm independently. In this way, interpretation of each line of evidence informs the interpretation of the others. For example, if an epidemiology analysis can be interpreted two ways, and animal studies can shed light on whether one is more plausible than the other, this should be considered when making judgments about the epidemiology study.

This approach evaluates how results of particular studies can inform potential similar causal processes in other studies, including studies in other realms of investigation. It is the potential for such commonality of causal processes that makes animal data useful evidence for potential effects in humans. EPA's current method of integrating judgments at the end of the evaluation does not allow data from one realm of evidence to influence conclusions from another.

The ISA should discuss uncertainties inherent in extrapolating from exposure concentrations well above ambient exposures

The draft IRP (US EPA, 2014, p. 4-10) states:

> Causal determinations are developed for major outcome categories (*e.g.*, respiratory effects) or more specific groups of related endpoints and for the range of exposure concentrations of oxides of nitrogen that are representative of those across various ambient microenvironments. Findings based on higher exposure concentrations may be considered if they inform biological plausibility and modes of action.

The draft IRP states that findings from studies using high exposure concentrations may be considered if they are informative regarding potential MoAs, but studies cited in the first draft ISA as providing such support were invariably conducted at exposure concentrations much higher than ambient concentrations. The first draft ISA did not discuss whether these MoAs would be biologically plausible in humans at ambient concentrations.

The ISA should discuss the extent to which findings at higher exposure concentrations may or may not be predictive of responses at lower concentrations, specifically in light of potential MoAs for NO_2. If these MoAs involve a threshold response, such as saturation of anti-oxidants in the airway extracellular lining fluid, the ISA should consider the likelihood that findings at higher exposure concentrations may have exceeded such a threshold response.

Regarding Quality Management (Section 4.3.5)

The IRP should be explicit regarding quality assurance measures that will be used to ensure accuracy of summarized information

The draft IRP (US EPA, 2014) states, "The ISA also can include information that is integrated or reduced from multiple sources to create new figures, tables, or summation, which is subject to rigorous quality assurance measures to ensure their accuracy." The IRP should provide additional detail regarding the

quality assurance measures; for example, will information be extracted and summarized by two different individuals and compared for accuracy, or will information extracted and summarized by one individual be checked by a second individual? In addition to accuracy, the ISA should ensure that information is not extracted or summarized in a selective manner.

Regarding Specific Issues to be Addressed in the ISA (Section 4.4)

The IRP should be more explicit with respect to evidence required for determining a lack of effect

The draft IRP (US EPA, 2014) indicates that two issues to be addressed in the ISA are "whether new evidence reinforces or calls into question the evidence presented and evaluated in the last primary NAAQS review with respect to factors such as the concentrations of oxides of nitrogen associated with health effects and plausibility of health effects caused by exposure to oxides of nitrogen" and "whether uncertainties from the last review have been reduced and/or whether new uncertainties have emerged."

EPA needs to describe how it 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). There is a tendency to conclude that because of the *possibility* for causation, the data supports causality. Instead, one should conclude the evidence is non-informative.

If similar studies come to similar conclusions but all indicate that lack of causation is as likely as causation, they should not contribute to a causal conclusion. As stated by Bradford Hill (1965), "the same results from precisely the same form of inquiry will not invariably strengthen the original evidence. I would myself put a good deal of weight upon similar results reached in quite different ways."

The IRP should state that the ISA should only conclude that an association is causal if such an association is *more* likely than not.

The IRP should include consideration of biological plausibility and coherence as a policy-relevant issue related to dosimetry and MoA

Given that many of the studies regarding dosimetry and MoA are conducted at exposure concentrations substantially higher than ambient exposures, the ISA should consider whether the postulated MoAs would be biologically plausible in humans at ambient exposure levels.

The IRP should be more explicit regarding evaluation of whether observed health effect associations are due to NO_2 *per se* or traffic-related pollutants in general

Although the draft IRP states that it would be considered as a specific policy-relevant question, the first draft ISA did not fully evaluate whether NO_2 is a surrogate for traffic-related pollution. Importantly, as noted in our comments regarding the first draft ISA (Gradient, 2014), studies are cited as supporting an association between NO_2 and specific health effects, despite conclusions by the study authors that associations were more likely due to traffic-related pollution (*e.g.*, Jerrett *et al*., 2008; McConnell *et al*., 2010; Oftedal, 2009). Although the first draft ISA stated that it evaluated the influence of traffic-related pollution, in many cases, the evaluation was based on a limited set of traffic-related pollutants, such as

particulate matter or various forms of carbon (*e.g.*, black, elemental and organic), but not other traffic-related pollutants, such as volatile organic compounds, that could contribute to adverse health effects.

The IRP should identify criteria that can be used for evaluating whether observed effects may be due to NO_2 *per se* or traffic-related pollutants in general. This evaluation should be conducted both for individual studies and when considering consistency of findings across different lines of evidence. For example, if findings from epidemiology studies are uncertain as to whether effects are due to NO_2 *per se* or traffic-related pollutants and findings from animal studies show effects of NO_2 only at concentrations well above those evaluated in the epidemiology studies, the findings from the epidemiology studies may be more likely due to traffic-related pollutants. Importantly, if the ISA relies on studies for which the study authors conclude that findings may be due to traffic-related pollutants, the ISA should clearly articulate why it believes the effects are instead specific to NO_2.

The IRP should also broaden the list of potential traffic-related pollutants to include compounds such as acetaldehyde, 1,3-butadiene, and formaldehyde, which may also cause adverse health effects (*e.g.*, HEI, 2009).

Regarding Consideration of Quantitative Assessments for This Review (Section 5.2)

The REA should fully evaluate use of a threshold concentration-response function

The draft IRP notes that "the presence or not of a threshold" is a key area of uncertainty related to the epidemiological-based risk assessment. Given that the proposed MoAs are likely to operate only above a threshold, the IRP should suggest the REA give serious consideration to use of a threshold concentration-response function.

Uncertainty should be addressed in the risk estimates

To reduce uncertainty overall, quantitative risk estimates should focus on endpoints for which there is strong evidence of 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 how and when the uncertainty would under- and overestimate risks.

Conclusions

Although the draft IRP presents a reasonable approach for conducting the ISA and REA, there are instances where the IRP could be more explicit and other instances where it should consider or evaluate additional issues for conducting the ISA or REA. Because the customary sequence of events was not followed for this current NO_x review (*i.e.*, release of the draft IRP prior to that of the first draft ISA), we were also able to identify instances where the first draft ISA (US EPA, 2013) did not strictly follow the approach outlined in the IRP. In addition to cases were the IRP could be more explicit, or were where additional issues should be considered, the EPA's causal framework should be revised to more fully evaluate Bradford Hill's "aspects of association."

Table 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 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 copollutant 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 (2013, Table II).

References

Boyles, AL. 2012. "NTP Update: Systematic Review Incorporating Data Mining with Graphical Tools for Researchers." Office of Health Assessment and Translation. 30p, March 6.

Gradient. 2014. "Comments on US EPA's Integrated Science Assessment for Oxides of Nitrogen - Health Criteria (First External Review Draft)." Report to American Petroleum Institute (API) (Washington, DC) 60p., January 15.

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.

Health Effects Institute (HEI). 2009. "Traffic-Related Air Pollution: A Critical Review of the Literature on Emissions, Exposure, and Health Effects." HEI Panel on the Health Effects of Traffic-Related Air Pollution, Special Report 17. 394p., May.

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

Hill, AB. 1965. "The environment and disease: Association or causation?" *Proc. R. Soc. Med.* 58:295-300.

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

Jerrett, M; Shankardass, K; Berhane, K; Gauderman, WJ; Künzli, N; Avol, E; Gilliland, F; Lurmann, F; Molitor, JN; Molitor, JT; Thomas, DC; Peters, J; McConnell, R. 2008. "Traffic-related air pollution and asthma onset in children: A prospective cohort study with individual exposure measurement." *Environ. Health Perspect.* 116(10):1433-1438.

McConnell, R; Islam, T; Shankardass, K; Jerrett, M; Lurmann, F; Gilliland, F; Gauderman, J; Avol, E; Kunzli, N; Yao, L; Peters, J; Berhane, K. 2010. "Childhood incident asthma and traffic-related air pollution at home and school." *Environ. Health Perspect.* 118(7):1021-1026.

National Toxicology Program (NTP). 2012. "Summary Minutes of NTP Board of Scientific Counselors Meeting, held June 21-22, 2012, in Research Triangle Park, North Carolina." Board of Scientific Counselors. 36p, June.

Oftedal, B; Nystad, W; Brunekreef, B; Nafstad, P. 2009. "Long-term traffic-related exposures and asthma onset in schoolchildren in Oslo, Norway." *Environ. Health Perspect.* 117(5):839-844.

Thayer, K. 2012. "Systematic Review and New Tools for Information Management and Data Display." Office of Health Assessment and Translation. 1p.

US EPA. 2013. "Integrated Science Assessment for Oxides of Nitrogen–Health Criteria (First External Review Draft)." National Center for Environmental Assessment, EPA/600/R-13/202. 878p., November.

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 Office of Research and Development, Office of Air Quality Planning and Standards Office of Air and Radiation, EPA-452/P-14-001. 92p., February.

Attachment B

February 2014 Presentation to EPA on the NO_x Integrated Science Assessment

Comments on the Integrated Science Assessment for Oxides of Nitrogen – Health Criteria (First External Review Draft)

Julie E. Goodman, Ph.D., DABT

US EPA Research Triangle Park
February 5, 2014

GRADIENT

Causal Determinations for Short-term NO_2 Exposure

Health Effect Category	Causal Determination	
	2008 ISA for Oxides of Nitrogen	2013 Draft ISA for Oxides of Nitrogen
Respiratory Effects	Sufficient to determine a likely causal relationship	Causal
Cardiovascular Effects	Inadequate to infer the presence or absence of a causal relationship	Likely causal
Total Mortality	Suggestive but not sufficient to infer a causal relationship	Likely causal

2
Copyright Gradient 2013

GRADIENT

Causal Determinations for Long-term NO_2 Exposure

Health Effect Category	Causal Determination	
	2008 ISA for Oxides of Nitrogen	2013 Draft ISA for Oxides of Nitrogen
Respiratory Effects	Suggestive	Likely causal
Cardiovascular Effects	Inadequate	Suggestive
Reproductive & Developmental Effects	Inadequate	Fertility, Reproduction, and Pregnancy: Suggestive
		Birth Outcomes: Suggestive
		Postnatal Development: Suggestive
Total Mortality	Inadequate	Suggestive
Cancer	Inadequate	Suggestive

3
Copyright Gradient 2013
GRADIENT

Causal Determinations Not Supported by Evidence

- Epidemiology studies of inadequate quality
  - Selection bias
  - Ecologic bias
  - Confounding by co-pollutants and other factors
  - Exposure and outcome measurement error
  - Uncertainty regarding whether NO_x is a proxy for traffic-related pollution
- Associations not always consistent (mostly null or close to null), coherent, or biologically plausible
- Lack of confirmed MoAs

4
Copyright Gradient 2013
GRADIENT

Outline

- Causal Framework
- Chapter 3. Mode of Action/Dosimetry
- Chapter 4. Short-term NO_x Exposure
- Chapter 5. Long-term NO_x Exposure
- Chapter 6. Effect modifiers

Copyright Gradient 2013

GRADIENT

Causal Framework and the "Roadmap for Reform"

Critical Reviews in Toxicology
http://informahealthcare.com/txc
ISSN: 1040-8444 (print), 1547-6898 (electronic)
Crit Rev Toxicol, 2013; 43(9): 753–784
© 2013 Informa Healthcare USA, Inc. DOI: 10.3109/10408444.2013.832727
informa healthcare

REVIEW

A survey of frameworks for best practices in weight-of-evidence analyses

Lorenz R. Rhomberg[1], Julie E. Goodman[1], Lisa A. Bailey[1], Robyn L. Prueitt[1], Nancy B. Beck[2], Christopher Bevan[3], Michael Honeycutt[4], Norbert E. Kaminski[5], Greg Paoli[6], Lynn H. Pottenger[7], Roberta W. Scherer[8], Kimberly C. Wise[9], and Richard A. Becker[2]

Critical Reviews in Toxicology
http://informahealthcare.com/txc
ISSN: 1040-8444 (print), 1547-6898 (electronic)
Crit Rev Toxicol, Early Online: 1–21
© 2013 Informa Healthcare USA, Inc. DOI: 10.3109/10408444.2013.837864
informa healthcare

REVIEW ARTICLE

Evaluation of the causal framework used for setting National Ambient Air Quality Standards

Julie E. Goodman, Robyn L. Prueitt, Sonja N. Sax, Lisa A. Bailey, and Lorenz R. Rhomberg

Copyright Gradient 2013

GRADIENT

NAAQS Process for Causal Determination

Step	Description
1	Conduct Literature search
2	Select studies for inclusion
3	Consider general limitations of each study type
4	Use modified Bradford Hill aspects to aid in judging causality
5	Evaluate evidence for major health outcome categories
6	Integrate evidence from across disciplines and across health endpoints
7	Weigh alternative views on controversial issues
8	Characterize strength of evidence into causal conclusions
9	Assess adversity of effects

Copyright Gradient 2013
GRADIENT

Weight-of-Evidence Best Practices

1 Define causal question
Develop study selection criteria

2 Develop and apply criteria for review of individual studies

3 Integrate and evaluate evidence

4 Draw conclusions based on inferences

Copyright Gradient 2013
GRADIENT

1 Define causal question
Develop study selection criteria

- Explicitly state the causal question
- Explicit guidance for study selection
- Provide justification for study inclusion/exclusion
- Don't eliminate studies based on quality or null/negative findings

Copyright Gradient 2013

GRADIENT

2 Develop and apply criteria for review of individual studies

Study Quality
Study design
Selection bias
Sample population/size
Outcome assessment
Exposure assessment
Statistical modeling
Control for confounders
Sensitivity analysis

Study Relevance
Exposure/dose
Cross-species extrapolation

Copyright Gradient 2013

GRADIENT

3 Integrate and evaluate evidence

Consider modified Bradford Hill aspects
Integrate evidence across disciplines
Integrate older studies
Weigh alternative views on controversial issues
Consider when is something not likely to increase risk
Assess adversity

Copyright Gradient 2013

GRADIENT

4 Draw conclusions based on inferences

IOM Causation Categories
Sufficient
Equipoise and above
Below equipoise
Against

NAAQS Causal Framework
Causal
Likely causal
Suggestive
Inadequate
Not likely causal

Effect Modifiers (At-risk factors)
Adequate evidence
Suggestive evidence
Inadequate evidence
Evidence of no effect

Copyright Gradient 2013

GRADIENT

Chapter 3. Mode of Action/Dosimetry

- Relevance of high-dose studies to humans unclear
- No direct evidence to support NO_2 reaction products or their metabolites cause pulmonary effects
- Insufficient evaluation of reaction products from inhaled *vs*. endogenous NO_x
- Insufficient evaluation of beneficial effects of NO_2 metabolites
- No support for plausible MoA

Does not increase weight of evidence in support of causation

13

Copyright Gradient 2013

GRADIENT

Chapter 4. Short-term NO_x Controlled Exposure Studies

- Lack of consistent statistically significant effects on airway hyper-responsiveness in studies with specific allergens
- Lack of effect in studies involving exercise
  - Little evidence for refractory period after exercise
    - Exercise intensity insufficient
    - Studies cited in ISA not comparable to NO_x studies
- Lack of concentration-response below 600 ppb

14

Copyright Gradient 2013

GRADIENT

Chapter 4. Short-term NO_x Exposure

Modified Bradford Hill Aspect	Respiratory Effects (Asthma Morbidity) – "Causal"
Consistency	No
Coherence	Controlled exp studies of lung function not coherent with asthma HA and ED studies
Biological plausibility	No known plausible MoA
Biological gradient	None in controlled exposure studies Difficult to evaluate in epi studies
Strength of association	Small – Issues with exposure measurement error, outcome misclassification, and/or confounding
Experimental evidence	N/A
Temporal relationship	Yes
Specificity	No
Analogy	None

Copyright Gradient 2013
GRADIENT

Chapter 4. Short-term NO_x Exposure

Modified Bradford Hill Aspect	Cardiovascular Effects – "Likely Causal"
Consistency	No (for CV hospitalization and mortality)
Coherence	Not among epi and controlled exposure biomarker studies
Biological plausibility	No known plausible MoA
Biological gradient	Difficult to evaluate in epi studies
Strength of association	Small – Issues with exposure measurement error, outcome misclassification, and/or confounding
Experimental evidence	N/A
Temporal relationship	Mostly (most 0 or 1 day lags, some longer)
Specificity	No
Analogy	None

Copyright Gradient 2013
GRADIENT

Chapter 4. Short-term NO_x Exposure

Modified Bradford Hill Aspect	Total Mortality – "Likely Causal"
Consistency	No (particularly across cities)
Coherence	No
Biological plausibility	No known plausible MoA
Biological gradient	Difficult to evaluate in epi studies
Strength of association	Small – likely due to ecological bias, exposure measurement error, and/or confounding
Experimental evidence	N/A
Temporal relationship	Mostly (most 0 or 1 day lags, some longer)
Specificity	No
Analogy	None

Copyright Gradient 2013
GRADIENT

Chapter 5. Long-term NO_x Exposure

Modified Bradford Hill Aspect	Respiratory Effects (Child Asthma) – "Likely Causal"
Consistency	No
Coherence	No – effects at high exposures in tox studies
Biological plausibility	No confirmed MoA
Biological gradient	Difficult to evaluate in epi studies
Strength of association	Small – Issues with exposure measurement error, outcome misclassification, and/or confounding
Experimental evidence	N/A
Temporal relationship	Possible
Specificity	No
Analogy	None

Copyright Gradient 2013
GRADIENT

Chapter 5. Long-term NO_x Exposure "Suggestive" Endpoints

Health Effect Category	Why 2013 "Suggestive" Determination Is Not Supported
CV Effects	Lower quality epi studies and findings inconsistent Lack of confirmed MoAs and limited tox evidence
Fertility, Reproduction, and Pregnancy	No coherence: Pre-eclampsia in 3 of 4 studies, but hypertension findings inconsistent
Birth Outcomes	Inconsistent findings regarding timing of NO_2 exposure Coherence not evaluated
Postnatal Development	Lower quality epi studies; findings inconsistent No coherence with tox studies Lack of confirmed MoAs
Total Mortality	Lower quality epi studies; findings inconsistent Limited coherence with respiratory and cardiovascular morbidity
Cancer	Lower quality epi studies; findings inconsistent Lack of confirmed MoA

Copyright Gradient 2013
GRADIENT

Chapter 6. Effect Modifiers

- Evidence classification scheme different than causal framework
- ISA states evidence is systematically evaluated across disciplines and strength of evidence is evaluated; no indication how
- Study quality and relevance not discussed in preamble or Chapter 6
- Results discussed briefly, no critical review
- No indication how this analysis will be applied to setting the NAAQS

Copyright Gradient 2013
GRADIENT

Conclusions

- Evidence does not support stronger causal determinations in first draft 2013 ISA *vs.* 2008 ISA
  - Associations close to null and likely explained by chance, bias, and/or confounding
  - Associations not always consistent, coherent, or biologically plausible
- No MoA to support findings
- Evidence is not "suggestive" of causation for endpoints classified as such
- Value of effect modifiers analysis unclear
- Causal framework not adequate

Copyright Gradient 2013

GRADIENT

Frameworks

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.

Copyright Gradient 2013
GRADIENT

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 copollutant 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.

Copyr

EPA's Classification of Evidence for Potential Effect Modifiers ("At-risk Factors")

Classification	Health Effects
Adequate evidence	There is substantial, consistent evidence within a discipline to conclude that a factor results in a population or lifestage being at increased or decreased risk of air pollutant-related health effect(s) relative to some reference population or lifestage. Where applicable this includes coherence across disciplines. Evidence includes multiple high-quality studies.
Suggestive evidence	The collective evidence suggests that a factor results in a population or lifestage being at increased or decreased risk of an air pollutant-related health effect relative to some reference population or lifestage, but the evidence is limited due to some inconsistency within a discipline or, where applicable, a lack of coherence across disciplines.
Inadequate evidence	The collective evidence is inadequate to determine if a factor results in a population or lifestage being at increased or decreased risk of an air pollutant-related health effect relative to some reference population or lifestage. The available studies are of insufficient quantity, quality, consistency, and/or statistical power to permit a conclusion to be drawn.
Evidence of no effect	There is substantial, consistent evidence within a discipline to conclude that a factor does not result in a population or lifestage being at increased or decreased risk of air pollutant-related health effect(s) relative to some reference population or lifestage. Where applicable this includes coherence across disciplines. Evidence includes multiple high-quality studies.

Copyright Gradient 2013

GRADIENT

Attachment C

Comments on US EPA's Integrated Science Assessment for Oxides of Nitrogen - Health Criteria (First External Review Draft)

Comments on US EPA's Integrated Science Assessment for Oxides of Nitrogen - Health Criteria (First External Review Draft)

Prepared for
American Petroleum Institute
1220 L Street, NW
Washington, DC 20005-4070

January 15, 2014

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 The Causal Framework 2
- 2.1 Evaluation of Study Quality Should Be Transparent 2
- 2.2 Weight of Evidence for Causal Determination Should Be Strengthened 3

3 Dosimetry and Mode of Action 5
- 3.1 NO_2 5
- 3.2 NO 6
- 3.3 Conclusion 7

4 Short-term Exposure 8
- 4.1 Introduction 8
- 4.2 Respiratory Effects 8
  - 4.2.1 Asthma Morbidity 8
  - 4.2.2 Respiratory Mortality 18
  - 4.2.3 Causal Determination 19
- 4.3 Cardiovascular Effects 21
  - 4.3.1 Hospital Admissions and Emergency Department Visits 21
  - 4.3.2 Cardiovascular Mortality 25
  - 4.3.3 Supporting Evidence 26
  - 4.3.4 Causal Determination 28
- 4.4 Total Mortality 29
  - 4.4.1 Epidemiology Studies 29
  - 4.4.2 Causal Determination 30

5 Long-term Exposure 31
- 5.1 Introduction 31
- 5.2 Respiratory Effects 31
  - 5.2.1 Epidemiology Studies of Asthma Incidence in Children 31
  - 5.2.2 Causal Determination 33
- 5.3 Cardiovascular Effects 35
- 5.4 Reproductive and Developmental Effects 35
  - 5.4.1 Fertility, Reproduction, and Pregnancy 36
  - 5.4.2 Birth Outcomes 36
  - 5.4.3 Postnatal Development 36

5.5 Total Mortality 37
5.6 Cancer 37

6 "At-risk Factors" 38

7 Conclusions 40

References 41

List of Tables

Table 1.1 Causal Determinations for Short- and Long-term NO_2 Exposure in the 2008 and 2013 ISA

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

Table 2.2 EPA's Weight of Evidence for Causal Determination

Table 2.3 EPA's Classification of Evidence for Potential "At-risk Factors"

Executive Summary

In November 2013, the United States Environmental Protection Agency (EPA) released the first external review draft of the *Integrated Science Assessment for Oxides of Nitrogen – Health Criteria* (referred to as the "draft ISA" throughout these comments) (US EPA, 2013). The draft ISA reviews epidemiology, controlled human exposure, toxicology, and mode-of-action studies of nitrogen dioxide (NO_2) and other nitrogen oxides (NO_x) to determine whether (and, if so, with what degree of confidence) the evidence indicates short- and long-term NO_2 exposure can cause respiratory effects, cardiovascular effects, and total mortality, or long-term NO_2 exposure can cause reproductive and developmental effects and cancer. For all of these outcomes, the draft ISA concludes evidence is stronger than it was in the previous ISA (US EPA, 2008), but this is not supported by the evidence presented.

It appears that EPA's increased confidence in causality for these endpoints is based on the draft ISA's application of its causal framework ("EPA's causal framework"), which is largely based on the consideration of a modification of the Bradford Hill aspects (*i.e.*, strength of association, consistency and coherence, biological plausibility, biological gradient or exposure-response, specificity, temporality of effect, and adversity).[1] However, the framework does not provide enough guidance so that studies are evaluated in a consistent manner using well-specified criteria in the draft ISA. Also, the framework does not require the draft ISA to determine whether, as a whole, the data constitute evidence for causation or are more likely indicative of an alternative hypothesis (*e.g.*, that chance, bias, or confounding account for observed statistical associations). Thus, EPA's causal framework is not applied using a true weight-of-evidence approach.

For each endpoint the draft ISA classified as causal or likely causal, we evaluated the evidence considering EPA's causal framework's modified Bradford Hill aspects. In general, relationships deemed to be causal (short-term NO_2 exposure and respiratory effects) and likely causal (short-term NO_2 exposure and cardiovascular effects and total mortality; long-term NO_2 exposure and respiratory effects) in the draft ISA were close to the null and more likely explained by chance, bias (*e.g.*, exposure or outcome misclassification), and/or confounding (*e.g.*, by co-pollutants, smoking, socioeconomic status). In some cases, associations were not consistent either within or among epidemiology studies or coherent with controlled exposure studies. For several reported short-term effects, associations were found at lag times that do not appear to be biologically plausible. For numerous endpoints, a mode or mechanism of action was not established; where identified, modes or mechanisms of action lacked biological plausibility. It was often difficult to assess the biological gradient of the effects of NO_2 because most studies assumed response linearity and reported effect estimates associated with increments in NO_2 concentrations. There is also uncertainty regarding whether observed associations are due to NO_2 *per se* or whether NO_2 is a surrogate for another pollutant or mixture of pollutants.

Although EPA concludes that evidence for long-term NO_2 exposure is suggestive of causation for cardiovascular effects, reproductive and developmental effects, total mortality, and cancer, the evidence presented in the draft ISA does not indicate that long-term NO_2 exposure is more likely than not to be a causal factor for these effects. Even if new, high-quality studies demonstrated statistically significant associations (which they do not), the results of all other relevant studies must be considered to determine whether these associations are likely indicative of causation. If one cannot resolve inconsistencies among

[1] From Hill (1965).

the data, it is inappropriate to conclude the data are suggestive of an association; rather, the weight of evidence indicates that the data are inadequate for drawing conclusions.

Finally, the draft ISA's framework for classifying potential effect modifiers (what EPA refers to as "at-risk factors"; *i.e.*, variables that differentially modify the observed effect of a risk factor on disease status) is not the same as the causal framework, and neither is sufficient for assessing causality. Also, the draft ISA does not conduct a systematic review of potential effect modifiers that considers study quality and relevance and weighs studies based on both before coming to conclusions. It is also not clear how the assessment of effect modifiers bears on the National Ambient Air Quality Standards, as these factors are already accounted for in studies evaluated in other sections of the draft ISA.

Overall, the evidence presented in the draft ISA does not indicate increased confidence in any causal classification. While not necessarily agreeing with the classifications regarding causality determined by the 2008 ISA, Gradient's opinion is that the draft ISA provides no evidence that the classifications should be strengthened for any of the endpoints reviewed.

1 Introduction

In November 2013, the United States Environmental Protection Agency (EPA) released the first external review draft of the *Integrated Science Assessment for Oxides of Nitrogen – Health Criteria* (referred to as the "draft ISA" throughout these comments) (US EPA, 2013). In the draft ISA, EPA reviews controlled human exposure, short- and long-term epidemiology, toxicology, and mode-of-action (MoA) studies of nitrogen dioxide (NO_2) and other nitrogen oxides (NO_x) and makes causal determinations regarding several health effects. These determinations are used to review and set National Ambient Air Quality Standards (NAAQS).

In its previous review of NO_2, EPA concluded that evidence was sufficient to establish a likely causal relationship between short-term NO_2 exposure and respiratory effects, inadequate to infer the presence or absence of a causal relationship between short-term NO_2 exposure and cardiovascular effects, and suggestive but not sufficient to infer a causal relationship between short-term NO_2 exposure and total mortality (US EPA, 2008) (Table 1.1). With regard to short-term exposure, the 2013 draft ISA indicated that evidence is now sufficient to conclude a causal relationship with respiratory effects and a likely causal relationship with cardiovascular effects and total mortality.

In addition, the 2008 ISA concluded that evidence was suggestive but not sufficient to assume a likely causal relationship between long-term NO_2 exposure and respiratory effects, and it was inadequate to infer the presence or absence of a causal relationship between long-term NO_2 exposure and cardiovascular effects, reproductive and developmental effects, total mortality, and cancer. In contrast, with the addition of a few new studies per endpoint, the 2013 draft ISA indicates that evidence for all of these endpoints is suggestive of a causal association with the exception of respiratory effects (for which the draft ISA concludes the evidence indicates a likely causal relationship).

In Section 2 of these comments, we discuss the framework set forth in the draft ISA for assessing causality (EPA's causal framework) and how it is applied in the draft ISA. Sections 3 and 4 review the evidence regarding health effects associated with short- and long-term NO_2 exposure, respectively. In Section 5, we discuss the draft ISA's evaluation of effect modifiers (what EPA refers to as "at-risk factors"; *i.e.*, variables that differentially modify the observed effect of a risk factor on disease status). Overall, we found that the evidence presented in the draft ISA does not support an increased confidence in causality for any endpoint.

2 The Causal Framework

As described in the draft ISA, EPA's causal framework describes scientific evidence used for evaluating general causation; characterizes the process for integrating and evaluating evidence for making causal determinations; identifies issues regarding uncertainty; and provides a framework for classifying and characterizing the weight of evidence (WoE) with respect to general causation. EPA notes that it evaluates strengths and weaknesses in studies across all relevant disciplines (*e.g.*, epidemiology, controlled exposure, toxicology, mechanistic studies), considering various sources of uncertainty.

EPA applies its causal framework in evaluating specific health endpoints potentially associated with NO_2. As such, appropriate design and application of the framework is critical for ensuring that causal determinations accurately reflect the WoE. As discussed below, there are several issues with both the design and application of EPA's causal framework in the draft ISA.

2.1 Evaluation of Study Quality Should Be Transparent

In the draft ISA, EPA notes that it evaluated study quality by assessing such things as the representativeness of the exposure assessment, adequacy of selecting the study population, appropriateness of statistical analyses, sufficient control of potential confounders, validity and reliability of health endpoints, and the overall biological coherence, internally and externally, of study findings. The draft ISA also lists several important questions it addresses regarding study quality.

The draft ISA needs to specify criteria for assessing these metrics to ensure they are assessed in a consistent manner across studies. For example, before evaluating any study results, the draft ISA should discuss all of the ways in which exposure can be measured, the strengths and limitations of each method, the possibility 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 and assumptions in Cox proportional hazard model are appropriate). Specific confounders should be addressed [*e.g.*, co-pollutants, socioeconomic status (SES), age, weather] in terms of how they are handled in different studies and their likely impact on results. Other factors that should be considered in detail include measurement bias, measurement precision, replicability of observations, data reliability, outliers, selective outcome reporting, and fraudulent studies.

It is crucial that these quality measures are evaluated in consistent manner across studies. If a particular statistical model is considered a limitation in one study, for example, it must be considered a limitation in all studies that use it, unless there is a reason to conclude otherwise (and, if this is the case, this reason should be made clear). Importantly, an evaluation of study quality should be independent of the results and funding source for that study. It should be based purely on the methods, in a consistent manner across studies, and studies with more robust methods should receive more weight in causal determinations. Currently, no explicit rationale is provided in the draft ISA for why certain studies are considered key evidence while others of similar quality are not.

2.2 Weight of Evidence for Causal Determination Should Be Strengthened

The EPA causal framework draws from language in sources across the federal government and scientific community, particularly the Institute of Medicine (IOM) report *Improving the Presumptive Disability Decision-making Process for Veterans* (IOM, 2008). Whereas the IOM recommended four categories for the level of evidence for causation (Table 2.1), EPA has five categories for causal relationships (Table 2.2). Based on these categories, EPA determines which health effects will be evaluated in quantitative risk assessments. Notably, the draft ISA uses another framework for classifying effect modifiers (which it calls "at-risk factors") that is much more similar to the IOM framework; however, EPA indicates that it is built on its own causal framework (Table 2.3, discussed below).

EPA's causal framework is supposed to be based on 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. In its current form, EPA's causal framework is not congruent with the judgments based on the original or modified Bradford Hill aspects. For example, the 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 incorporate aspects such as coherence, biological plausibility, biological gradient, and strength of association.

The draft ISA states that evidence is sufficient to conclude a *causal* relationship if "chance, bias, and confounding [can] be ruled out with reasonable confidence" (US EPA, 2013), yet there is no guidance on what constitutes "reasonable confidence." Based on the current framework, the draft ISA cannot make that determination reliably because it does not fully explore chance, bias, and confounding in a consistent manner. The draft ISA suggests "controlled human exposure studies that demonstrate consistent effects" constitute evidence for a causal relationship (US EPA, 2013), but it should indicate that this is only true if the exposures are at concentrations relevant to ambient exposure and the results are coherent with other lines of evidence. The draft ISA also indicates that "observational studies that cannot be explained by plausible alternatives" constitute evidence for a causal relationship (US EPA, 2013). Yet, the draft ISA does not fully explore alternative explanations for study results. Currently, the draft ISA sets forth a hypothesis (*i.e.*, a criteria pollutant causes a particular health effect) and determines whether the data support that hypothesis. The draft ISA does not, but should, fully explore whether and to what degree the data support *other* hypotheses (*e.g.*, a confounder, rather than the criteria pollutant, causes a particular health effect). It is only in this manner that alternative hypotheses can truly be ruled out.

The draft ISA states that evidence is sufficient to conclude a likely *causal* relationship if "copollutant 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 if "animal toxicological evidence from multiple studies from different laboratories…demonstrate effects, but limited or no human data are available" (US EPA, 2013). The draft ISA concludes evidence is *suggestive of a causal relationship* if "at least one high-quality epidemiologic study shows an association with a given health outcome but the results of other studies are inconsistent" or if "a well-conducted toxicological study, such as those conducted in the National Toxicology Program (NTP), shows effects in animal species" (US EPA, 2013).

For making determinations regarding causality, it is important to evaluate all available data, including positive, null, and negative evidence, in what is referred to as a WoE evaluation. 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 studies are inconsistent, the WoE may indicate a lack of causality or inadequate evidence to assess causation. If positive effects in high-dose animal studies cannot be related 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 the draft ISA should not place undue weight on studies that are not.

It is notable that the EPA 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 should be reviewed using the same criteria and one should conclude a suggestive causal association only if the WoE indicates that a causal association is more likely than not based on all the data combined. In situations where there are multiple, but inconsistent, high-quality studies, the appropriate conclusion is that the evidence is "below equipoise" (IOM, 2008).

Finally, evaluating the data as a whole means that one should evaluate not only how much evidence can be adduced to support (or to counter) 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 data that supports or counters a given hypothesis.

Many of the issues noted above could be resolved by updating the draft ISA's categories for causal determination to be more consistent with the IOM framework (on which it was based originally), outlined in Table 2.1. Ultimately, the draft ISA should evaluate all of the data in a consistent manner using well-specified criteria and determine whether, as a whole, they constitute evidence for causation or are 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 more supportive of causation (*i.e.*, equipoise and above in the IOM framework).

3 Dosimetry and Mode of Action

The draft ISA presents a relatively detailed chapter on the dosimetry[2] and MoA for NO_x exposures, which is a new and welcome addition to the assessment. What emerges from a close reading, however, is a high degree of inconsistency and uncertainty in the dosimetry evidence. In addition, the draft ISA does not provide plausible or cohesive MoAs for the various effects it ascribes to NO_x exposure at ambient concentrations. Consequently, as acknowledged in the draft ISA, there are no supported MoAs for effects of exposure to ambient NO_2, which undermines the draft ISA's assignment of more stringent causality designations.

3.1 NO_2

The draft ISA reviewed a number of controlled human exposure, animal, and *in vitro* studies that may lend support to potential MoAs for NO_2, including formation of secondary oxidation products, activation of neural reflexes, initiation of inflammation, alteration of epithelial barrier function, enhancement of bronchial smooth muscle reactivity, modification of innate/adaptive immunity, remodeling of airways and alveoli, and transduction of extrapulmonary responses. A common characteristic of these studies is exposure to NO_2 at very high concentrations, some reaching two to three orders of magnitude higher than the current NAAQS. The draft ISA states that results from studies with concentrations similar to ambient concentrations were considered preferentially but, in reality, few studies cited in the draft ISA used low (*i.e.*, ambient) concentrations. As a result, the relevance of any biological consequences observed is unclear. The draft ISA should discuss the uncertainties involved when extrapolating findings from exposure to high laboratory doses to much lower ambient doses.

According to the draft ISA, the effects of NO_2 are "likely" due to a product derived from the initial extracellular lining fluid (ELF) substrate and/or formation of secondary oxidants (US EPA, 2013, p. 3-27). Because NO_2 is a highly reactive gas and is quickly reduced to nitrite upon dissolution, it is unlikely that inhaled ambient NO_2 penetrates the ELF and interacts directly with respiratory tract epithelial cells, and even less likely that NO_2 itself is systemically distributed. As stated in the draft ISA, given the rapid reactions of inhaled NO_2 with various biological substrates, the short half-life of some reaction products, and the continuous turnover of the ELF, specific chemical species are not likely to persist for any appreciable time. This finding is supported by controlled human exposure studies at NO_2 concentrations much higher than the current NAAQS in which only temporarily reduced levels of antioxidants were found in the ELF. The draft ISA speculates that the "compounds thought responsible for pulmonary effects of inhaled NO_2 are the reaction products themselves or the metabolites of these products in the ELF." However, there is a lack of direct evidence to support this hypothesis.

The draft ISA notes that NO_2 is believed to be produced endogenously in the lung and organisms tend to be less sensitive to endogenously produced oxidants. It remains virtually impossible to distinguish reaction products from endogenous NO_2 and those from inhaled ambient NO_2. Endogenous levels of nitrate and nitrite are highly dependent on the diet, as diet represents the primary source for both. This is consistent with the results of *in vitro* studies using isolated lungs, which failed to establish changes of appreciable magnitude in nitrite levels following exposures to 10,000-20,000 ppb NO_2. As concluded in

[2] Defined in the draft ISA as "the measurement or estimation of the amount of a compound, or its reaction products, absorbed and/or generated at specific sites in the respiratory tract during an exposure."

the ISA, "it is unclear to what extent environmental exposures at current ambient NO_2 concentrations might affect the overall balance of nitrite and nitrate or how ambient NO_2 uptake compares with endogenous production rates/amounts" (US EPA, 2013, p. 3-23). Until these issues are resolved to an appreciable degree, it is not possible to develop or test potential MoAs for NO_2.

As stated in the draft ISA, current dosimetry models for NO_2 do not adequately consider reactive absorption and secondary reactions that affect the probability of oxidants/cytotoxic products reaching target sites. Further, it is unclear to what extent exposures at current ambient NO_2 concentrations might affect the overall balance of nitrite and nitrate or how ambient NO_2 uptake compares with endogenous production rates/amounts (US EPA, 2013, p. 3-23). In addition, although some scientists have postulated that N-nitrosamines, known carcinogens, can be formed when nitrite from NO_2 reacts with secondary amines, nitrosamines are not detected in the tissues of animals exposed by inhalation to NO_2 under normal conditions. It is thus speculative, according to the draft ISA, to suggest that NO_2 exposure may lead to nitrosamine formation. Finally, since ambient NO_2 contributes only modest amount of oxides of nitrogen relative to dietary intake, substantial contribution to systemic nitrosamine formation is not likely. Therefore, an MoA for carcinogenicity is not supported, and cancer is unlikely to be related to NO_2 exposure.

The adversity of NO_2 oxidation products (or, in some cases, lack thereof) should be discussed in more detail in the draft ISA. As stated in the draft ISA, the major metabolites of NO_2 and NO include nitrites and nitrates, S-nitrosothiols, and nitrated fatty acids, lipids, amino acids, and proteins. With the exception of nitrate, all are biologically active, yet the draft ISA states there is little evidence to "support a role for these metabolites in mediating the effects of NO_2 and NO" (US EPA, 2013, p. 3-55). A plausible MoA needs to account for the active metabolites; at present, the draft ISA does not accomplish this.

Further, there is evidence that nitrite, which is formed from inhaled NO_2 in the ELF, may have beneficial effects. Numerous studies have examined the effects of systemic nitrite on various tissues and organs, including respiratory and CV systems, although the majority of these studies are not adequately reviewed in the ISA. Findings from a few studies that the draft ISA mentions only briefly support a *beneficial* role of nitrite in preventing injuries to the heart and airways and relaxing smooth muscle. Some nitro-fatty acids, also metabolites of NO_2, are anti-inflammatory and vasculoprotective. The effects of these metabolites are not consistent with the adverse MoAs for NO_2 proposed in the draft ISA. There is no discussion on how to incorporate these beneficial biological effects into the overall health effects of NO_2 or how current literature is inadequate to delineate the complex biological process following NO_2 exposure.

3.2 NO

In addition to NO_2, NO is evaluated as an air pollutant in some epidemiology studies. The evaluation is complicated by the fact that humans produce NO endogenously, primarily as a result of dietary intake. Concentrations of NO measured in expired breath (eNO) range from ~5 to > 300 ppb while nasal NO concentrations are much higher, spanning from several hundred to several thousand ppb. Given the fact that endogenous NO exists at much higher concentrations in the respiratory tract than ambient NO, as stated in the draft ISA, "exposure to ambient NO may not generally affect the overall sequelae of its absorption, metabolism, or downstream impacts on vascular homeostasis or on lung and systemic biological processes." Although some epidemiology studies have reported associations between NO and health effects, it can be reasonably concluded that inhaled ambient NO would not play a significant role in any health outcomes associated with exposure to NOx.

3.3 Conclusion

In summary, the draft ISA does not provide a plausible or cohesive MoA or dosimetry for the various effects it ascribes to NO_x exposure at ambient concentrations, limiting its utility in supporting causal inference on health effects related to NO_2 exposure. A comprehensive delineation of MoA or dosimetry would require reliable and consistent information leading to plausible scenarios for how NO_2 causes various effects at ambient levels of exposure. At a minimum, such an analysis would need to explain how NO_2 interacts with the body to produce compounds that consistently cause specific effects at ambient concentrations, taking into consideration endogenous NO_2 concentrations – especially from the diet, the biggest source of NO_2 in the body. Consequently, as acknowledged in the draft ISA, there are no supported MoAs for effects of exposure to ambient NO_2, which undermines the draft ISA's increased confidence in causality compared to the 2008 ISA.

4 Short-term Exposure

4.1 Introduction

The 2008 ISA concluded that evidence was sufficient to (a) determine a likely causal relationship between short-term NO_2 exposure and respiratory effects, (b) inadequate to infer the presence or absence of a causal relationship between short-term NO_2 exposure and cardiovascular effects, and (c) suggestive but not sufficient to infer a causal relationship between short-term NO_2 exposure and total mortality (US EPA, 2008). The 2013 draft ISA indicates that evidence is now sufficient to conclude a causal relationship with respiratory effects and a likely causal relationship with cardiovascular effects and total mortality. As discussed below, the increased confidence in causality for these endpoints is not supported by the weight of evidence.

4.2 Respiratory Effects

4.2.1 Asthma Morbidity

The draft ISA concludes that there is a causal relationship between short-term NO_2 exposure and respiratory effects, specifically asthma morbidity in children. This conclusion is based primarily on evidence from epidemiology studies that evaluated hospital admissions and emergency department (ED) visits for asthma, respiratory symptoms, and lung function. The draft ISA also discusses supporting evidence from controlled human exposure studies and studies in both humans and animals regarding plausible biological mechanisms. Overall, in contrast to what is stated in the draft ISA, chance, bias, and confounding cannot be ruled out with reasonable confidence as possible explanations for NO_2-associated outcomes in these studies.

4.2.1.1 Epidemiology Studies

The draft ISA's causal determination regarding the relationship between short-term NO_2 exposure and asthma morbidity is based on evidence from several epidemiology studies EPA considered to be "key" (*i.e.*, of high quality). These particular studies reported that NO_2 exposure is associated with increased hospital admissions and ED visits for asthma, increased respiratory symptoms in children with asthma, and lung function decrements in children.

Overall, the epidemiology studies do not provide robust evidence of an association between NO_2 and either hospital admissions or ED visits for asthma. Effect estimates in all studies were small (with increases of 7% or less). In studies that evaluated associations in models with co-pollutants, NO_2 effect estimates were substantially reduced and, in most cases, no longer significant. Other studies did not evaluate effects in co-pollutant models, despite fairly high correlations between NO_2 and other pollutants, or findings of stronger associations for other pollutants. Biological plausibility was an issue in several studies; associations were stronger for exposures occurring several days prior to a hospital admission or ED visit than for same-day exposures, as well as observations of significant effects only for specific age groups that differed among pollutants. Finally, as noted by several of the study authors, it is not clear

whether NO_2 *per se* is causally associated with the outcomes evaluated in these studies or whether NO_2 is a surrogate for other pollutants or mixtures of pollutants.

4.2.1.1.1 Hospital Admissions and Emergency Department Visits

The draft ISA (Table 4-23) states that consistent associations from multiple high-quality epidemiology studies at relevant concentrations are the basis for a causal determination between short-term NO_2 exposure and asthma morbidity. In Section 4.2.7 of the draft ISA, EPA specifically discusses evidence from three key studies regarding hospital admissions (Samoli *et al*., 2011; Ko *et al*., 2007; Son *et al*., 2013) and three key studies regarding ED visits (Strickland *et al*., 2010; Villeneuve *et al*., 2007; Jalaludin *et al*., 2008). The draft ISA also lists two studies (Ito *et al*., 2007, and Li *et al*., 2011b, both as cited in US EPA, 2013) as key references in Table 4-23, but these studies are not discussed in the related text (Section 4.2.7.3).

Hospital Admissions

Samoli *et al*. (2011) conducted a single-city time series analysis in Athens, Greece, and reported a 4.13% increase in pediatric hospital admissions per interquartile (IQR) increase in NO_2 concentrations on the day of admission (lag 0). However, the increase was not statistically significant [95% confidence interval (CI): -2.50, 11.21] and, as noted in the draft ISA, the magnitude of the association was small compared to those for sulfur dioxide (SO_2; 7.84%) and coarse particulate matter (PM_{10}; 5.66%), both of which were statistically significant. The association with NO_2 was attenuated by approximately 50% in a co-pollutant model with PM_{10}, and it was negative in a co-pollutant model with SO_2. Given the stronger associations for PM_{10} and SO_2, and the attenuation and reversal of the association in co-pollutant models with PM_{10} and SO_2, respectively, the study by Samoli *et al*. (2011) does not clearly indicate that NO_2 is associated with hospital admissions for asthma. In fact, Samoli *et al*. (2011) concluded that their findings "provide limited evidence of an association between NO_2 exposure and asthma exacerbation." Samoli *et al*. (2011) further stated that NO_2 "probably acts as a proxy indicator of a mixture of pollutants and independently only to a small degree."

Ko *et al*. (2007) reported a significant association between NO_2 and emergency hospital admissions for asthma in Hong Kong, with a relative risk (RR) of 1.039 (95% CI: 1.051, 1.006) for children ages 1-14 years in a single pollutant model using a 5-day average concentration (lag 0-4). Although the RR was significant using other lags, the risk was greatest using lag 0-4. In a two-pollutant model with ozone (O_3), the association for NO_2 was greatly attenuated and no longer significant [RR for NO_2 reduced from 1.039 (95% CI: 1.028, 1.051) to 1.006 (95% CI: 0.998, 1.015)]. Moreover, the asthmatic response to an environmental agent is likely to be concurrent with the exposure rather than a consequence of exposures occurring over several days (Samoli *et al*., 2011). Hence, the observation that risks were greatest for lag 0-4 is not consistent with the biological basis for exacerbation of asthma by environmental triggers. Overall, the results from this study do not indicate that NO_2 is associated with emergency hospital admissions.

In an evaluation conducted in eight Korean cities, Son *et al*. (2013) observed a significant increase in asthma hospital admissions with an IQR increase in NO_2 concentrations (2.2%; 95% CI: 0.3%; 4.1%). The authors also observed a significant association between hospital admissions and O_3 that was much stronger than the association for NO_2 (increase of ~7.5%; 95% CI: ~11%, ~5%, for an IQR increase in O_3 concentration).[3] However, Son *et al*. (2013) did not report whether O_3 concentrations were correlated with NO_2 concentrations, and they did not evaluate whether the association with NO_2 remained significant

[3] Estimated from Figure 1; actual values were not provided in text.

in a co-pollutant model. Overall, this study does not provide robust evidence of an association between short-term NO_2 exposure and asthma hospital admissions.

Emergency Department Visits

Strickland *et al*. (2010) observed a significant association (RR = 1.066; 95% CI: 1.038, 1.095) between NO_2 exposure and ED visits for asthma among children ages 5-17 years old during the warm season in Atlanta, GA, using a single-pollutant model and a 3-day moving average concentration (lag 0-2). Significant associations in single-pollutant models were also observed for O_3, carbon monoxide (CO), and fine particulate matter ($PM_{2.5}$). The association with NO_2 remained significant in a two-pollutant model with O_3. However, despite a fairly high correlation between NO_2 and CO (Spearman correlation coefficient of 0.59), Strickland *et al*. (2010) did not evaluate whether the effect of NO_2 remained significant in a multi-pollutant model with CO. Hence, this study does not provide definitive evidence that NO_2 is associated with ED visits.

In a study conducted in Sydney, Australia, Jalaludin *et al*. (2008) observed a significant association between an IQR increase in NO_2 (lag 0) and ED visits for asthma among children ages 1-4 years (3% increase; 95% CI: 1.8%, 4.2%). Although the association with NO_2 remained significant in two-pollutant models with PM_{10}, CO, and SO_2, the association was not significant in a model with O_3 (OR = 2.1; 95% CI: 0.9, 3.4); it was only marginally significant in model with $PM_{2.5}$ (OR = 2.3; 95% CI: 1.0, 3.6). Further, the study authors note that "interpretation of results may yet be complicated by modeled air pollutants acting as a proxy for some poorly or unmeasured air pollutant" (Jalaludin *et al*., 2008). Considering the uncertainty regarding whether the observed associations with NO_2 may be due to other pollutants, this study does not provide robust evidence of an association between NO_2 and ED visits for asthma.

In a study by Villeneuve *et al*. (2007), ED visits for asthma for children ages 2-4 years were significantly associated with 1-day lag [odds ratio (OR) = 1.07; 95% CI: 1.03, 1.10], 3-day average (OR = 1.09; 95% CI: 1.04, 1.13), or 5-day average NO_2 concentrations (OR = 1.14; 95% CI: 1.09, 1.20), but not with same-day NO_2 concentrations (lag 0) (OR = 1.01; 95% CI: 0.98, 1.05). The lack of a significant effect with same-day NO_2 concentrations is not consistent with the time course for environmental exposures associated with asthma, which are expected to occur shortly after exposure rather than at a lag of 24 hours or more (Samoli *et al*., 2011). Five-day average CO and NO_2 were the only pollutants included in a two-pollutant model in this study; in this model, there was a significant effect of NO_2 only for ages 2-4 and a significant effect of CO only for ages 15-44. There is no clear biological basis to explain why different age groups would be sensitive to different air pollutants. Moreover, as noted above, the lack of a significant effect of NO_2 for same-day exposures is not consistent with the biological basis for exacerbation of asthma by environmental triggers. Given that the results are not coherent in terms of the biological understanding of pollutant exposures and asthma, this study does not provide robust evidence of an association between NO_2 and ED visits for asthma.

In a study by Li *et al*. (2011), NO_2 was significantly associated with ED visits for asthma among children ages 2-18 years in Detroit, MI, based on a 5-day lag (RR = 1.038, 95% CI: 1.005, 1.072) but not a 3-day lag (RR = 1.013; 95% CI: 0.981, 1.047). Li *et al*. did not provide results for other lags, including the most biologically plausible (0-day lag). Based on results from other epidemiology studies, as well as results from controlled exposure studies, it is not likely that there would be a 5-day delay to observe effects of NO_2 exposure. Moreover, Li *et al*. (2011) did not evaluate associations in multi-pollutant models. Hence, this study does not provide evidence that NO_2 is associated with ED visits for asthma.

4.2.1.1.2 Respiratory Symptoms

As stated in Table 4-23, key supporting evidence for a causal determination between short-term NO_2 exposure and asthma morbidity includes increases in respiratory symptoms in children with asthma in diverse populations (in association with 24-hour average, 2- to 4-hour average, and 1-hour max with lags 0 or 3- to 6-day average NO_2) in previous and recent multicity studies based on four key references (Mortimer *et al*., 2002; Mann *et al*., 2010; Gent *et al*., 2003; Schildcrout *et al*., 2006, all as cited by US EPA, 2013 as well as Section 4.2.6.1 and Figure 4-3 of the draft ISA. A review of the draft ISA, however, indicates that these statements are not supported by the studies cited, only one of which was published since the 2008 ISA.

As summarized in Table 4-18 of the draft ISA, "Epidemiologic studies of respiratory symptoms and asthma medication use in children" (Gent *et al*., 2003,[4] as cited in US EPA, 2013) provided no quantitative data. As shown in Table 4-18, Mann *et al*. (2010, as cited in US EPA, 2013) found a significant but very modest (OR = 1.24; 95% CI 1.05, 1.48) increase in wheezing associated with NO_2 (2-day lag), while Mortimer *et al*. (2002, as cited in US EPA, 2013) found a statistically significant increase in morning symptoms (OR = 1.80; 95% CI 1.03, 3.15). Schildcrout *et al*. (2006, as cited in US EPA, 2013) found a borderline significant increase with asthma symptoms (OR = 1.06; 95% CI 1.00, 1.12, 0-day lag) or rescue inhaler use (OR = 1.04; 95% CI 1.00, 1.08, 0-day lag). None of these studies are included in the draft ISA as having effects on either lung function or exhaled nitric oxide (eNO; Tables 4-7 and 4-14), further indicating a general lack of effects in these studies. In addition, all four studies used exposure data from central-site monitors; for example, Schildcrout *et al*. (2006, as cited in US EPA, 2013) used the average of multiple sites within 50 miles of a zip code, while Mortimer *et al*. (2002, as cited in US EPA, 2013) used an average of all city monitors. This contributes uncertainty to study results, since central-site monitors do not adequately reflect concentrations at other locations (such as where children live, play, and go to school).

Table 4-18 in the draft ISA summarizes symptoms for children with asthma exposed to NO_2. Three studies found no associations with NO_2 exposure (Sarnat *et al*., 2012; Holguin *et al*., 2007; Zora *et al*., 2013, all as cited in US EPA, 2013). As represented in Table 4-18, most studies evaluating wheezing or coughing (two of the most commonly assessed respiratory symptoms) had marginally positive findings with ORs slightly above 1: O'Connor *et al*., 2008 (OR = 1.17; 95% CI: 0.99, 1.37) (wheeze-cough); Spira-Cohen *et al*., 2011 (OR = 1.10; 95% CI: 0.87, 1.39) (wheeze); Barraza-Villarreal *et al*., 2008 [OR = 1.09; 95% CI: 1.03, 1.15 (wheeze), OR = 1.09; 95% CI: 1.04, 1.14 (cough)]; Escamilla-Nunez *et al*., 2008 [OR = 1.07; 95% CI: 1.02, 1.12 (cough), OR = 1.08; 95% CI: 1.02, 1.14 (wheeze)]; Jalaludin *et al*., 2004 [OR = 1.02; 95% CI: 0.86, 1.22 (wheeze), OR = 1.13; 95% CI: 1.00, 1.27 (wet cough)]; and Patel *et al*., 2010 (OR = 1.16; 95% CI: 0.93, 1.45) (wheeze). Although Mann *et al*. (2010, as cited in US EPA, 2013) found a higher OR for wheezing of 1.24 (95% CI: 1.05, 1.48), the NO_2 measurements were taken at a central-site monitor that was within 20 km from of homes and therefore not likely to provide representative air concentration data. The remaining studies either found no association or borderline increases in wheezing or coughing. These small increases, which are consistent across studies, are more likely the result of bias or confounding by other pollutants and do not support a causal relationship between NO_2 exposure and cough or wheezing in children with asthma.

As stated in the draft ISA (US EPA, 2013), a major uncertainty regarding the respiratory effects of short-term ambient NO_2 exposure was the high correlations of NO_2 with other traffic-related pollutants and the potential for NO_2 to serve primarily as an indicator for another or mixture of combustion-related pollutants. As shown in Table 4-18, among the studies that observed significant associations between

[4] The correct citation appears to be Gent *et al*. (2009).

NO_2 and respiratory symptoms (*e.g.*, O'Connor *et al*., 2008; Schildcrout *et al*., 2006, both as cited in US EPA, 2013), NO_2 may be serving as a surrogate for traffic-related pollution (TRP). According to the draft ISA, in addition to NO_2, most studies found associations for respiratory symptoms with co-pollutants such as PM_{10}, $PM_{2.5}$, $PM_{10-2.5}$, elemental carbon (EC), black carbon (BC), and black soot (BS), SO_2, and O_3, which showed a wide range of correlations with NO_2 (r = 0.23-0.64). Higher correlations existed for CO and BS/BC (r = 0.54-0.92) (US EPA, 2013, p. 4-137). Multicity studies analyzed multi-pollutant models (three pollutants) but had limited implications because of potential multi-collinearity (O'Connor *et al*., 2008; Mortimer *et al*., 2002, both as cited in US EPA, 2013) or joint effect models (Schildcrout *et al*., 2006, as cited in US EPA, 2013). In two studies covering smaller regions, although NO_2-wheeze associations were robust to adjustment for $PM_{10-2.5}$ among children in California (OR = 1.14; 95% CI: 0.95, 1.37 per 20-ppb increase in lag day 2 of 24-hour average NO_2; Mann *et al*., 2010, as cited in US EPA, 2013) and a source apportionment factor comprising EC, zinc, lead, copper, and selenium in New Haven County, CT [OR = 1.08; 95% CI: 0.99, 1.18 per unit increase in lag 0 NO_2 (unit not reported); Gent *et al*., 2003, as cited in US EPA, 2013], the increases in both studies were marginal. The draft ISA concludes that the few co-pollutant-adjusted results from these studies provide evidence for a consistent and independent association of NO_2. However, this conclusion is not supported by the data presented, which show few significant or clinically relevant associations with NO_2 in this group of studies.

4.2.1.1.3 Lung Function Decrements

Table 4-23 in the draft ISA presents arguments for a causal association between asthma morbidity and short-term NO_2 exposure, stating that consistent associations from multiple, high-quality epidemiology studies at relevant concentrations demonstrate a causal association between decrements in lung function in children with asthma in recent studies (Table 4-23 cites O'Connor *et al*., 2008; Greenwald *et al*., 2013; Holguin *et al*., 2007; Delfino *et al*., 2008a, all as cited in US EPA, 2013, in addition to Section 4.2.3.1, Figure 4-1). However, as demonstrated in Table 4-7 of the draft ISA, the changes in forced expiratory volume in 1 second (FEV_1) in many studies were not statistically significant (*e.g.*, Greenwald *et al*., 2013; Holguin *et al*., 2007; Spira-Cohen *et al*. 2011; Dales *et al*., 2009a; Barraza-Villareal *et al*., 2008; Hernandez-Cadena *et al*., 2009; Gillespie-Bennett *et al*., 2011, all as cited in US EPA, 2013) and may be attributable to chance or bias. In addition, findings were inconsistent: for example, in Greenwald *et al*. (2013), FEV_1 was positively associated with NO_2 in School A, especially indoors (38%), but FEV_1 was negatively associated with NO_2 at School B to a similar degree indoors (-14%) and outdoors (-17%). These inconsistent findings and lack of significance do not support a causal relation between NO_2 exposure and FEV_1 decrements in children with asthma.

Table 4-7 in the draft ISA, which summarizes lung function changes in epidemiology studies of children with asthma, includes data for Holguin *et al*. (2007), O'Connor *et al*. (2008), and Delfino *et al*. (2008a), all as cited in US EPA (2013). These studies are three of the four studies that form the basis of the draft ISA's conclusion that there is a causal association between asthma morbidity and short-term NO_2 exposure. All three studies reported small decrements in FEV_1 (ranging from -1.2% to -1.7%), which are unlikely to have any clinical relevance. Holguin *et al*. (2007, as cited in US EPA, 2013) reported an FEV_1 decrement of -1.2% in a small group of children with asthma (n = 31) with a 0-6 day average lag. O'Connor *et al*. (2008, as cited in US EPA, 2013) found a statistically significant decrease in predicted FEV_1 of -1.3% with a 1-5 day lag, and Delfino *et al*. (2008a, as cited in US EPA, 2013) found a significant decrease in predicted FEV_1, but the decreases were consistently small (-1.3% for central-site NO_2 with 0 lag, -1.5% for personal NO_2 with 0 lag, and -1.7% for personal NO_2 and a 0-1 day average lag). In addition, these three studies are not consistent with the Greenwald *et al*. (2013, as cited in US EPA, 2013) results, which included improvement in FEV_1. Finally, these four key studies reported effects associated with a range of lag times (0 lag, 0-1 day, 0-4 day average, 0-6 day average, and 1-5 day average), which does not support a common biological basis.

Table 4-7 of the draft ISA presents epidemiology studies of lung function in children with respiratory disease. The table includes the four key studies relied on in the draft ISA, as well as 12 additional studies. Many studies have small populations (less than 53) (Greenwald *et al*., 2013; Martins *et al*., 2012; Spira-Cohen *et al*., 2011; Yamazaki *et al*., 2011; Delfino *et al*., 2008a; Wiwatanadate and Trakultivakon, 2010, all as cited in US EPA, 2013), and some studies have no quantitative effect estimates (*e.g.*, Mortimer *et al*., 2002; Just *et al*., 2002; Yamazaki *et al*., 2011, all as cited in US EPA, 2013). A number of studies either did not adjust for co-pollutants or found attenuated associations with adjustment for co-pollutants (*e.g.*, Greenwald *et al*., 2013; Holguin *et al*., 2007; Martins *et al*., 2012; Liu *et al*., 2009b; Barraza-Villareal *et al*., 2008; Hernandez-Cadena *et al*., 2009; Wiwatanadate and Trakultivakon, 2010; Gillespie-Bennett *et al*., 2011, all as cited in US EPA, 2013). Greenwald *et al*. (2013) and Holguin *et al*. (2007), both as cited in US EPA (2013), for example, did not use co-pollutant models.

In studies evaluating co-pollutants, the co-pollutants evaluated are inconsistent and not comprehensive. As discussed in Section 2.6.4.1 and summarized in Table 4-7 of the draft ISA, a number of studies report Pearson or Spearman correlations of ambient NO_2 with other criteria pollutants, mainly focusing on those related to traffic sources ($PM_{2.5}$, CO, PM_{10}). In Table 4-7, for example, about half (7/16) of the studies evaluated the association of $PM_{2.5}$ with FEV_1. A few studies explored associations between NO_2 and other traffic-related pollutants, such as EC, ultrafine particulate matter (UFP), and, rarely, volatile organic compounds (VOCs). Some of these co-pollutants have stronger associations with FEV_1 than NO_2 (*e.g.*, BTEX in Greenwald *et al*., 2013, as cited in US EPA, 2013) and may be partially or wholly responsible for observed associations. Taken as whole, the evaluation of the impact of co-pollutants in the draft ISA is inconsistent and incomplete.

4.2.1.2 Indoor NO_2 Exposure Studies

Regarding respiratory effects, the draft ISA (Table 4-23) states that chance, confounding, and other biases can be ruled out with reasonable confidence by epidemiology evidence from studies that evaluated exposure to NO_2 indoors and cites in four key studies (Hansel *et al*., 2008; Lu *et al*., 2013; Sarnat *et al*., 2012; Greenwald *et al*., 2013, all as cited in US EPA, 2013) and three key evidence points: 1) Indoor NO_2 associated with increases in respiratory effects in children with asthma in previous and recent studies; 2) Consistent results across various lags of exposure and outcomes examined in previous and recent studies; and 3) Previous and recent panel studies of children examine representative populations recruited from schools.

In Greenwald *et al*. (2013), FEV_1 was positively associated with NO_2 exposure in School A (especially indoors – 38%), but at School B, located in a high traffic area, FEV_1 was negatively associated with NO_2 exposure both indoors (-14%) and outdoors (-17%), although FEV_1 (% predicted) was not significantly different between the schools. Similarly, eNO decreased in School A (both indoors and outdoors), while it increased in School B (both indoors and outdoors). In addition, there were highly significant differences between the schools in asthma control scores, with children from School B scoring higher on both the Pediatric Asthma Control Test (PACT) and the Asthma Control questionnaire (ACQ), indicating that children from School B had more poorly controlled asthma than children from School A. These differences could be related to a number of variables including socioeconomic factors.

In the Sarnat *et al*. (2012) study, children's guardians completed daily diaries that provided information on the incidence of respiratory symptoms, school absenteeism, and medication use each week. The authors concluded that school outdoor NO_2 exposure was not associated with modeled respiratory symptoms. In contrast, they observed positive overall associations between eNO and many of the measured pollutant metrics and reported that both traffic-related and non-traffic-related particles were typically more robust predictors of eNO than was NO_2, for which associations were "highly sensitive" to model specifications

(Sarnat *et al*., 2012). The authors concluded that the results suggested small subclinical changes in airway inflammation (as assessed by eNO) associated with air pollution exposures, but no corresponding changes in respiratory symptoms.[5]

Studies by Lu *et al*. (2013) and Hansel *et al*. (2008) (both as cited in US EPA, 2013) are not included in any table in the draft ISA and only referred to in passing in the text. They studied young children in Baltimore, MD, and evaluated symptoms based only on indoor exposure. Indoor exposures are generally higher than ambient exposures and may not be appropriate for evaluating effects for exposure outdoors, especially since indoor NO_2 is arguably a "marker" for pollutants from the combustion of gas in heaters and stoves. For example, Hansel *et al*. (2008) noted that NO_2 concentrations were significantly associated with use of a gas stove, space heater, and/or stove/oven for heat. The authors concluded that, because the presence of a gas stove or heater and the use of a space heater or gas stove or oven for heat were associated with higher NO_2 concentrations, ambient NO_2 concentrations were only minimally correlated with indoor levels; changes to home heating and cooking devices may be a feasible means to reduce the burden of asthma. Lu *et al*. found no associations between NO_2 levels and any of the asthma symptoms outcomes among normal weight participants, although NO_2 was associated with some asthma symptoms among overweight and obese participants (Lu *et al*., 2013, Table VI). The only statistically significant increase was for symptoms related to exercise ($p = 0.05$) (Lu *et al*., 2013, Table VI). There were no differences in relationships between NO_2 levels and asthma-related health care use, lung function, and eNO values among body-mass index categories.

Another concern is that eNO was evaluated as a marker of inflammation following NO_2 exposure, but eNO production was not consistently associated with NO_2 concentrations. However, NO is also produced endogenously in various cells along the respiratory tract by three enzymes called nitric oxide synthases (NOS). One of them, inducible NOS (iNOS), can be induced by inflammation, resulting in higher eNO levels in patients with upper respiratory tract infections or asthma than in patients without. As noted in the draft ISA, eNO is affected by a variety of factors, including disease state, diet, sex (or height), species, smoking history, and environmental exposure. As discussed above in Section 3, diet is particularly important for eNO production; the draft ISA point out that "endogenous NO production, even during inflammatory states, is at best modest compared to dietary intake," underscoring the lack of an impact of ambient exposure on endogenous NO production. Yet these factors have not been consistently or comprehensively evaluated in epidemiology studies. Studies with exhaled breath condensate (EBC), which captures aerosolized materials contained in exhaled air (including those directly related to reactive nitrogen chemistry, such as nitrate and nitrite), have resulted in concentrations of eNO and EBC constituents not being related in many cases, further calling into question the relationship between eNO and NO_2 and its metabolites. Thus, eNO is unlikely to be associated with NO_2 at ambient concentrations.

While it is unclear how heavily the draft ISA relies on the studies by Lu *et al*. (2013) and Hansel *et al*. (2008), neither provides strong evidence that ambient NO_2 causes respiratory effects related to asthma or that reducing ambient NO_2 concentrations would affect respiratory function/symptoms in children.

4.2.1.3 Controlled Human Exposure Studies

The draft ISA concludes that results from the controlled human exposure studies of airway hyper-responsiveness (AHR) provide supporting evidence to rule out chance, confounding, and bias in the epidemiology studies. Specifically, the draft ISA discusses the potential importance of AHR in terms of exposure to specific allergens, and it speculates that the lack of an NO_2 effect on AHR in studies involving exercise is due to a refractory period following exercise and that studies with multiple exposure

[5] Table 4-14 in the draft ISA reports that Sarnat *et al*. (2012) found a significant increase of 6.3% eNO as a 24-hour average, but these data could not be found in the journal article (Sarnat *et al*., 2012).

concentrations provide evidence of concentration-response. On closer inspection, the evidence does not indicate that NO_2 increases AHR for specific allergens or that the paradoxical lack of an effect in studies involving exercise is due to a refractory period, nor do the studies with multiple exposure concentrations provide clear, consistent evidence of a concentration-response curve. Overall, based on the weight of evidence, the association between NO_2 and AHR at NO_2 concentrations below 600 ppb remains questionable.

4.2.1.3.1 Lack of Statistically Significant Effect for Studies with Specific Allergens

The draft ISA notes that the "lack of statistical significance in Table 4-4 [which shows the fraction of study participants with increased AHR following exposure to a specific allergen] does not necessarily diminish the potential importance of allergen exposure" and "although rare, severe bronchoconstriction can occur with inhalation of very low allergen concentrations" (US EPA, 2013). The issue, however, is whether NO_2 exacerbates effects of allergen exposure, not whether allergen exposure is important. As evidence regarding the importance of an NO_2-exacerbation of allergen exposure, the draft ISA notes that studies by Jenkins *et al*. (1999), Strand *et al*. (1998), and Tunnicliffe *et al*. (1994) observed NO_2-induced effects on allergen exposure at NO_2 exposures as low as 260 ppb. Only the study by Strand *et al*. (1998) observed an effect of NO_2 at 260 ppb; Jenkins *et al*. observed an effect at 400 ppb but not 200 ppb. Tunnicliffe *et al*. observed an effect at 400 but not 100 ppb, and the authors specifically noted that the effect was "small." In contrast to these three studies, other studies that used specific allergen challenges did not observe increased AHR with NO_2 exposure, including studies by Ahmed *et al*. (1983, as cited in Folinsbee, 1992), Barck *et al*. (2002), HEI (2012), Strand *et al*. (1997), and Witten *et al*. (2005). Given that the majority of studies indicate that NO_2 does not have an effect on AHR for specific allergens, the impact of NO_2 on responses to allergen exposures is at best inconclusive.

4.2.1.3.2 Lack of an Effect in Studies Involving Exercise

The draft ISA indicates that the lack of a significant effect on AHR in studies where participants were exposed to NO_2 while exercising may be attributable to "the development of a refractory period following bouts of exercise," during which time response to an airway challenge is diminished. However, it is not clear that this actually explains the lack of an effect, as there is little evidence of such an effect either in the NO_2 AHR studies or other studies that evaluated AHR using comparable study designs.

As evidence of an exercise-induced refractory period, the draft ISA notes that – whereas Jorres and Magnussen (1990) found statistically significant increase in AHR to SO_2 following 30-minute exposure to 250 ppb at rest – Rubinstein *et al*. (1990) found no change in responsiveness to SO_2 inhalation following 30-minute exposure to 300 ppb NO_2, including 20 minutes of exercise. This observation, however, does not provide an explanation for why exposure to NO_2 at rest, but not while exercising, would have an effect on AHR, and it does not consider that there may be critical study differences other than exercise that could explain the results. For example, Jorres and Magnussen exposed study participants *via* mouthpiece, while Rubinstein *et al*. used an exposure chamber. As discussed by Goodman *et al*. (2009), effects are more likely to occur *via* mouthpiece exposures, but this is not a representative exposure method. Hence, the significant effect observed by Jorres and Magnussen (1990) could be related to the exposure method rather than whether study participants were exercising.

As further evidence of an exercise-induced refractory period, the draft ISA refers to studies by Inman *et al*. (1990) and Freedman *et al*. (1988). Inman *et al*. found that the protection afforded by exercise (in terms of reduced bronchoconstriction) increases with the level of exercise, as indicated by a correlating

increase in methacholine PC20. Freedman *et al*. found that methacholine-induced bronchoconstriction was reversed more rapidly following periods of exercise or hyperventilation than following periods of rest. The studies by Inman *et al* (1990) and Freedman *et al*. (1988), however, are not analogous to the NO_2 AHR studies, in which the bronchial challenge agent was administered following a period of exercise. In the Inman *et al*. study, methacholine was administered while study participants were exercising; in Freedman *et al*., methacholine was administered prior to exercise. Hence, results from these two studies do not necessarily apply to the NO_2 AHR studies.

In contrast to the studies by Inman *et al*. (1990) and Freedman *et al*. (1988), Boulet *et al*. (1987), Hahn *et al*. (1984), and Magnussen *et al*. (1986) used study designs in which the bronchial challenge agent was administered after a period of rest or exercise. Neither Boulet *et al*. (1987) or Hahn *et al*. observed an effect of exercise on subsequent AHR As discussed in the draft ISA, Magnussen *et al*. (1986) found that a diminished airway response to exercise correlated with a diminished response to a methacholine challenge. However, overall, there was no difference in the provocative dose before or after exercise. Moreover, study participants also inhaled cold, dry air, which is known to induce bronchoconstriction.

In addition, results from several of the NO_2 AHR studies do not provide evidence of a refractory period for AHR following exercise. The draft ISA notes that, in studies by Jorres and Magnussen (1991) and Strand *et al*. (1996), there was a "slight tendency for the PD_{100} to be lower following the filtered air exposures relative to control (no exposure, no exercise), with roughly 53% of the individuals having a lower PD_{100} following filtered air (with exercise)," concluding, "these two studies do not support an effect of exercise on AHR in studies evaluating effects of NO_2 exposure." Similarly, in the study by Roger *et al*. (1990), the concentration of methacholine required to induce a doubling in specific airway resistance was greater when tested for inclusion in the study (presumably without exercise) than following exercise (mean ± SE = 5 ± 1.0 *vs*. 3.3 ± 0.7).

Finally, the intensity of exercise in many of the NO_2 AHR studies may not have been sufficient to induce a refractory period. For example, in the study by Bauer *et al*. (1986), study participants exercised on a bicycle ergometer for only 10 minutes at a workload of 300 kpm/min, which is approximately equivalent to walking 3 miles per hour (mph) at a 5% grade (Martin, 1999). For comparison, a workload of 540 kpm/min is considered to be a low-intensity workload (Marra *et al*., 2005). By this metric, exercise intensity in the study by Rubinstein *et al*. (1990), which was approximately 480 kpm/min, would also have been considered to be low intensity.[6] Study participants in the study by Jorres and Magnussen (1991) exercised at a workload designed to achieve a ventilation rate of approximately 30 L/min. According to Martin (1999), a ventilation rate of 30 L/min would be observed in a well-trained jogger after walking on a treadmill for 5 minutes at 2.5 mph. In the study by Kleinman *et al*. (1983), the exercise intensity was aimed to achieve a ventilation rate twice that of resting ventilation, which is approximately 6 L/min in a well-trained jogger (Martin, 1999).

Importantly, the refractory effect does not necessarily abolish the AHR effect but, rather, increases the threshold for an effect (*e.g.*, Inman *et al*., 1990); any effect of exercise should apply to both NO_2 and clean air exposures. If there is an allergen response following clean air, then one should still observe a more pronounced effect following NO_2 exposure if NO_2 were associated with AHR. In fact, AHR was observed in the five NO_2 AHR studies that provided information to evaluate the extent of airway responsiveness following NO_2 *vs*. filtered air exposure, either as a decrease in FEV_1 or a resting provocative dose (Avol *et al*., 1988, 1989; HEI, 1989; Roger *et al*., 1990; Witten *et al*., 2005). Among these studies in which AHR was observed following exercise, only the study by Avol *et al*. (1989) observed a difference in AHR following exposure to NO_2 *vs*. air.

[6] Study participants exercised at a workload of 60-80 watts, with 1 watt = 6.12 kpm/min (ExRx.net).

4.2.1.3.3 Lack of a Concentration-Response

The draft ISA notes that studies evaluating multiple exposure concentrations, specifically those using resting exposure (*e.g.*, Bylin *et al*., 1988; Orehek *et al*., 1976; Tunnicliffe *et al*., 1994), provide evidence of an NO_2 concentration-response for AHR. Overall these studies provide limited, if any, evidence of a concentration-response. The draft ISA notes that Bylin *et al*. observed statistically significant effects on AHR at 270 ppb but not 140 ppb. However, the response at 530 ppb, as reflected by either fraction affected or the provocative dose, was lower than the response at either 140 or 270 ppb (Bylin *et al*., 1988). Hence, over the full range of study concentrations, Bylin *et al*. did not observe an increase in AHR with increasing NO_2 concentrations. As discussed in the draft ISA, the PD100 was similar at both 100 and 200 ppb NO_2 for three of four individuals in the study by Orehek *et al*. (for which there was responsiveness data) and was doubled for the fourth individual. Moreover, the fraction affected at these two exposure concentrations was comparable (0.7; 95% CI: 0.46, 0.88 at 100 ppb NO_2 *vs*. 0.75; 95% CI: 0.19, 0.81).[7] Based on the limited data from Orehek *et al*. (1976), there does not appear to be a clear concentration-response for NO_2. In the study by Tunnicliffe *et al*. (1994), although the fraction affected following NO_2 exposure was greater at 400 ppb (0.75; 95% 0.35, 0.97) than 100 ppb (0.63; 95% CI: 0.25, 0.92), the decrease in FEV_1 following an airway challenge was comparable at the two exposure concentrations (-4.90%; 95% CI: -9.62. -0.18 at 100 ppb *vs*. -5.29; 95% CI: -9.09, -1.49). Hence, this study does not provide evidence of a concentration-response.

4.2.1.4 Mode of Action

Table 4-23 in the draft ISA states that increases in eNO in children with asthma are associated with 2- and 24-hour average NO_2, based on Delfino *et al*. (2006), Sarnat *et al*. (2012), and Martins *et al*. (2012), all as cited in US EPA (2013). Elsewhere, the draft ISA states that, across studies, associations varied in strength and precision; however, most results indicated a pattern of increasing eNO with increasing short-term NO_2 exposure (US EPA, 2013, Figure 4-2 and Table 4-14).

According to the draft ISA, Sarnat *et al*. (2012, as cited in US EPA, 2013) found a modest, significant increase of 6.3% (95% CI: 2.5, 10%) with 24-hour NO_2 in school outdoor air with a 0- to 4-day average lag.[8] Delfino *et al*. (2006, as cited in US EPA, 2013) found a nonsignificant and smaller increase in eNO (1.1%; 95% CI: -2.0 to 4.3%) with a 0-day lag and significant increase of 7.5% with a 0- to 1-day lag, both for 24-hour average personal NO_2, while Martins *et al*. (2012, as cited in US EPA, 2013) found a nonsignificant increase of 14% (95% CI: -12%, 40%) in eNO at 0- to 6–day average lag. The range of lag times (0, 0-1, 0-4, and 0-6 days), the range of findings (1.1% to 6.3% for statistically significant increases), and the modest size of the increases call into question the biological plausibility and coherence of these findings. A number of additional studies summarized in Table 4-14 of the draft ISA report "no association" (*e.g.*, Holguin *et al*., 2007, as cited in US EPA, 2013) or found eNO changes that were not statistically significant (*e.g.*, Greenwald *et al*., 2013; Liu *et al*., 2009b; Berhane *et al*., 2011, all as cited in US EPA, 2013).

As discussed above, there are many issues with studies evaluating eNO. Most (eight of 11) of the studies summarized in Table 4-14 of the draft ISA did not adjust eNO analyses for co-pollutants [*e.g.*, Greenwald *et al*., 2013; Holguin *et al*., 2007; Martins *et al*., 2012 (co-pollutant examination for EBC pH but not for eNO); Flamant-Hulin *et al*., 2010; Liu *et al*., 2009b (co-pollutant examination for thiobarbituric acid reactive substances, *i.e.*, TBARS, but not for eNO); Barraza-Villareal *et al*., 2008; Romieu *et al*., 2008; Berhane *et al*., 2011, all as cited in US EPA, 2013]. Co-pollutants (*e.g.*, BC, VOCs, BTEX, $PM_{2.5}$, PM_{10},

[7] 95% CIs are as reported by Goodman *et al*. (2009).
[8] These data could not be found in the journal article.

acetaldehyde, O_3) were associated with eNO in some of the studies that did not evaluate associations in models with NO_2 and other pollutants (*e.g.*, Barraza-Villareal *et al*., 2008; Berhane *et al*., 2011; Greenwald *et al*., 2013; Flamant-Hulin *et al*., 2010; all as cited in US EPA, 2013), so associations in single-pollutant models may be unreliable and not be indicative of causation. Although the draft ISA discusses the importance of analyzing the impact of co-pollutants on results, it nonetheless bases its analysis and conclusions on a number of studies without co-pollutant analysis.

Studies with a small number of people are less reliable than those conducted with larger populations. Many of the studies summarized in Table 4-14 of the draft ISA have small numbers of individuals; about half had 51 or fewer, with even smaller numbers in subgroups (Delfino *et al*., 2006; Sarnat *et al*., 2012; Greenwald *et al*., 2013; Martins *et al*., 2012; Lin *et al*., 2011, all as cited in US EPA, 2013). The remainder had populations of 104-194 except for Berhane *et al*. (2011, as cited in US EPA, 2013), which was a cross-sectional study with 2,240 children. The only statistically significant finding presented in Table 4-14 of the draft ISA for Berhane *et al*. (2011, as cited in US EPA, 2013) was an increase in eNO for those with respiratory allergy (but not for those with asthma or no respiratory allergy); no co-pollutant model was used.

4.2.2 Respiratory Mortality

The draft ISA concludes that there are consistent positive associations between short-term NO_2 exposure and respiratory mortality, noting that there is evidence that the magnitude of the effect is larger compared to total and cardiovascular mortality. This conclusion is based primarily on recent multicity studies conducted in Asia (Wong *et al*., 2008), China (Chen *et al*., 2012), and Italy (Chiusolo *et al*., 2011; Bellini *et al*., 2007).

4.2.2.1 Epidemiology Studies

In a time-series analysis of air pollution and mortality in four Asian cities, Wong *et al*. (2008) reported a small combined increase in respiratory mortality (1.48%, 95% CI: 0.68-2.28) for a 10 $\mu g/m^3$ increase in NO_2 (lag 0-1) across the four cities in single-pollutant models. In a separate study, Chen *et al*. (2012) conducted a time-series analysis of short-term NO_2 exposure and mortality in 17 Chinese cities. A significant positive association between ambient NO_2 concentrations at lag 0-1 day and respiratory mortality was observed in the single-pollutant model (2.5%, 95% CI: 1.4-3.6 per 10 $\mu g/m^3$ increase in 24-hour average NO_2); this association was reduced and no longer statistically significant in two-pollutant models with PM_{10} or SO_2. Although the majority of city-specific effect estimates were positive, effect estimates were null for five of the 17 cities evaluated.

In a meta-analysis of findings from 15 Italian cities, Bellini *et al*. (2007) reported a null association between respiratory mortality and ambient NO_2 concentrations (1.55%, 95% CI: -2.22-5.38) for a 10 $\mu g/m^3$ increase in NO_2 at lag 0-1 days in the year-round analysis; the authors reported a small, statistically significant increase in the summer-only analysis (1.66%, 95% CI: 2.35-1.88). The study did not report results for multi-pollutant analyses. In another study of mortality in Italy, Chiusolo *et al*. (2011) used a case-crossover design to evaluate mortality and short-term NO_2 exposures in 10 cities. The authors reported a larger and statistically significant risk of respiratory mortality in a single-pollutant model (3.48%, 95% CI: 0.75-6.29 per 10 $\mu g/m^3$ increase in NO_2) for a lag of 1-5 days in the year-round analysis. The results were significantly higher for the April-September analysis (9.63%, 95% CI: 4.08-15.47) and robust to inclusion of PM_{10} and O_3. The study did not provide city-specific effect estimates.

4.2.3 Causal Determination

We evaluated the current evidence on short-term exposure to NO_2 and respiratory effects using the modified The Bradford Hill aspects, which aid in judging causality. These aspects include evaluation of strength of association, consistency and coherence, biological plausibility, biological gradient or exposure-response, specificity, temporality of effect, and adversity. Based on this evaluation, the evidence in the draft ISA is not sufficiently robust or coherent to indicate that the association between NO_2 exposure and short-term respiratory effects is causal. Further, several epidemiologists (*e.g.*, Delfino *et al*., 2008; Ito *et al*., 2007) have raised the possibility that NO_2 may be a surrogate marker for other air pollution components.

4.2.3.1 Asthma Morbidity

For key epidemiology studies of hospital admissions and ED visits for asthma, the effect estimates for NO_2 exposure were small and not robust for adjustment for co-pollutants. For epidemiology studies that evaluated lung function (specifically FEV_1) decrements, the effect estimates for ambient NO_2 exposure were small and largely non-statistically significant.

Common sources of bias in these epidemiology studies include selection bias, exposure measurement error, and outcome misclassification. Selection bias may be of concern when participants are recruited at a single clinic or study center. The extent of exposure measurement error is often substantial when measurements of central air monitors are used as surrogates for individual exposure.

Confounding contributes a significant amount of uncertainty to the findings of epidemiology studies. The influence of co-pollutants was often poorly controlled for, if at all, in studies relied upon in the draft ISA. As some pollutants showed similar effects to those of NO_2, it remains a challenge to identify the causal agent for the observed effects. Moreover, controlled human exposure and animal studies have shown that supplementation with antioxidant vitamins is protective against exposure to very high doses of NO_2 by reducing lipid peroxidation and increasing lung antioxidant enzyme activity. Therefore, dietary intake and supplemental use of antioxidants confound and modify the health effects of NO_2, particularly at ambient exposure concentrations. In the majority of epidemiology studies on NO_2, antioxidant intakes were seldom evaluated or considered as a confounder, which could result in considerable uncertainty in the results.

Evidence from epidemiology studies evaluating lung function decrements in children and controlled exposure studies evaluating AHR is not coherent with associations between short-term NO_2 exposure and asthma hospital admission and ED visits for children reported in epidemiology studies.

It is difficult to assess the biological gradient for effects of NO_2 in epidemiology studies, as response linearity was usually assumed and estimates associated with an increment in NO_2 concentrations were reported. There is a lack of biological gradient in controlled human exposure studies of AHR, and the concentrations of NO_2 exposure were much higher than ambient concentrations.

The animal toxicological and *in vitro* studies do not provide a biologically plausible MoA for the effects of NO_2. Furthermore, inhaled NO_2 generates a spectrum of reaction products and metabolites, some of which may protect against airway injuries and relax smooth muscle, but such effects have not been incorporated into the MoAs of NO_2.

4.2.3.2 Respiratory Mortality

Three of the four studies employed a time-series design. As discussed by Bukowski (2008), ecological bias may arise when causal inference on disease risk at the individual level is made based on results from time-series analyses. In Chiusolo *et al*. (2011), which employed a case-crossover design, a major flaw was the selection of control days; this likely biased the results.

The magnitude of respiratory mortality increases was small, with increases of 3% or less in year-round analyses (Bellini *et al*., 2007; Chiusolo *et al*., 2011; Chen *et al*., 2012; Wong *et al*., 2008). These estimates are more likely due to chance, bias, or confounding (*e.g.*, see Boffetta *et al*., 2008; Fewell *et al*., 2007) than indicative of a causal association.

Exposure measurement error is likely in the studies; city-level NO_2 concentrations were determined from one or a small number of central-site air monitors. These monitors measure background ambient NO_2 levels and do not reflect exposures that may be affected by nearby sources (such as traffic emissions) or indoor sources. In fact, Wong *et al*. (2008) noted that:

> Among the major limitations of our study was the difference in monitoring locations among the cities. In densely populated cities such as Hong Kong and Shanghai, the monitors tend to be close to major roadways, whereas in Bangkok and Wuhan the monitors are located farther from major pollutant sources.

Errors can also be introduced for covariates such as temperature or humidity, which are often measured only at a single location. Outcome measurement error often occurs as a result of misclassification of the cause of death; this also could have biased results.

Confounding is a major source of uncertainty that may have impacted reported respiratory mortality risks. For example, few studies evaluated the impact of considering co-pollutants as potential confounders. Also, significant seasonal differences were observed in several of the studies (*e.g.*, Bellini *et al*., 2007; Chiusolo *et al*., 2011), which may indicate that there are additional confounding factors (*e.g.*, temperature) that were not properly or fully controlled for. As noted by Bellini *et al*. (2007), the "huge seasonal differences cast some doubts on the fairness of the attribution of all of the effects to any specific pollutant."

Although epidemiology studies have generally reported small excess risks of respiratory mortality, there are inconsistencies both within and across studies. Within studies that reported city specific estimates, there was unexplained heterogeneity in effect estimates across cities (Wong *et al*., 2008; Chen *et al*., 2012). For example, in the study by Wong *et al*. (2008), associations between NO_2 and respiratory mortality were null in Bangkok and three times higher in Wuhan than in Hong Kong or Shanghai. In the study by Chen *et al*. (2012), effect estimates varied from negative to positive across cities, ranging from approximately -2% to 7% (Figure 1 in the study). Although the studies in Asia reported similar respiratory mortality estimates, the studies conducted in Italy reported different results, null associations in the study by Bellini *et al*. (2007), and a significant increase in respiratory mortality in the study by Chiusolo *et al*. (2011). In fact, Chiusolo *et al*. (2011) noted that mortality estimates were almost double those reported by Bellini *et al*. (2007) for the same cities, despite decreases in NO_2 levels. Also, Wong *et al*. (2008) noted that:

> In the combined four-city analysis, the excess risks per 10-μg/m^3 increase in NO_2 were 2–3 times greater than those derived from the APHEA (Air Pollution and Health: A European Approach) project (Samoli *et al*. 2006) for mortality at all ages due to all natural causes, cardiovascular disease, and respiratory disease (1.23% *vs*. 0.3%, 1.36% *vs*. 0.4%, and 1.48% *vs*. 0.38%, respectively).

Most studies reported effects for lag times of 0 and 1 day with the exception of Chiusolo *et al*. (2011), which found that effects remained up to a lag of 5 days. Given that the mode of action or biological processes by which NO_2 exposure leads to mortality are unknown, and the numerous uncertainties that call into question the epidemiology findings, they provide little support of a causal effect from NO_2 exposures.

It is difficult to assess the biological gradient of the effects of NO_2 on mortality because most studies assumed linearity and reported effect estimates associated with increments in NO_2 concentrations. Two studies presented exposure-response curves for respiratory mortality (Chen *et al*., 2012; Wong *et al*., 2008). For studies that were conducted in Asia (Chen *et al*., 2012; Wong *et al*., 2008), the results indicated a linear or J-shaped exposure-response curve that were different for each city (*e.g*., Wong *et al*., 2008). The shape of these curves could vary if co-pollutants were included in the analysis. Exposure measurement error could also linearize the exposure-response function, which may in fact be non-linear – particularly at lower exposure levels (Rhomberg *et al*., 2011).

Lastly, in terms of specificity, there are many causes of respiratory mortality, and it is difficult to separate an association specific to NO_2.

4.3 Cardiovascular Effects

The draft ISA concludes that the association between cardiovascular (CV) health effects and short-term exposure to NO_2 is likely to be causal, "based primarily on epidemiologic studies of adults that consistently demonstrate NO_2-associated hospitalizations and ED visits for cardiovascular effects and mortality from cardiovascular disease." In contrast to the draft ISA, we conclude that the evidence does not support a likely causal relationship.

4.3.1 Hospital Admissions and Emergency Department Visits

4.3.1.1 Studies of CV hospital admissions or ED visits reviewed in the 2008 ISA

Epidemiology studies reviewed in the 2008 ISA reported mixed findings regarding the associations between ambient NO_2 concentrations and hospital admissions or ED visits for CV effects, with most studies reporting positive associations. Yet, several of the studies reviewed in the 2008 ISA were considered to be key evidence in the 2013 draft ISA regarding associations between short-term NO_2 exposure and hospital admissions or ED visits for cardiovascular disease (CVD) in adults. Methodological limitations and inconsistent findings in these studies limit the confidence in causal inference.

Three studies conducted in potentially vulnerable populations reported that increased risks of hospital admissions or ED visits for CVD or ischemic heart disease (IHD) were significantly associated with increases in same day ambient NO_2 concentrations (Mann *et al*., 2002; von Klot *et al*., 2005; Peel *et al*., 2007). However, the results were only from single-pollutant models, so confounding by co-pollutants cannot be ruled out.

Two studies conducted in general populations reported significant positive associations between risks of hospitalization or ED visits for CVD and ambient NO_2 levels (Barnett *et al*., 2006; Tolbert *et al*., 2007). However, the observed associations were not robust against adjustment for co-pollutants such as CO and $PM_{2.5}$.

A multi-city study in Spain reported a significant positive association between CV hospitalization and ambient NO_2 concentrations (Ballester *et al*., 2006). The association was attenuated when adjusting for CO, SO_2, or PM in two-pollutant models, but remained statistically significant.

Clearly, studies that were considered key evidence in the 2008 ISA reported mixed findings. The effect estimates for NO_2 were close to null and thus susceptible to the influence of bias, confounding, and chance. In fact, one of the major uncertainties cited in causal determinations in the 2008 ISA was whether the results truly indicated the effects of NO_2 or were influenced by the confounding co-pollutants. Since 2008, additional epidemiology studies assessed the association between short-term NO_2 exposure and hospital admissions or ED visits for all or specific CV effects, but the results are still mixed.

4.3.1.2 Recent Studies of Hospital Admission or ED Visits for all CV Causes

Among studies that examined hospital admissions or ED visits for all CV causes, five studies are of higher quality. Ito *et al*. (2011) was a time-series analysis that primarily focused on the associations between $PM_{2.5}$ and CVD hospitalization and mortality in New York City. Risks of CVD hospitalization was reported to be positively associated with ambient NO_2 concentrations at lag 0 and 1 day. The results for NO_2 were from single-pollutant models without adjustment for potential confounders such as holidays and influenza epidemics. Several pollutants, such as elemental carbon, SO_2, and CO, showed similar patterns as NO_2 in associations with CVD hospitalizations. The effect estimates for NO_2 were close to the null value, making it difficult to rule out chance findings. Confounding by unmeasured factors and co-pollutants, as well as chance, contribute to the uncertainty of the NO_2-specific results. In fact, the authors of the study indicated that the pollutants may be important indicators of emission sources rather than directly responsible for the observed health effects.

Guo *et al*. (2009) employed a case-crossover design to investigate the association between ambient air pollution and ED visits for CVD at a single hospital in Beijing, China. This study design automatically controls for time-invariant individual confounders, but not time-varying confounders such as meteorological factors, co-pollutants, physical activity, and stress level. In this study, the control period was selected as the same day from the last week of the case period. This unidirectional selection of control may be biased because it does not take into account temporal trends of hospitalization rates and air pollution levels. City-level ambient NO_2 concentrations were assigned to cases as surrogates for individual exposure, resulting in substantial errors in exposure measurement. Results from single-pollutant models showed that risk of hospital ED visits for CVD significantly increased with higher NO_2 concentrations on the same day (OR = 1.014, 95% CI: 1.001-1.028). However, this association became null after adjustment for $PM_{2.5}$ and SO_2 (OR = 0.998, 95% CI: 0.979-1.018). Methodological limitations such as selection bias, residual confounding, and exposure measurement error undermine the validity of the results.

Chen *et al*. (2010) conducted a time-series analysis on ambient air pollution and hospital admission in Shanghai, China. The study design and statistical approaches were similar to those of Ito *et al*. (2011). However, no association was observed between CV hospital admissions and ambient NO_2 at lag 0 or 1 day in single-pollutant models. Instead, significant positive associations with CV hospital admissions were observed for NO_2 concentrations at lag 4 (% increase = 1.23, 95% CI: 0.53, 1.93) and 5 days (% increase = 0.80, 95% CI: 0.10, 1.49). The positive association at lag 5 day persisted after controlling for

PM_{10} (% increase = 0.71, 95% CI: 0.00, 1.41) but became null with adjustment for SO_2 (% increase = 0.28, 95% CI: −0.76, 1.32). However, residual confounding by unmeasured factors such as stress level may still be an issue, and chance cannot be ruled out. In the smoothing plots of pollutants against CV hospital admission, NO_2 showed a J-shaped exposure-response relationship, and there were no adverse effects observed until NO_2 reached levels higher than the current NAAQS.

Andersen *et al*. (2008) conducted a time-series analysis of air pollution and hospital admissions in the elderly in Copenhagen, Denmark. No association between ambient NO_2 concentrations at lag 0-3 days and CVD hospital admission was observed (RR = 1.00, 95% CI: 0.98-1.03). The results did not change with adjustment for the total number concentration of ultrafine and accumulation mode particles. This study controlled for a number of potential confounders in the statistical analysis, including temperature, temporal trends, day of week, holiday, and influenza epidemics. However, residual and unmeasured confounding cannot be ruled out completely.

Larrieu *et al*. (2007) was a multi-city time series analysis of air pollution and hospital admissions for CVD in France. A significant positive association between ambient NO_2 concentrations and hospital admission for CVD was observed (excess RR = 0.5, 95% CI: 0.1–1.0) from single-pollutant models. The effect of NO_2 appeared to be stronger for hospital admission for IHD among the elderly (excess RR = 2.4, 95% CI: 1.4-3.5). A number of potential confounders, including temperature, temporal trends, day of week, holiday, and influenza epidemics, were adjusted for in the statistical analysis. However, residual confounding by co-pollutants, as well as other unmeasured factors such as stress level, contributed to the uncertainty of the findings.

The draft ISA reviews three additional studies of hospital admission or ED visits for all CV causes. Pereira Filho *et al*. (2008) was a time-series analysis of air pollution and secondary health data from a single hospital in São Paulo, Brazil. Significant increases in risk of CV ED visits were observed with increased NO_2 concentrations at the same day. This positive association was stronger among those with diabetes compared to non-diabetics. The results were from single-pollutant models and important confounders, such as day of week and holiday, were not adjusted for in the models. Selection bias, confounding, and chance contribute to the uncertainty of the results.

Linares and Diaz (2010) was a time-series analysis of air pollution and CV admission in the elderly at a single hospital in Madrid, Spain. No association with NO_2 was observed using single-pollutant models. A number of potential confounders, including temperature, temporal trends, day of week, holiday, and influenza epidemics, were adjusted for in the statistical analysis; however, residual confounding by co-pollutants and selection bias may undermine the validity of the results.

Chan *et al*. (2008) conducted a time-series analysis of air pollution and CV admissions at a single hospital in Taipei, Taiwan. A non-significant positive association between NO_2 and CV hospital admissions was observed from single-pollutant models; however, some key confounders, such as day of week and holiday, were not adjusted for in the statistical models. Selection bias and confounding contribute to the uncertainty of the results.

Overall, these studies had several methodological limitations, as they usually utilized secondary health data from a single hospital, employed only single-pollutant models, and omitted certain key confounders. Selection bias and confounding severely undermine the validity of their findings.

4.3.1.3 Recent studies of hospital admission or ED visits for cardiac causes

Among studies that examined the effect of NO_2 on hospital admissions for cardiac effects, two studies (Bhaskaran *et al*., 2011; Turin *et al*., 2012) were considered less susceptible to outcome misclassification in the draft ISA because they identified cases of myocardial infarction (MI) through clinical registries rather than administrative data. Both studies employed the case-crossover design, which allows causal inference on individual disease risk. Null associations between ambient NO_2 concentrations at lag 0-2 days and hospitalization of MI were observed. The associations remained null with additional adjustment for multiple co-pollutants. Although exposure measurement error and residual confounding may have contributed to the uncertainty of the findings, the studies suggest that ambient NO_2 levels did not affect risk of MI.

Bell *et al*. (2008) employed time-series analyses to examine the effects of air pollution on cause-specific admissions at a single hospital in Taipei, Taiwan. A significant positive association was observed for same-day NO_2 levels with hospitalization for IHD from single-pollutant models. Different selection of central-site air monitors did not affect this association. However, the time series of hospitalizations were collected from a single hospital; therefore, selection bias may be an issue. Confounding by co-pollutants, as well as holiday, influenza epidemics, and stress level, further contribute to the uncertainty of the results.

Two related Canadian multi-city studies investigated the effects of NO_2 on ED visits for angina or acute MI and for chest pain (Stieb *et al*., 2009; Szyszkowicz, 2009). Both studies reported significant positive associations between rates of hospitalization and ambient NO_2 concentrations on the same day; however, the results were from single-pollutant models. Stieb *et al*. (2009) found the association between NO_2 and ED visits for angina or acute MI was entirely driven by data from a single city. In addition to potential confounding by co-pollutants, other key confounders, such as influenza epidemics and stress level, were not controlled for in the statistical analysis.

An Italian multicity study employed a case-crossover design to assess the association between ambient air pollution and risk of hospitalization for acute MI (Nuvolone *et al*., 2011). A significant positive association was observed using single-pollutant models. However, the association was attenuated and did not maintain statistical significance in the two-pollutant models with PM_{10}, and became null in the two-pollutant model adjusting for CO. A study in New Jersey also reported a significant positive association between NO_2 and hospitalization for transmural infarctions from single-pollutant models; this association became null when adjusting for $PM_{2.5}$ (Rich *et al*., 2010).

Several other studies reported positive associations between ambient NO_2 concentrations and hospitalizations or ED visits for MI, IHD, and acute coronary syndrome (Szyszkowicz, 2007; Thach *et al*., 2010; Vencloviene *et al*., 2011; Wichmann *et al*., 2012). However, only single-pollutant models were used in these studies, and it is unclear whether the observed associations would remain robust if co-pollutants were accounted for.

Two case-crossover studies conducted in two cities in Taiwan both reported significant positive associations between hospitalizations for MI and ambient NO_2 levels at lag 0-2 days from single-pollutant models (Cheng *et al*., 2009; Hsieh *et al*., 2010). The associations were robust in two-pollutant models with PM_{10}, SO_2, CO, or ozone. Substantial exposure measurement error and residual confounding by stress level and physical activity contribute to the uncertainty of the results.

The draft ISA states that epidemiology studies have reported broadly consistent findings on the positive associations between ambient NO_2 concentrations and risk of hospitalization or ED visits for CV effects,

and the associations are robust in co-pollutant models. Such statements are not supported by the literature. Recent epidemiology studies on air pollution and hospital admission or ED visits for CV causes suffered from methodological limitations similar to those in earlier studies. The findings are inconsistent and the effect estimates are small and close to the null value. The majority of observed positive associations from single-pollutant models are not robust with further adjustment for co-pollutants, suggesting the effect of NO_2 on risk of hospitalization or ED visits for CVD is not causal.

4.3.2 Cardiovascular Mortality

The draft ISA discusses four recent multi-city studies on ambient NO_2 concentrations and CV mortality. Bellini *et al*. (2007) conducted a meta-analysis of 15 Italian single-city time-series studies on short-term effects of air-pollution. A null association between CV mortality and ambient NO_2 concentrations at lag 0-1 days was observed in the all-year analysis. However, an increased risk of CV deaths was found to be significantly associated with increases in ambient NO_2 levels in summer. The results for NO_2 were from single-pollutant models.

Wong *et al*. (2008) was a time-series analysis of air pollution and mortality in four Asian cities. Increased risks of CV mortality were significantly associated with increases in ambient NO_2 concentrations from single-pollutant models.

Chen *et al*. (2012) conducted a time-series analysis of short-term NO_2 exposure and mortality in 17 Chinese cities. A significant positive association between ambient NO_2 concentrations at lag 0-1 day and CV mortality was observed in a single-pollutant model; this association remained robust in two-pollutant models with PM_{10} or SO_2. The majority of city-specific effect estimates were positive, and only three out of 17 cities showed null associations.

Chiusolo *et al*. (2011) was a case-crossover study of short-term effects of NO_2 on mortality in 10 Italian cities. A significant positive association between cardiac mortality and ambient NO_2 concentrations at lag 0-5 days was observed in the all-year analysis. This association was robust against adjustment for PM_{10}. The overall effects of NO_2 on cardiac mortality were stronger in summer and robust when controlling for ozone. City-specific effect estimates were not provided.

Multicity epidemiology studies of NO_2 and CV mortality have consistently reported a positive association that remained robust to adjustment for PM_{10}. However, in studies of hospitalization or ED visits, the effects of NO_2 on CV outcomes were primarily confounded by CO. None of the above mortality studies controlled for CO in their analyses, so confounding may have been an issue.

Three of the above studies were time-series analyses, which are ecological in nature. Ecological bias may arise when causal inference on disease risk at the individual level is made based on results from time-series analyses (Bukowski, 2008). Time series analyses also involved highly complex statistical models. The estimated coefficients are often sensitive to model choice or specification (Bukowski, 2008).

One common limitation in these studies is the use of averaged area-level NO_2 concentrations and meteorological factors measured at a single location. Ambient NO_2 concentrations were usually measured at single or multiple air monitors and averaged across an entire city/area. Area-level exposures do not take into account spatial variation, and the resulting errors in exposure measurement may be substantial. Meteorological factors were usually measured at a single location and used for an entire city/area, which do not reflect conditions all residents in the city/area experience (Bukowski, 2008).

In addition to co-pollutants, there are many acute risk factors of CV outcomes, such as physical exertion, emotional distress, sleep disturbance, smoking, alcohol consumption, and sexual activity (Bukowski, 2008). Many of these factors are conceivably associated with exposure to air pollution. For example, smokers often go outside to smoke and are thus more likely to be exposed to outdoor NO_2 than non-smokers. In the case-crossover studies, it is crucial to account for these confounders because they vary with time and are not controlled for by study design. Although time-series analyses do not need to adjust for individual risk factors, ecological-level confounders, such as traffic stress and industrial activity, may influence the results (Bukowski, 2008). These important confounders were almost never considered or measured in the draft ISA epidemiology studies of short-term exposure to NO_2 and CV health outcomes, resulting in significant residual confounding.

Overall, the effect estimates for NO_2 were generally small and more likely the result of bias, confounding, or chance than a causal relationship.

4.3.3 Supporting Evidence

4.3.3.1 Heart Rate/Heart Rate Variability

The draft ISA reviews a number of epidemiology studies and several controlled human exposure studies that evaluated NO_2 and heart rate variability (HRV) and concludes "current evidence suggests that among participants with pre-existing or elevated risk for CVD, ambient NO_2 concentrations are associated with alterations in cardiac autonomic control as assessed by indices of HRV." However, the draft ISA does not systematically evaluate study quality or consider inconsistencies in the study findings, which may have impacted causal inference.

Several methodological limitations were present in the epidemiology studies. Selection bias was usually of concern when the participants were recruited from a single clinic or hospital. Exposure measurement errors were often substantial when area-level NO_2 concentrations were used as surrogates for individual exposures.

Also, HRV measurements are highly error-prone (Sandercock *et al*., 2005) and less reliable when taken following exercise interventions (*vs*. when subjects are at rest). In addition, the reproducibility of HRV is poor in populations with pre-existing conditions. Most epidemiology studies conducted in populations with pre-existing conditions reported an association between short-term NO_2 exposure and HRV, while studies of healthy individuals usually reported null findings. Since most studies did not include a power calculation to incorporate such variability and it is unknown whether the measurement errors in HRV were independent, it may be that the observed associations in populations with pre-existing conditions were chance findings.

Confounding by co-pollutants remains a significant source of uncertainty in these studies. Most studies that reported changes in HRV measurements only employed single-pollutant models. Among studies that employed multi-pollutant models, the results were inconsistent with adjustments for co-pollutants. In Timonen *et al*. (2006), adjustment for $PM_{2.5}$ in the two-pollutant model did not change effect estimates of NO_2; however, potential confounding by accumulation model particles or CO, both of which are correlated with NO_2 and show similar or stronger effects on HRV, was not assessed. In Huang *et al*. (2012a), NO_2 was only evaluated as a confounder for $PM_{2.5}$ in two-pollutant models; the study did not report NO_2-specific results. In Zanobetti *et al*. (2010), NO_2 was associated with non-significant increases in r-MSSD and significant decreases in the high frequency component of HRV (HF) when adjusting for $PM_{2.5}$. However, two-pollutant models with adjustment for BC yielded significant increases in r-MSSD

and non-significant increases in HF associated with increased NO_2. Taken together, the reported effects of NO_2 on indices of HRV do not appear to be robust against adjustment for co-pollutants.

Another major limitation of these studies is residual confounding by unmeasured factors. Many mental or psychological factors, such as depression, anxiety, and stress, may cause changes in HRV (Sandercock *et al*., 2005). Certain medications, such as β-blockers and antiarrhythmic drugs, may also influence HRV (European Heart Journal, 1996). These factors are prevalent in populations with previous CV events or existing CV risk factors, and they were usually unaccounted or unadjusted for in the epidemiology studies of air pollution and HRV used in the draft ISA.

These limitations undermine the validity of the findings, and limit the ability to make causal inference on the association between ambient NO_2 and HRV.

Even setting aside these issues, results among the epidemiology studies are not consistent. For example, two studies conducted in patients with coronary artery disease assessed the associations between short-term NO_2 exposures and HRV measurements in two-pollutant models with adjustment for $PM_{2.5}$ (Timonen *et al*., 2006; Zanobetti *et al*., 2010). Timonen *et al*. reported a non-significant inverse association between HF and 5-day average of NO_2, while Zanobetti *et al*. observed a positive association between HF and 3-day average of NO_2.

Results on NO_2 and heart rate are also mixed. Three studies examined heart rate in populations with pre-existing heart conditions or cardiovascular risk factors (Barclay *et al*., 2009; Goldberg *et al*., 2008; Williams *et al*., 2012). Barclay *et al*. and Goldberg *et al*. observed non-significant increases in pulse rate associated with increased NO_2 concentrations among patients with stable heart failure, while Williams *et al*. reported a significant inverse association between heart rate and NO_2 levels in a population with a high prevalence of CV risk factors (*i.e.*, hypertension, diabetes, hyperlipidemia).

In addition, findings from controlled human exposure studies are inconsistent with those from epidemiology studies. Many epidemiology studies showed that increases in ambient NO_2 were associated with decreases in time domain variables of HRV, measures derived from instantaneous heart rate, or intervals between adjacent QRS complexes. However, controlled human exposure studies conducted in healthy adults (Huang *et al*., 2012b) and individuals with coronary heart disease (Scaife *et al*., 2012) did not show significant alterations in HRV time domain intervals following exposure to NO_2 at concentrations four or five times higher than the current NAAQS.

Finally, chronic reduction in HRV predicts CV events. However, it is unclear whether changes in HRV observed in the studies discussed above may persist and progress to adverse CV outcomes. Temporary fluctuations in HRV may be reversible and not clinically significant.

4.3.3.2 Blood Biomarkers

The draft ISA reviews a number of epidemiology studies and a few controlled human exposure studies of blood biomarkers. The most commonly studied biomarkers are those associated with inflammation, oxidative stress, cell adhesion, coagulation, and thrombosis. Very few biomarkers were studied in more than one study, making comparison of the effects of NO_2 on specific biomarkers across studies very difficult.

Several studies evaluated IL-6 and fibrinogen with mixed results. Increases, decreases, and no changes in fibrinogen levels have been reported to be associated with increases in ambient NO_2 concentrations in epidemiology studies. Several epidemiology studies reported a positive association between IL-6 and

ambient NO_2, but controlled human exposure studies failed to observe any statistically significant changes in IL-6 with exposure to NO_2 at concentrations much higher than ambient levels.

An important limitation of biomarker studies is that some of the endpoints analyzed may not be clinically relevant (*i.e.*, associated with disease). That is, many of the biomarkers may not be indicative of or on the pathway to adverse CV effects but rather are part of homeostatic processes. In addition, many of the biomarkers, particularly those associated with inflammation and coagulation, are indicators of an overall systemic response from blood rather than specific local responses, so they may not be related to CV injury. Moreover, many factors related to study design can influence the measured concentrations of circulating biomarkers, including the time of day that blood is collected (due to the influence of circadian patterns), feeding state of participants (such as after a high-fat meal), and level of physical activity of the participants.

Taken as a whole, these studies do not provide adequate supporting evidence for an MoA for CV effects.

4.3.4 Causal Determination

We evaluated the current evidence on the CV effects of short-term exposure to NO_2 using the modified Bradford Hill aspects to aid in judging causality (as outlined in Section 4.2.3).

The majority of effect estimates are small and very close to the null value. Given the various methodological limitations in the epidemiology studies, these estimates are more likely the result of bias, confounding, or chance than a causal relationship between NO_2 and CV health outcomes. This is supported by the fact that findings from epidemiology studies on CV hospitalization and mortality were not consistent.

There are three main sources of bias in NO_2 epidemiology studies: selection bias, exposure measurement error, and outcome misclassification. Selection bias and outcome misclassification were accounted for in some studies but an important source of bias in others. For studies where selection bias or outcome misclassification were identified, the magnitude and direction of the bias was difficult to discern. Exposure measurement error was a major source of bias that likely impacted the findings across all epidemiology studies (Rhomberg *et al*., 2011).

Several known risk factors for acute CV outcomes, such as stress and physical exertion, were often not controlled for in statistical analyses. More importantly, the presence of multiple correlated and confounding co-pollutants makes it virtually impossible to identify the true causal agent for any observed health effect. As noted in the draft ISA, "[u]ntil more reliable methods to adjust for multiple co-pollutants simultaneously become available, there is potential for residual confounding due to unmeasured co-pollutants (Section 1.5). Without consistent and reproducible experimental evidence that is coherent with the effects observed in epidemiologic studies, some uncertainty still exists concerning the role of correlated pollutants in the associations observed with oxides of nitrogen" (US EPA, 2013, p. 250).

Results are also not coherent among epidemiology studies evaluating different outcomes or controlled human exposure studies evaluating HRV and blood biomarkers. Also, the draft ISA proposes that alterations in cardiac autonomic control and increases in oxidative stress and inflammation may be the MoAs for the CV effects of NO_2, but the evidence does not support either of these suggestions.

It is difficult to assess the biological gradient of the potential CV effects of NO_2 because most studies assumed linearity and reported effect estimates associated with increments in NO_2 concentrations. However, the most robust effects of NO_2 were usually observed in Asian countries or certain European

countries with ambient NO_2 that is much higher than in the US, suggesting a potential threshold for CV effects.

Although the majority of the observed effects were for 0- or 1-day lags, which are biologically plausible, there appeared to be inconsistency in reported seasonal patterns of NO_2 effects. The reported associations between NO_2 and CV hospitalization were stronger in colder conditions, while the reported association with CV mortality was more pronounced in summer. It is not clear why this should be the case.

It is also notable that none of the CV effects examined in this analysis are specific to NO_2. Each outcome has several other risk factors that are much more likely to contribute to CV effects that were not fully accounted for in most analyses.

Overall, the conclusion in the draft ISA that there is likely to be a causal relationship between short-term exposure to NO_2 and cardiovascular health effects is not supported by the evidence when evaluated using the modified Bradford Hill aspects.

4.4 Total Mortality

4.4.1 Epidemiology Studies

The draft ISA concludes that there is likely to be a causal relationship between short-term NO_2 exposure and total mortality, based primarily on "the robust evidence observed across various multicity studies." However, the draft ISA did not systematically evaluate study quality or how methodological limitations may have impacted study findings.

All recent multicity studies, except for Chiusolo *et al*. (2011), employed time-series analysis. As discussed above in Section 3.3.2, time-series analyses are ecological in nature, and causal inference on risk at the individual level is often subject to ecological bias. Chiusolo *et al*. (2011) employed a case-crossover study design that allows causal inference at the individual level. However, a serious flaw of the study was the selection of control days. The authors used a time-stratified approach, in which the study period was divided into monthly strata and the controls days were chosen on the same days of the week as the case day in the stratum. Such an approach is biased for mortality studies when days after the case day are selected as controls, because a case cannot be exposed after he or she is deceased.

Exposure measurement error was usually substantial in these studies when city-level NO_2 concentrations were calculated from measurements at a single or a small number of central air monitors. The air monitors generally measure the background ambient NO_2 levels and do not account for emissions from traffic or industrial sources. In the case-crossover study, city-level NO_2 concentrations were used as surrogates for individual exposure, likely resulting in substantial error. There was also likely measurement error of the covariates (*e.g.*, meteorological factors). For example, temperature was often measured at a single location that did not reflect the area-level temperature accurately or the individual exposure to any "heat islands" (Bukowski, 2008).

A considerable source of uncertainty in the time-series studies is confounding. Although reported associations were not measurably confounded by co-pollutants (although residual confounding is possible), many key risk factors were not controlled for. For example, daily stress, at the city or at individual level, was not considered in any of these studies, and may have confounded results.

4.4.2 Causal Determination

We evaluated the current evidence on short-term exposure to NO_2 and total mortality using the modified Bradford Hill aspects to aid in judging causality.

The majority of the effect estimates were small. Given the various methodological limitations in the epidemiology studies, these estimates were more likely the results of bias, confounding, or chance than a causal relationship between NO_2 and total mortality. As discussed above, exposure measurement error and selection bias severely undermined the validity of the findings. Also, although the effects of NO_2 on mortality appeared to be robust against certain co-pollutants, residual confounding by co-pollutants and, more importantly, other unmeasured factors remained an issue in these studies.

Epidemiology studies have generally reported significant overall excess risks of total mortality, but there are inconsistencies both within and across studies. Considerable heterogeneity was present among all studies that reported city-specific effect estimates (Bellini *et al.*, 2007; Berglind *et al.*, 2009; Wong *et al.*, 2008; HEI, 2010) except one (Chen *et al*., 2012). In Bellini *et al.* (2007), the association between NO_2 and total mortality was null or negative in 10 out of 15 Italian cities examined. The overall excess risk was driven by five positive city-specific estimates. In Berglind *et al.* (2009), the association between NO_2 and mortality was null in three out of five European cities examined. The overall excess risk was driven by the estimates for Augsburg and Barcelona. Wong *et al.* (2008) and HEI (2010) reported consistent positive associations in four Asian cities from single-pollutant models. The city-specific associations remained robust against co-pollutants except for Bangkok, for which the association became null with adjustment for PM_{10}.

The majority of the observed effects were for 0- and 1-day lags, which is biologically plausible. Also, the draft ISA proposes that the CV and respiratory effects of NO_2 may lead to CV and respiratory mortality that account for over 40% of total mortality. However, as discussed above, the evidence does not support a causal relationship between NO_2 and these health effects. It was also noted in the draft ISA that the evidence "does not clearly describe a biologically plausible mechanism to support NO_2-induced cardiovascular mortality" and "the biological mechanism that explains the continuum of effects that could lead to respiratory-related mortality also remains unclear."

It is difficult to assess the biological gradient of any effects of NO_2 on mortality because most studies assumed response linearity and reported effect estimates associated with increments in NO_2 concentrations. Three studies presented exposure-response curves (Moolgavkar *et al.*, 2013; Chen *et al.*, 2012; Wong *et al.*, 2008; HEI, 2010). For studies that were conducted in Asia (Chen *et al.*, 2012; Wong *et al.*, 2008; HEI, 2010), linear or J-shaped exposure-response curves were presented with adverse mortality outcomes at concentrations exceeding the current NAAQS. Moolgavkar *et al.* (2013), which studied 72 cities in the US, reported an exposure-response curve based on estimates from single-pollutant models. The effect of NO_2 on total mortality derived from the exposure-response curve may be exaggerated because the point estimate for NO_2 decreased by 40% when adjusting for co-pollutants. In addition, the association between NO_2 and total mortality in Berglind *et al.* (2009) was null in the two cities (Helsinki and Stockholm) with the lowest ambient NO_2 levels and the city (Rome) with the highest ambient NO_2 levels.

Finally, total mortality encompasses many disease processes, and it is certainly not specific to NO_2.

The conclusion in the draft ISA that there is likely to be a causal relationship between short-term NO_2 exposure and total mortality is not supported by the evidence when evaluated using the modified Bradford Hill aspects.

5 Long-term Exposure

5.1 Introduction

For respiratory effects, the 2008 ISA concluded that evidence was *suggestive*, but not sufficient to infer a causal relationship with long-term NO_2 exposure. The current draft ISA concludes that there is *likely to be a causal relationship* between long-term NO_2 exposure and respiratory effects, based primarily on epidemiology studies regarding asthma incidence in children. Overall, however, these epidemiology studies do not provide sufficient evidence that long-term NO_2 exposure is associated with respiratory effects. In addition, there is on-going uncertainty regarding whether the observed effects are due to TRP, rather than NO_2 *per se*.

The 2008 ISA reported that evidence was *inadequate to infer the presence or absence of a causal relationship* between long-term NO_2 exposure and cardiovascular effects, reproductive and developmental effects, total mortality, and cancer. In contrast, with the addition of a few new studies per endpoint, the 2013 draft ISA indicates that evidence for all of these is *suggestive of a causal association*.[9] Based on the review of these endpoints in the draft ISA, however, it appears that the evidence may be below equipoise in all cases, indicating the criteria used in the draft ISA to classify evidence as suggestive instead indicate the evidence is inadequate to draw a conclusion regarding causality.

Below, we evaluate the evidence regarding respiratory effects using the modified Bradford Hill aspects to aid in judging causality and briefly discuss issues with the draft ISA evaluation of endpoints classified as *suggestive of a causal association*.

5.2 Respiratory Effects

5.2.1 Epidemiology Studies of Asthma Incidence in Children

Key studies discussed in the draft ISA to support the causal determination include both large prospective studies (McConnell *et al*., 2010; Jerrett *et al*., 2008; Gehring *et al*., 2010; Clark *et al*., 2010; Gruzieva *et al*., 2013; Lee *et al*., 2012; Shima *et al*., 2002; Nishimura *et al*., 2013; Oftedal *et al*., 2009; Carlsten *et al*., 2011; Clougherty *et al*., 2007) and several new cross-sectional studies. Of these, the draft ISA appears to rely more heavily on findings from the prospective studies; we summarize these studies below.

McConnell *et al*. (2010) conducted a study using the Children's Health Study (CHS) cohort of Southern California communities. The authors evaluated 120 children 4-8 years old enrolled between 2002-2003 and followed them for three years. The authors assessed associations between new-onset asthma with annual average NO_2 concentrations from central-site community monitors and TRP estimated at school and home using a dispersion model. The authors reported a significant association between NO_2 and asthma onset (HR = 2.18, 95% CI: 1.18-4.0), but the estimate was significantly attenuated with the inclusion of TRP (HR = 1.37, 95% CI, 0.69-2.71).

[9] The draft ISA divides reproductive and developmental effects into the categories of fertility, reproduction, and pregnancy; birth outcomes; and postnatal development.

Jerrett *et al*. (2008) also conducted a study using the prospective CHS cohort. The authors evaluated associations between new-onset asthma and TRP in children 10-18 years of age living in 11 of the 12 California communities that were part of the CHS study. The study measured local NO_2 concentrations outside the children's homes in winter and summer using Palmes tubes.[10] The authors evaluated both within- and between-community asthma health risks to assess independent contributions of local *vs*. regional NO_2. Within-community incident asthma risk was not statistically significantly associated with annual average NO_2 (HR = 1.31, 95% CI: 0.76-2.26 per IQR of NO_2), whereas between-community estimates were significant (HR = 1.28, 95% CI: 1.05-1.57). The authors noted that, because the within- and between-community estimates were similar in magnitude, this may indicate that both regional and local air pollutants contribute to the observed association. They further suggested that this provides evidence that NO_2 may be a surrogate for a mixture of regional and local traffic-related air pollutants.

Gehring *et al*. (2010) analyzed more than 3,000 children from the Prevention and Incidence of Asthma and Mite Allergy (PIAMA) study in the Netherlands to assess effects of TRP on asthma, allergy, and related symptoms. The study estimated annual levels of NO_2 and other traffic-related air pollutants at the address of each participant using a land-use regression model. The authors reported no association between NO_2 exposures and bronchial hyperresponsiveness or allergy. Results for asthma symptoms were not statistically significant, but statistically significant effects were reported for asthma incidence, asthma prevalence, and wheeze.

Gruzieva *et al*. (2013) assessed associations between asthma symptoms, wheeze, and incident asthma at 12 years of age in a Swedish cohort that was part of the Children, Allergy, Milieu, Stockholm Epidemiology Survey (BAMSE). The authors used dispersion models to estimate NO_2 and PM_{10} concentrations for all addresses for the years 1994-2008. The authors reported no association between asthma symptoms and exposures after the first year of life. Only wheeze (three or more episodes) at 12 years of age was significantly associated with NO_2. For incident asthma, the authors reported an effect that was not statistically significant (OR = 1.87, 95% CI: 1.0-3.44) per 46.8 $\mu g/m^3$ of NO_2.

A study by Lee *et al.* (2012) examined pulmonary function and respiratory diseases, including asthma, in Taiwan by comparing rates in low and high NO_2 concentration areas. The authors reported an incidence rate ratio of asthma that was nominally higher, but still null, in communities with higher NO_2 concentrations (RR = 0.97, 95% CI: 0.85-1.11) compared to communities with lower NO_2 concentrations (RR = 0.73, 95% CI: 0.64-083).

Carlsten *et al.* (2011) evaluated the association between early exposure to TRP (estimated using land-use regression models) and new-onset asthma in a high-risk birth cohort of 7-year-olds. Associations between new onset-asthma and NO and NO_2 exposure were not statistically significant (OR = 1.2, 95% CI: 0.9-1.7 and OR = 1.5, 95% CI: 0.9-2.5 per IQR of NO and NO_2, respectively).

The study by Oftedal *et al*. (2009) assessed the association between TRP and new-onset asthma in a retrospective study of 9- to 10-year-old children who were part of the Oslo Norway birth cohort. The study estimated lifetime exposures, based on residential address, using dispersion modeling. The authors reported no association between exposure to NO_2 in the first year of life (averaged from birth) to asthma onset or from previous year's exposure and asthma onset.

Shima *et al*. (2002) evaluated approximately 1,900 children in eight communities in Japan. Central-site ambient air monitors were used to estimate NO_2 exposures. The authors reported a significant incidence

[10] Palmes tubes are passive diffusion tubes that are used to measure concentrations of gaseous air pollutants. The tubes are about the size and shape of a pencil and are made of glass and contain a chemically treated substrate that reacts with the air pollutant.

of asthma associated with NO_2 exposures (OR = 3.62, 95% CI: 1.11-11.87) when comparing the least and most polluted communities.

More than 4,000 participants (8-21 years old) were investigated for asthma onset in a multi-city study that included Chicago, IL, Bronx, NY, Houston, TX, San Francisco, CA, and Puerto Rico (Nishimura *et al*., 2013). The study estimated exposure based on inverse distance-squared weighted averages from closest area monitors to the residences. The combined OR for asthma was 1.17 (95% CI: 1.04-1.31) for NO_2 exposures. Heterogeneity in asthma risks was observed across cities, with a significant risk observed for San Francisco but none of the other cities.

A case-control study conducted in British Columbia in a population of 3- to 4-year-old children (Clark *et al*., 2010) evaluated associations between incidence of asthma and exposure to air pollutants (CO, NO, NO_2, PM_{10} and $PM_{2.5}$, SO_2, BC, wood smoke, and ozone). The study determined exposures for gestational period and first-year of life from high-resolution modeled concentrations based on monitor data, as well as land-use regression models to account for temporal variability in pollutant concentrations. The authors reported increased risks of asthma diagnosis for CO, NO, NO_2, PM_{10}, SO_2, and BC. The adjusted OR was 1.08 (95% CI: 1.04-1.12) for a 10 $\mu g/m^3$ increase of NO and 1.12 (95% CI = 1.07-1.17) for a 10 $\mu g/m^3$ increase in NO_2.

In a study in East Boston, MA, Clougherty *et al*. (2007) evaluated the risk of asthma onset in 413 children. The study used a geographic information system (GIS) model to estimate exposure retrospectively. The authors reported significantly increased risk of asthma per 4.3 ppb increase in NO_2, but only among children with elevated exposure to violence (OR = 1.63, 95% CI: 1.14-2.33).

5.2.2 Causal Determination

We evaluated the current evidence on long-term exposure to NO_2 and respiratory effects using the modified Bradford Hill aspects to aid in judging causality.

The majority of the effect estimates from prospective studies were small. While the prospective study design allows for control of many individual potential confounders, effects of this magnitude may be due to chance, bias, or unmeasured or residual confounding (*e.g.*, see Boffetta *et al*., 2008; Fewell *et al*., 2007).

In studies of asthma, a major source of bias is outcome misclassification. Most of the long-term studies relied on questionnaires to assess asthma occurrence, but these data may be affected by access to care and differential diagnostic practices by physicians. Asthma incidence was also evaluated in different ways. Some studies used a pediatric allergist (*e.g.*, Carlsten *et al.*, 2011), which is likely to be associated with reduced outcome misclassification. In other studies (*e.g.*, Gruzieva *et al.*, 2013), however, asthma was defined as having a certain number of episodes of wheeze in the last year or a combination of one episode of wheeze and prescriptions for inhaled corticosteroids. In these studies, there is a much higher chance of exposure misclassification. In addition, it is generally more difficult to diagnose asthma in infants or very young children (Gruzieva *et al*., 2013). As noted by Braback and Forsberg (2009): "Misclassification of the outcome is often a concern in questionnaire-based studies of respiratory disease, particularly in young children with less distinctive symptoms. Lower respiratory illness and wheezing symptoms are frequent in early childhood but very few of these children develop asthma." Outcome measurement error is highlighted in studies that show inconsistencies across different measures of asthma; for example, Gruzieva *et al*. (2013) reported marginally significant effects for wheeze (*e.g.*, > 3 episodes) associated with NO_x exposure in a cohort followed from birth to 12 years, whereas doctor-diagnosed asthma was not associated with NO_x (*e.g.*, Gruzieva *et al*., 2013).

Exposure measurement error is another significant source of bias in epidemiology studies. NO_x data for the long-term studies were collected in various ways, including from central-site monitors and land-use regression or air dispersion models, to estimate residential concentrations. All of these methods introduce exposure measurement error to varying degrees, and this source of bias is largely ignored in epidemiology investigations. For example, the use of Palmes tubes is subject to a positive bias in the measurement of NO_2 concentrations (Jimenez *et al*., 2011 as cited in US EPA, 2013). Similarly, dispersion models that include NO_x chemistry could potentially overestimate NO_2 concentrations, and central-site monitors can fail to capture spatial variability in the concentrations. EPA notes that these biases would tend to bias estimates toward the null; however, these errors may bias estimates in either direction (Rhomberg *et al*., 2011).

Confounding is likely the most important source of uncertainty in the epidemiology studies of long-term NO_x exposure relied upon in the draft ISA. Most studies did not control for co-pollutants or other important confounders such as smoking (*e.g.*, Jerrett *et al*., 2008; Carlsten *et al*., 2011; Gruzieva *et al*., 2013), and those that considered multi-pollutant models found attenuated effects. For example, McConnell *et al.* (2010) evaluated childhood incident asthma and TRP using dispersion modeling and NO_2 from central-site monitors. The authors reported an association with NO_2 and new-onset asthma per 10 ppb increase in NO_2 (HR = 1.39, 95% CI: 1.07-1.80) but, in models with both NO_2 and TRP, the estimate with NO_2 was attenuated (HR = 1.14, 95% CI: 0.85, 1.53); this indicates confounding by TRP. In fact, many of the studies suggest that NO_x is likely a proxy for other traffic-related air pollutants that are the more likely causal agents for observed occurrence of asthma:

- "Although some toxicologic evidence supports the potential for NO_2 as a direct causative agent for respiratory health effects at high concentrations…the more likely interpretation is that NO_2 was a proxy for other components of the local air pollution mixture with high toxic potential" (Jerrett *et al*., 2008);
- "In the current study, the attenuated association between asthma and NO_2 continuously measured at the community monitor in models with adjustment for TRP also suggests that NO_2 was not causally related to asthma. However, interpretation of this result is not clear because exposure measured at the community monitor may have misclassified exposures of children in parts of the community with significant local sources of NO_2" (McConnell *et al*., 2010); and
- "NO_2 is not considered a causal agent for respiratory effects, but may in this study act as a general assessment of traffic-related pollution" (Oftedal *et al*., 2009).

As noted in the draft ISA, association between long-term NO_2 and increases in asthma prevalence and incidence were inconsistent. For example, although results from the CHS in California (Jerrett *et al*., 2008) and a cohort in the Netherlands (Gehring *et al*., 2007) found positive associations, many other studies reported null effects (*e.g.*, Clougherty *et al*., 2007; Oftedal *et al*., 2009; Gruzieva *et al*., 2013; Lee *et al*., 2012).

With regard to biological plausibility, as discussed in the draft ISA, there is limited evidence from studies in laboratory animals regarding a mechanism of action for the associations observed in the epidemiology studies. Notably, effects in animals were observed only at concentrations substantially higher (*e.g.*, $\geq$ 1,000 ppb) than ambient NO_2 concentrations. As discussed in Section 4.2.3 above, certain metabolites of NO_2 are beneficial to the respiratory system, and these are not incorporated into the proposed MoAs of NO_2 in the draft ISA.

As noted in the draft ISA, the shape of the concentration-response relationship has not been fully evaluated in long-term NO_2 studies. Therefore, there is little evidence for determining whether there is a biological gradient for NO_2 respiratory effects.

Specificity is difficult to establish for respiratory effects, as there are numerous risk factors for asthma. Establishing temporality is also challenging in long-term studies. The draft ISA notes that epidemiology evidence appears to be stronger for early life exposures but it is not necessarily clear that this is the case, as there is heterogeneity in the findings from epidemiology studies with regard to effects from early life exposures.

Overall, the conclusion in the draft ISA that evidence is sufficient to conclude there is a causal relationship between long-term NO_2 exposure and respiratory health effects is not supported by the recent studies it relies upon. Considering the modified Bradford Hill aspects, NO_2 epidemiology studies do not demonstrate a clear causal link between long-term NO_2 exposure and respiratory health effects, and many epidemiologists have suggested that NO_2 may be a surrogate marker for other air pollution components.

5.3 Cardiovascular Effects

The draft ISA emphasizes several studies that reported statistically significant associations with various cardiovascular morbidity and mortality endpoints but indicates, "[a]lthough epidemiologic studies generally considered important confounders, studies with similar designs and approaches for control of confounding did not consistently report associations with cardiovascular effects." The draft ISA also concludes that there was limited toxicology evidence at relevant exposures and weak evidence for key events to inform a mode of action. The draft ISA should discuss whether studies indicating no effect are of higher quality than those that reported associations, particularly in light of the limited toxicology evidence, to determine whether the weight of evidence regarding causation is at or above equipoise.

5.4 Reproductive and Developmental Effects

The draft ISA states, "There is limited recent evidence for NO_x-related effects on fertility, reproduction, and pregnancy indicated as increases in pre-eclampsia. There is limited recent evidence for NO_x-related effects on birth outcomes indicated as fetal growth restriction. For postnatal development, the strongest evidence is for partially irreversible decreases in lung function growth in children with less consistent evidence for decreases in cognitive function in children." It also states, "Previous epidemiologic evidence was limited and inconsistent for effects on birth outcomes and there was weak evidence in experimental animals to provide biological plausibility. This weak biological plausibility remains an uncertainty for all three subcategories of reproductive and developmental effects, as there are no recent animal toxicological studies to consider. Another uncertainty that pertains to all three subcategories is the lack of association found in some recent high-quality epidemiologic studies."

The draft ISA does not provide sufficient justification that new data change the balance of the weight of evidence regarding causation from inadequate to suggestive. In addition, the draft ISA makes conclusions about these broad reproductive and developmental categories based on one endpoint each (*e.g.*, fertility, low birth weight, gestational age), even though the weight of the evidence does not appear to be above equipoise for these endpoints and, as acknowledged in the draft ISA, is below equipoise for other endpoints.

5.4.1 Fertility, Reproduction, and Pregnancy

The draft ISA concludes that there is limited but consistent epidemiology evidence for associations between NO_2 or NO_x and pre-eclampsia, and that evidence for an association with pregnancy-induced hypertension is inconsistent. Regarding pre-eclampsia, the draft ISA references three studies that show associations and one that does not (van den Hooven *et al.*, 2011, as cited in US EPA, 2013); the draft ISA does not discuss the quality or relative weight of these studies. It is not clear whether the trimester in which exposure occurred is consistent among studies. Also, the draft ISA indicates that evidence for a linear concentration-response relationship is limited and, although there are generally high correlations between NO_x and $PM_{2.5}$ exposures, pre-eclampsia has not been evaluated in co-pollutant models. Also, pregnancy-induced hypertension is generally an early sign of pre-eclampsia; the lack of a consistent association for this endpoint indicates a lack of coherence among the data and that the weight of evidence does not support NO_x as a causal factor for pre-eclampsia.

5.4.2 Birth Outcomes

The draft ISA states that epidemiology studies found associations between higher prenatal NO_x exposure and fetal growth restriction, but associations with other birth outcomes (*e.g.*, preterm birth or low birth weight) are inconsistent. The draft ISA should discuss whether fetal growth restriction is expected to lead to or be associated with other outcomes and whether these outcomes are observed to be associated with NO_x. If not, this may shift the weight of evidence below equipoise.

In addition, the draft ISA indicates that the evidence is "less certain when it comes to assessing the time period of pregnancy when exposure to NO_x is associated with the highest risks. Some studies find the highest risks associated with NO_x when NO_x is measured in early pregnancy, while in other studies the time period associated with the greatest risk is toward the end of pregnancy. Others find the greatest risk when exposure is assigned for the entire pregnancy period" (US EPA, 2013). If studies find risks differ by stage of pregnancy, this constitutes inconsistent evidence; it is not biologically plausible that NO_x increases risk at different stages of pregnancy in different women.

5.4.3 Postnatal Development

The draft ISA evaluates evidence from recent epidemiology studies and, to a lesser degree, animal toxicological studies for three major postnatal development outcomes: neurodevelopment, developmental respiratory effects, and early life mortality. Findings from epidemiology studies were inconsistent for most neurodevelopmental endpoints and early life mortality. In addition to inconsistency, there was a lack of coherence between epidemiology and animal toxicological studies. With regard to developmental respiratory effects, exposure to NO_2 did not alter lung function in experimental animals, especially juvenile ones. The draft ISA should take into account the inconsistent results and lack of biological plausibility when relying on these studies to make determinations regarding causality.

Further, the draft ISA does not fully evaluate study quality or incorporate deficiencies in study quality into the causal determination process. For example, in the review of epidemiology studies regarding cognitive function, the draft ISA notes that "[w]hile some studies considered birth outcomes and noise exposure, none of the studies considered polycyclic aromatic hydrocarbons or lead, both of which are well-characterized risk factors for neurodevelopmental decrements." The draft ISA, however, does not address how such omissions would impact the results or introduce uncertainty in the causal determinations. For autism, the draft ISA concludes that "[a]utism risk increased with exposure to NO_2 during pregnancy and the first year of life, with more pronounced risk found for late gestation and early

life exposure." This conclusion is based on findings from two case-control studies, which are known to be vulnerable to selection bias, yet the draft ISA provides no discussion regarding potential methodological issues.

5.5 Total Mortality

The draft ISA indicates that there are several epidemiology studies indicating positive associations between long-term NO_x exposure and total mortality, but there are no such associations in the American Cancer Society (ACS) or California Seventh-day Adventists cohort (AHSMOG) cohorts and limited coherence with respiratory and cardiovascular morbidity endpoints. The majority of mortality studies included in the draft ISA are cohort studies that suffered several common methodological limitations. Control for key confounders (*e.g.*, smoking, physical activity, and SES) was generally inadequate or incomplete. GIS-facilitated modeling was often employed to interpolate individual exposure to NO_2, but the accuracy and validity of the models were seldom addressed. Confounding effects of co-pollutants also contributed to the uncertainty of the results, as some observed associations between NO_2 and total mortality became null when adjusting for co-pollutants. Exposure estimates based on stationary air monitor measurements and residential addresses usually did not account for indoor exposure, or exposure while commuting. Loss to follow-up was substantial in some cohort studies. These methodological issues have considerable impact on the validity of the findings but were not considered in the draft ISA.

The draft ISA should evaluate the evidence across studies, including ACS and AHSMOG, to determine whether the reported associations are more likely to be indicative of causation or attributable to other factors (*e.g.*, confounding, bias). This evaluation should be based on an evaluation of all relevant data, including other epidemiology, toxicology, and MoA studies. If it appears more likely than not that the positive associations are indicative of causation, an above equipoise/suggestive causal classification is warranted. Otherwise, a below equipoise/inadequate classification is more appropriate.

5.6 Cancer

The draft ISA indicates that evidence is suggestive of a causal relationship between long-term NO_x exposure and cancer based primarily on associations between ambient NO_x or NO_2 concentrations and lung cancer in some previous and recent high-quality epidemiology studies, but not in other studies of comparable quality. However, as discussed above, the epidemiology studies considered to be of high-quality in the draft ISA were not without limitations. One of the key issues that may impact the association with lung cancer is confounding by smoking and SES, as these two important risk factors were either uncontrolled or inadequately adjusted for in the statistical analyses. The draft ISA also indicates that biological plausibility of a causal relationship with cancer is limited. There are findings for lung tumor promotion and hyperplasia of lung epithelial cells with NO_2 exposure, some at higher than ambient-relevant concentrations, but no evidence for direct effects on carcinogenesis. Because the evidence is not consistent and there is uncertainty owing to limited biological plausibility, these data should be systematically reviewed to determine whether it is more likely than not that NO_x is carcinogenic.

6 "At-risk Factors"

Chapter 6 of the draft ISA evaluates what it calls "at-risk factors." EPA describes these as factors that "potentially lead to some populations or lifestages being at increased or decreased risk of air pollution-related health effects" (US EPA, 2013, p. lii). These include both intrinsic (*e.g.*, genetic or developmental factors, race, sex, life-stage, the presence of pre-existing diseases) and extrinsic, non-biological factors (*e.g.*, SES, occupation). An "at-risk factor" is more accurately described as an effect modifier, which is a technical term defined in epidemiology as a variable that differentially modifies the observed effect of a risk factor (in this case, NO_2) on disease status. Because "at-risk factors" are evaluated in the draft ISA based on epidemiology studies, it may be more appropriate for the draft ISA to refer to them as effect modifiers, as we do herein.

The draft ISA states that effect modifiers are identified in studies that conduct stratified analyses and then, "[b]uilding on the causal framework…conclusions are reached regarding the strength of evidence across scientific disciplines for each at-risk factor that may contribute to increased or decreased risk of an air pollutant-related health effect." The draft ISA states that the modified Bradford Hill aspects are considered in drawing conclusions regarding whether there is *adequate evidence*, *suggestive evidence*, *inadequate evidence*, or *evidence of no effect*.

Although EPA's frameworks for both causal determination (Table 2.2) and classification of evidence for potential "at-risk factors" (Table 2.3) are based on consideration of modified Bradford Hill aspects, 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 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 causal framework; there does not appear to be an equivalent to the *likely to be a causal relationship*. The issues with the weight of evidence for causal determination (discussed above in Section 2) apply to the effect modifier classifications as well. For example, the draft 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 weight of evidence is below equipoise. In these circumstances, the evidence is inadequate to make a determination as to whether a factor is an effect modifier for a health endpoint.

The draft ISA provides no explanation for why the framework for effect modifiers differs from that for causal determination. Although one is an assessment of direct causation and the other an assessment of factors that can contribute to (or prevent) causation from NO_x, in both cases, the goal is to critically, systematically, and transparently review the weight of scientific evidence. 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. The draft ISA should adopt the IOM (2008) recommended categories for the level of evidence for causation, which consider whether the weight of evidence is above or below equipoise.

The draft ISA Preamble states: "To identify at-risk factors that potentially lead to some populations or lifestages being at increased or decreased risk of air pollution-related health effects, the evidence is systematically evaluated across relevant scientific disciplines (*i.e.*, exposure sciences, dosimetry, toxicology, and epidemiology)." It then describes the types of studies that are evaluated and goes on to discuss how the strength of evidence is evaluated. There is no indication of how the evaluation was

systematic and what was done to ensure that all studies were evaluated in the same manner. Study quality and relevance are not discussed, and a review of Chapter 6 indicates that these factors were not considered. Chapter 6 briefly summarizes study results and makes effect modifier classifications based on whether results are consistent or not, regardless of study quality and the relative weight of studies from which they were derived. Even if a study is discussed in another part of the draft ISA, a critical review that includes an evaluation of study quality and its impact on the interpretation of results also must be considered in the evaluation of effect modifiers.

The draft ISA should indicate whether studies that indicate people at different lifestages are at increased risk for health effects related to NO_x exposure are of sufficient quality to be considered "adequate evidence." The draft ISA should also evaluate study quality and the relative weights of studies evaluating genetic background, pre-existing asthma or COPD, SES, sex, and diet to determine whether studies that are of higher quality and carry more weight indicate whether factors are more likely than not to be effect modifiers.

Finally, it is unclear how the evaluation of effect modifiers will be applied to setting the NAAQS for NO_x. The Clean Air Act requires the EPA Administrator to set NAAQS at the maximum ambient air level that, in his or her judgment, are "requisite to protect the public health," including for sensitive populations. Presumably, Chapter 6 is intended to identify effect modifiers that will then define sensitive, susceptible, or vulnerable populations, yet all of the studies regarding effect modifiers are already considered in the evaluation of specific health endpoints – *i.e.*, these potential populations are already considered. Thus, it is not evident how or if the effect modifier evaluation from the draft ISA should be applied to establishing the NAAQS.

In conclusion, the classification of evidence for potential effect modifiers should use the same framework as the causal framework, and both should be revised to four categories: *causal*, *equipoise and above*, *below equipoise*, and *not causal*. EPA should conduct a systematic review that considers study quality and relevance and weighs studies based on both before coming to conclusions about potential effect modifiers, and explain the bearing of this assessment on the NAAQS.

7 Conclusions

The causal determinations in the draft ISA are stronger than they were in the 2008 ISA for several endpoints (Table 1.1); however, there is insufficient evidence to support this increased confidence in causality. For relationships the draft ISA concludes are causal (short-term NO_2 exposure and respiratory effects) and likely causal (short-term NO_2 exposure and cardiovascular effects and total mortality, long-term NO_2 exposure and respiratory effects), associations in epidemiology studies were close to the null and more likely explained by chance, bias (*e.g.*, exposure or outcome misclassification), and/or confounding (*e.g.*, by co-pollutants, smoking, SES). In some cases, associations were not consistent either within or among epidemiology studies or coherent with controlled exposure studies. For several reported short-term effects, associations were found at lag times that do not appear to be biologically plausible. For numerous endpoints, a mode or mechanism of action was not established or, if present, lacked biological plausibility. It was often difficult to assess the biological gradient of the effects of NO_2 because most studies assumed response linearity and reported effect estimates associated with increments in NO_2 concentrations. There is also uncertainty regarding whether observed associations are due to NO_2 *per se* or whether NO_2 is a surrogate for another pollutant or mixture of pollutants.

Although EPA concludes that evidence for long-term NO_2 exposure is suggestive of causation for cardiovascular effects, reproductive and developmental effects, total mortality, and cancer, the evidence presented in the draft ISA does not indicate that long-term NO_2 exposure is more likely than not to be a causal factor for these effects. Even if new, high-quality studies consistently demonstrated statistically significant associations (which they do not), the results of all other relevant studies must be considered to determine whether these associations are likely indicative of causation. If one cannot resolve inconsistencies among the data, it is inappropriate to conclude the data are suggestive of an association; rather, the weight of evidence indicates that the data are inadequate for drawing conclusions.

It appears that the increased confidence in causality for these endpoints (as compared to the 2008 ISA) is based on the draft ISA's application of its causal framework. The framework does not provide enough guidance so that studies are evaluated in a consistent manner using well-specified criteria in the draft ISA. Also, the causal framework does not require the draft ISA to determine whether, as a whole, the data constitute evidence for causation or are more likely indicative of an alternative hypothesis (*e.g.*, that chance, bias, or confounding account for observed statistical associations). Similarly, the draft ISA's framework for classifying potential effect modifiers is also not sufficient for assessing causality (and it is not clear how the assessment of effect modifiers bears on the NAAQS).

Overall, the evidence presented in the draft ISA does not indicate increased confidence in any causal classification. While not necessarily agreeing with the causal classifications determined by the 2008 ISA, Gradient's opinion is that the draft ISA offers no basis to strengthen any of the classifications.

References

Andersen, ZJ; Wahlin, P; Raaschou-Nielsen, O; Ketzel, M; Scheike, T; Loft, S. 2008. "Size distribution and total number concentration of ultrafine and accumulation mode particles and hospital admissions in children and the elderly in Copenhagen, Denmark." *Occup. Environ. Med.* 65(7):458-466.

Avol, EL; Linn, WS; Peng, RC; Valencia, G; Little, D; Hackney, JD. 1988. "Laboratory study of asthmatic volunteers exposed to nitrogen dioxide and to ambient air pollution." *Am. Ind. Hyg. Assoc. J.* 49(4):143-149.

Avol, E; Linn, WS; Peng, RC; Whynot, JD; Shamoo, DA; Little, DE; Smith, MN; Hackney, JD. 1989. "Experimental exposures of young asthmatic volunteers to 0.3 ppm nitrogen dioxide and to ambient air pollution." *Toxicol. Ind. Health* 5(6):1025-1034.

Ballester, F; Rodriguez, P; Iniguez, C; Saez, M; Daponte, A; Galan, I; Taracido, M; Arribas, F; Bellido, J; Cirarda, FB; Canada, A; Guillen, JJ; Guillen-Grima, F; Lopez, E; Perez-Hoyos, S; Lertxundi, A; Toro, S. 2006. "Air pollution and cardiovascular admissions association in Spain: Results within the EMECAS project." *J. Epidemiol. Community Health* 60(4):328-336.

Barck, C; Sandstrom, T; Lundahl, J; Hallden, G; Svartengren, M; Strand, V; Rak, S; Bylin, G. 2002. "Ambient level of NO_2 augments the inflammatory response to inhaled allergen in asthmatics." *Respir. Med.* 96(11):907-917.

Barclay, JL; Miller, BG; Dick, S; Dennekamp, M; Ford, I; Hillis, GS; Ayres, JG; Seaton, A. 2009. "A panel study of air pollution in subjects with heart failure: Negative results in treated patients." *Occup. Environ. Med.* 66:325-334.

Barnett, AG; Williams, GM; Schwartz, J; Best, TL; Neller, AH; Petroeschevsky, AL; Simpson, RW. 2006. "The effects of air pollution on hospitalizations for cardiovascular disease in elderly people in Australian and New Zealand cities." *Environ. Health Perspect.* 114(7):1018-1123.

Bauer, MA; Utell, MJ; Morrow, PE; Speers, DM; Gibb, FR. 1986. "Inhalation of 0.30 ppm nitrogen dioxide potentiates exercise-induced bronchospasm in asthmatics." *Am. Rev. Respir. Dis.* 134(6):1203-1208.

Bell, ML; Levy, JK; Lin, Z. 2008. "The effect of sandstorms and air pollution on cause-specific hospital admissions in Taipei, Taiwan." *Occup. Environ. Med.* 65(2):104-111.

Bellini, P; Baccini, M; Biggeri, A; Terracini, B. 2007. "The meta-analysis of the Italian studies on short-term effects of air pollution (MISA): Old and new issues on the interpretation of the statistical evidences." *Environmetrics* 18(3):219-229.

Berglind, N; Bellander, T; Forastiere, F; von Klot, S; Aalto, P; Elosua, R; Kulmala, M; Lanki, T; Lowel, H; Peters, A; Picciotto, S; Salomaa, V; Stafoggia, M; Sunyer, J; Nyberg, F. 2009. "Ambient air pollution and daily mortality among survivors of myocardial infarction." *Epidemiology* 20(1):110-118.

Bhaskaran, K; Hajat, S; Armstrong, B; Haines, A; Herrett, E; Wilkinson, P; Smeeth, L. 2011. "The effects of hourly differences in air pollution on the risk of myocardial infarction: Case crossover analysis of the MINAP database." *BMJ* 343:d5531.

Boffetta, P; McLaughlin, JK; La Vecchia, C; Tarone, RE; Lipworth, L; Blot, WJ. 2008. "False-positive results in cancer epidemiology: A plea for epistemological modesty." *J. Natl. Cancer Inst.* 100:988-995.

Boulet, LP; Legris, C; Turcotte, H. 1987. "Bronchial responsiveness to histamine after repeated exercise-induced bronchospasm." *Respiration* 52(4):237-245.

Braback, L; Forsberg, B. 2009. "Does traffic exhaust contribute to the development of asthma and allergic sensitization in children: Findings from recent cohort studies." *Environ. Health* 8:17.

Bukowski, J. 2008. "Do pollution time-series studies contain uncontrolled or residual confounding by risk factors for acute health events?" *Regul. Toxicol. Pharmacol.* 51:135-140.

Bylin, G; Hedenstierna, G; Lindvall, T; Sundin, B. 1988. "Ambient nitrogen dioxide concentrations increase bronchial responsiveness in subjects with mild asthma." *Eur. Respir. J.* 1(7):606-612.

Carlsten, C; Dybuncio, A; Becker, A; Chan-Yeung, M; Brauer, M. 2011. "Traffic-related air pollution and incident asthma in a high-risk birth cohort." *Occup. Environ. Med.* 68(4):291-295.

Chan, CC; Chuang, KJ; Chen, WJ; Chang, WT; Lee, CT; Peng, CM. 2008. "Increasing cardiopulmonary emergency visits by long-range transported Asian dust storms in Taiwan." *Environ. Res.* 106(3):393-400.

Chen, R; Chu, C; Tan, J; Cao, J; Song, W; Xu, X; Jiang, C; Ma, W; Yang, C; Chen, B; Gui, Y; Kan, H. 2010. "Ambient air pollution and hospital admission in Shanghai, China." *J. Hazard. Mater.* 181(1-3):234-240.

Chen, R; Samoli, E; Wong, CM; Huang, W; Wang, Z; Chen, B; Kan, H. 2012. "Associations between short-term exposure to nitrogen dioxide and mortality in 17 Chinese cities: The China Air Pollution and Health Effects Study (CAPES)." *Environ. Int.* 45:32-38.

Cheng, MF; Tsai, SS; Yang, CY. 2009. "Air pollution and hospital admissions for myocardial infarction in a tropical city: Kaohsiung, Taiwan." *J. Toxicol. Environ. Health A* 72(19):1135-1140.

Chiusolo, M; Cadum, E; Stafoggia, M; Galassi, C; Berti, G; Faustini, A; Bisanti, L; Vigotti, MA; Dessì, MP; Cernigliaro, A; Mallone, S; Pacelli, B; Minerba, S; Simonato, L; Forastiere, F. 2011. "Short term effects of nitrogen dioxide on mortality and susceptibility factors in ten Italian cities: The EpiAir Study." *Environ. Health Perspect.* 119(9):1233-1238.

Clark, NA; Demers, PA; Karr, CJ; Koehoorn, M; Lencar, C; Tamburic, L; Brauer, M. 2010. "Effect of early life exposure to air pollution on development of childhood asthma." *Environ. Health Perspect.* 118(2):284-290.

Clougherty, JE; Levy, JI; Kubzansky, LD; Ryan, PB; Suglia, SF; Canner, MJ; Wright, RJ. 2007. "Synergistic effects of traffic-related air pollution and exposure to violence on urban asthma etiology." *Environ. Health Perspect.* 115(8):1140-1146.

Delfino, RJ; Staimer, N; Tjoa, T; Gillen, D; Kleinman, MT; Sioutas, C; Cooper, D. 2008. "Personal and ambient air pollution exposures and lung function decrements in children with asthma." *Environ. Health Perspect.* 116:550-558.

European Heart Journal. 1996. "Heart rate variability. Standards of measurement, physiological interpretation, and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology." *Eur. Heart J.* 17(3):354-381.

Fewell, Z; Smith, GD; Sterne, JA. 2007. "The impact of residual and unmeasured confounding in epidemiologic studies: A simulation study." *Am. J. Epidemiol.* 166:646-655.

Folinsbee, LJ. 1992. "Does nitrogen dioxide exposure increase airways responsiveness?" *Toxicol. Ind. Health* 8:273-283.

Freedman, S; Lane, R; Gillett, MK; Guz, A. 1988. "Abolition of methacholine induced bronchoconstriction by the hyperventilation of exercise or volition." *Thorax* 43(8):631-636.

Gehring, U; Wijga, AH; Brauer, M; Fischer, P; de Jongste, JC; Kerkhof, M; Oldenwening, M; Smit, HA; Brunekreef, B. 2010. "Traffic-related air pollution and asthma onset in children: A prospective cohort study with individual exposure measurement." *Am. J. Respir. Crit. Care Med.* 181(6):596-603.

Gent, JF; Koutrakis, P; Belanger, K; Triche, E; Holford, TR; Bracken, MB; Leaderer, BP. 2009. "Symptoms and medication use in children with asthma and traffic-related sources of fine particle pollution." *Environ. Health Perspect.* 117(7):1168-1174.

Goldberg, MS; Giannetti, N; Burnett, RT; Mayo, NE; Valois, MF; Brophy, JM. 2008. "A panel study in congestive heart failure to estimate the short-term effects from personal factors and environmental conditions on oxygen saturation and pulse rate." *Occup. Environ. Med.* 65(10):659-666.

Goodman, JE; Chandalia, JK; Thakali, S; Seeley, M. 2009. "Meta-analysis of nitrogen dioxide exposure and airway hyper-responsiveness in asthmatics." *Crit. Rev. Toxicol.* 39(9):719-742.

Greenwald, R; Sarnat, SE; Raysoni, AU; Li, WW; Johnson, BA; Stock, TH; Holguin, F; Sosa, T; Sarnat, JA. 2013. "Associations between source-indicative pollution metrics and increases in pulmonary inflammation and reduced lung function in a panel of asthmatic children." *Air Qual. Atmos. Health* 6(2):487-799.

Gruzieva, O; Bergstrom, A; Hulchiy, O; Kull, I; Lind, T; Melen, E; Moskalenko, V; Pershagen, G; Bellander, T. 2013. "Exposure to air pollution from traffic and childhood asthma until 12 years of age." *Epidemiology* 24(1):54-61.

Guo, Y; Jia, Y; Pan, X; Liu, L; Wichmann, HE. 2009. "The association between fine particulate air pollution and hospital emergency room visits for cardiovascular diseases in Beijing, China." *Sci. Total Environ.* 407(17):4826-4830.

Hahn, AG; Nogrady, SG; Tumilty, DM; Lawrence, SR; Morton, AR. 1984. "Histamine reactivity during the refractory period after exercise induced asthma." *Thorax* 39(12):919-923.

Hansel, NN; Breysse, PN; McCormack, MC; Matsui, EC; Curtin-Brosnan, J; Williams, DL; Moore, JL; Cuhran, JL; Diette, GB. 2008. "A longitudinal study of indoor nitrogen dioxide levels and respiratory symptoms in inner-city children with asthma." *Environ. Health Perspect.* 116(10):1428-1432.

Health Effects Institute. 1989. "Responses of Susceptible Subpopulations to Nitrogen Dioxide." HEI Research Report No. 23, 101p.

Health Effects Institute. 2010. "Public Health and Air Pollution in Asia (PAPA): A Combined Analysis of Four Studies of Air Pollution and Mortality." HEI Research Report 154 Part 5. In *Public Health and Air Pollution in Asia (PAPA): Coordinated Studies of Short-Term Exposure to Air Pollution and Daily Mortality in Four Cities*. Health Effects Institute, Boston, MA, p377-431.

Health Effects Institute (HEI). 2012. "Allergic Inflammation in the Human Lower Respiratory Tract Affected by Exposure to Diesel Exhaust." HEI Research Report 165, 82p., February.

Hill, AB. 1965. "The environment and disease: Association or causation?" *Proc. R. Soc. Med.* 58:295-300.

Hsieh, YL; Yang, YH; Wu, TN; Yang, CY. 2010. "Air pollution and hospital admissions for myocardial infarction in a subtropical city: Taipei, Taiwan." *J. Toxicol. Environ. Health A* 73(11):757-765.

Huang, W; Zhu, T; Pan, X; Hu, M; Lu, SE; Lin, Y; Wang, T; Zhang, Y; Tang, X. 2012a. "Air pollution and autonomic and vascular dysfunction in patients with cardiovascular disease: Interactions of systemic inflammation, overweight, and gender." *Am. J. Epidemiol.* 176(2):117-126.

Huang, YC; Rappold, AG; Graff, DW; Ghio, AJ; Devlin, RB. 2012b. "Synergistic effects of exposure to concentrated ambient fine pollution particles and nitrogen dioxide in humans." *Inhal. Toxicol.* 24(12):790-797.

Inman, MD; Watson, RM; Killian, KJ; O'Byrne, PM. 1990. "Methacholine airway responsiveness decreases during exercise in asthmatic subjects." *Am. Rev. Respir. Dis.* 141:1414-1417.

Institute of Medicine (IOM). 2008. *Improving the Presumptive Disability Decision-Making Process for Veterans*. Committee on Evaluation of the Presumptive Disability Decision-Making Process for Veterans, Board on Military and Veterans Health. *National Academies Press*, Washington, DC, 781p.

Ito, K; Thurston, GD; Silverman, RA. 2007. "Characterization of $PM_{2.5}$, gaseous pollutants, and meteorological interactions in the context of time-series health effects models." *J. Expo. Sci. Environ. Epidemiol.* 17(Suppl. 2):S45-S60.

Ito, K; Mathes, R; Ross, Z; Nadas, A; Thurston, G; Matte, T. 2011. "Fine particulate matter constituents associated with cardiovascular hospitalizations and mortality in New York City." *Environ. Health Perspect.* 119(4):467-473.

Jalaludin, B; Khalaj, B; Sheppeard, V; Morgan, G. 2008. "Air pollution and ED visits for asthma in Australian children: A case-crossover analysis." *Int. Arch. Occup. Environ. Health* 81(8):967-974.

Jenkins, HS; Devalia, JL; Mister, RL; Bevan, AM; Rusznak, C; Davies, RJ. 1999. "The effect of exposure to ozone and nitrogen dioxide on the airway response of atopic asthmatics to inhaled allergen: Dose- and time-dependent effects." *Am. J. Respir. Crit. Care Med.* 160:33-39.

Jerrett, M; Shankardass, K; Berhane, K; Gauderman, WJ; Künzli, N; Avol, E; Gilliland, F; Lurmann, F; Molitor, JN; Molitor, JT; Thomas, DC; Peters, J; McConnell, R. 2008. "Traffic-related air pollution and asthma onset in children: A prospective cohort study with individual exposure measurement." *Environ. Health Perspect.* 116(10):1433-1438.

Jorres, R; Magnussen, H. 1990. "Airways response of asthmatics after a 30 minute exposure, at resting ventilation, to 0.25 ppm NO_2 or 0.5 ppm SO_2." *Eur. Respir. J.* 3(2):132-137.

Jorres, R; Magnussen, H. 1991. "Effect of 0.25 ppm nitrogen dioxide on the airway response to methacholine in asymptomatic asthmatic patients." *Lung* 169:77-85.

Kleinman, MT; Bailey, RM; Linn, WS; Anderson, KR; Whynot, JD; Shamoo, DA; Hackney, JD. 1983. "Effects of 0.2 ppm nitrogen dioxide on pulmonary function and response to bronchoprovocation in asthmatics." *J. Toxicol. Environ. Health* 12(4-6):815-826.

Ko, FW; Tam, W; Wong, TW; Lai, CK; Wong, GW; Leung, TF; Ng, SS; Hui, DS. 2007. "Effects of air pollution on asthma hospitalization rates in different age groups in Hong Kong." *Clin. Exp. Allergy* 37(9):1312-1319.

Larrieu, S; Jusot, JF; Blanchard, M; Prouvost, H; Declercq, C; Fabre, P; Pascal, L; Tertre, AL; Wagner, V; Riviere, S; Chardon, B; Borrelli, D; Cassadou, S; Eilstein, D; Lefranc, A. 2007. "Short term effects of air pollution on hospitalizations for cardiovascular diseases in eight French cities: The PSAS program." *Sci. Total Environ.* 387(1-3):105-112.

Lee, YL; Hwang, BF; Chen, YA; Chen, JM; Wu, YF. 2012. "Pulmonary function and incident bronchitis and asthma in children: A community-based prospective cohort study." *PLoS ONE* 7(3):e32477.

Li, S; Batterman, S; Wasilevich, E; Wahl, R; Wirth, J; Su, FC; Mukherjee, B. 2011. "Association of daily asthma emergency department visits and hospital admissions with ambient air pollutants among the pediatric Medicaid population in Detroit: Time-series and time-stratified case-crossover analyses with threshold effects." *Environ. Res.* 111(8):1137-1147.

Linares, C; Diaz, J. 2010. "Short-term effect of concentrations of fine particulate matter on hospital admissions due to cardiovascular and respiratory causes among the over-75 age group in Madrid, Spain." *Public Health* 124(1):28-36.

Lu, KD; Breysse, PN; Diette, GB; Curtin-Brosnan, J; Aloe, C; Williams, DL; Peng, RD; McCormack, MC; Matsui, EC. 2013. "Being overweight increases susceptibility to indoor pollutants among urban children with asthma." *J. Allergy Clin. Immunol.* 131(4):1017-1023.

Magnussen, H; Reuss, G; Jorres, R. 1986. "Airway response to methacholine during exercise induced refractoriness in asthma." *Thorax* 41(9):667-670.

Mann, JK; Tager, IB; Lurmann, F; Segal, M; Quesenberry, CP; Lugg, MM; Shan, J; Van Den Eeden, SK. 2002. "Air pollution and hospital admissions for ischemic heart disease in persons with congestive heart failure or arrhythmia." *Environ. Health Perspect.* 110:1247-1252.

Marra, C; Bottaro, M; Oliveira, RJ; Novaes, JS. 2005. "Effect of moderate and high intensity aerobic exercise on the body composition of overweight men." *J. Exerc. Physiol. Online* 8(2):39-45.

Martin, L. 1999. "Exercise physiology." In *Pulmonary Physiology in Clinical Practice*. Lakeside Press.

McConnell, R; Islam, T; Shankardass, K; Jerrett, M; Lurmann, F; Gilliland, F; Gauderman, J; Avol, E; Kunzli, N; Yao, L; Peters, J; Berhane, K. 2010. "Childhood incident asthma and traffic-related air pollution at home and school." *Environ. Health Perspect.* 118(7):1021-1026.

Moolgavkar, SH; McClellan, RO; Dewanji, A; Turim, J; Luebeck, EG; Edwards, M. 2013. "Time-series analyses of air pollution and mortality in the United States: A subsampling approach." *Environ. Health Perspect.* 121(1):73-78.

Nishimura, KK; Galanter, JM; Roth, LA; Oh, SS; Thakur, N; Nguyen, EA; Thyne, S; Farber, HJ; Serebrisky, D; Kumar, R; Brigino-Buenaventura, E; Davis, A; Lenoir, MA; Meade, K; Rodriguez-Cintron, W; Avila, PC; Borrell, LN; Bibbins-Domingo, K; Rodriguez-Santana, JR; Sen, S; Lurmann, F; Balmes, JR; Burchard, EG. 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(3):309-318.

Nuvolone, D; Balzi, D; Chini, M; Scala, D; Giovannini, F; Barchielli, A. 2011. "Short-term association between ambient air pollution and risk of hospitalization for acute myocardial infarction: Results of the cardiovascular risk and air pollution in Tuscany (RISCAT) study." *Am. J. Epidemiol.* 174(1):63-71.

O'Connor, GT; Neas, L; Vaughn, B; Kattan, M; Mitchell, H; Crain, EF; Evans, R III; Gruchalla, R; Morgan, W; Stout, J; Adams, GK; Lippman, M. 2008. "Acute respiratory health effects of air pollution on children with asthma in US inner cities." *J. Allergy Clin. Immunol.* 121(5):1133-1139.

Orehek, J; Massari, JP; Gayrard, P; Grimaud, C; Charpin, J. 1976. "Effect of short-term, low-level nitrogen dioxide exposure on bronchial sensitivity of asthmatic patients." *J. Clin. Invest.* 57:301-307.

Peel, JL; Metzger, KB; Klein, M; Flanders, WD; Mulholland, JA; Tolbert, PE. 2007. "Ambient air pollution and cardiovascular emergency department visits in potentially sensitive groups." *Am. J. Epidemiol.* 165(6):625-633.

Pereira Filho, MA; Pereira, LA; Arbex, FF; Arbex, M; Conceicao, GM; Santos, UP; Lopes, AC; Saldiva, PH; Braga, AL; Cendon, S. 2008. "Effect of air pollution on diabetes and cardiovascular diseases in São Paulo, Brazil." *Braz. J. Med. Biol. Res.* 41(6):526-532.

Rhomberg, LR; Chandalia, JK; Long, CM; Goodman, JE. 2011. "Measurement error in environmental epidemiology and the shape of exposure-response curves." *Crit. Rev. Toxicol.* 41(8):651-671.

Rich, DQ; Kipen, HM; Zhang, J; Kamat, L; Wilson, AC; Kostis, JB. 2010. "Triggering of transmural infarctions, but not nontransmural infarctions, by ambient fine particles." *Environ. Health Perspect.* 118(9):1229-1234.

Roger, LJ; Horstman, DH; McDonnell, W; Kehrl, H; Ives, PJ; Seal, E; Chapman, R; Massaro, E. 1990. "Pulmonary function, airway responsiveness, and respiratory symptoms in asthmatics following exercise in NO(2)." *Toxicol. Ind. Health* 6(1):155-171.

Rubinstein, I; Bigby, BG; Reiss, TF; Boushey, HA. 1990. "Short-term exposure to 0.3 ppm nitrogen dioxide does not potentiate airway responsiveness to sulfur dioxide in asthmatic subjects." *Am. Rev. Respir. Dis.* 141:381-385.

Samoli, E; Nastos, PT; Paliatsos, AG; Katsouyanni, K; Priftis, KN. 2011. "Acute effects of air pollution on pediatric asthma exacerbation: Evidence of association and effect modification." *Environ. Res.* 111(3):418-424.

Sandercock, GR; Bromley, PD; Brodie, DA. 2005. "The reliability of short-term measurements of heart rate variability." *Int. J. Cardiol.* 103(3):238-247.

Sarnat, SE; Raysoni, AU; Li, WW; Holguin, F; Johnson, BA; Flores Luevano, S; Garcia, JH; Sarnat, JA. 2012. "Air pollution and acute respiratory response in a panel of asthmatic children along the U.S.-Mexico border." *Environ. Health Perspect.* 120(3):437-444.

Scaife, A; Barclay, J; Hillis, GS; Srinivasan, J; Macdonald, DW; Ross, JA; Ayres, JG. 2012. "Lack of effect of nitrogen dioxide exposure on heart rate variability in patients with stable coronary heart disease and impaired left ventricular systolic function." *Occup. Environ. Med.* 69(8):587-591.

Shima, M; Nitta, Y; Ando, M; Adachi, M. 2002. "Effects of air pollution on the prevalence and incidence of asthma in children." *Arch. Environ. Health* 57(6):529-535.

Son, JY; Lee, JT; Park, YH; Bell, ML. 2013. "Short-term effects of air pollution on hospital admissions in Korea." *Epidemiology* 24(4):545-554.

Stieb, DM; Szyszkowicz, M; Rowe, BH; Leech, JA. 2009. "Air pollution and emergency department visits for cardiac and respiratory conditions: A multi-city time-series analysis." *Environ. Health* 8:25.

Strand, V; Salomonsson, P; Lundahl, J; Bylin, G. 1996. "Immediate and delayed effects of nitrogen dioxide exposure at an ambient level on bronchial responsiveness to histamine in subjects with asthma." *Eur. Respir. J.* 9:733-740.

Strand, V; Rak, S; Svartengren, M; Bylin, G. 1997. "Nitrogen dioxide exposure enhances asthmatic reaction to inhaled allergen in subjects with asthma." *Am. J. Respir. Crit. Care Med.* 155(3):881-887.

Strand, V; Svartengren, M; Rak, S; Barck, C; Bylin, G. 1998. "Repeated exposure to an ambient level of NO_2 enhances asthmatic response to a nonsymptomatic allergen dose." *Eur. Respir. J.* 12:6-12.

Strickland, MJ; Darrow, LA; Klein, M; Flanders, WD; Sarnat, JA; Waller, LA; Sarnat, SE; Mulholland, JA; Tolbert, PE. 2010. "Short-term associations between ambient air pollutants and pediatric asthma emergency department visits." *Am. J. Respir. Crit. Care Med.* 182(3):307-316.

Szyszkowicz, M. 2007. "Air pollution and emergency department visits for ischemic heart disease in Montreal, Canada." *Int. J. Occup. Med. Environ. Health* 20(2):167-173.

Szyszkowicz, M. 2009. "Air pollution and ED visits for chest pain." *Am. J. Emerg. Med.* 27(2):165-168.

Thach, RQ; Wong, CM; Chan, KP; Chau, YK; Thomas, GN; Ou, CQ; Yang, L; Peiris, JSM; Lam, TH; Hedley, AJ. 2010. "Air pollutants and health outcomes: Assessment of confounding by influenza." *Atmos. Environ.* 44(11):1437-1442.

Timonen, KL; Vanninen, E; De Hartog, J; Ibald-Mulli, A; Brunekreef, B; Gold, DR; Heinrich, J; Hoek, G; Lanki, T; Peters, A; Tarkiainen, T; Tiittanen, P; Kreyling, W; Pekkanen, J. 2006. "Effects of ultrafine and fine particulate and gaseous air pollution on cardiac autonomic control in subjects with coronary artery disease: The ULTRA study." *J. Expo. Anal. Environ. Epidemiol.* 16:332-341.

Tolbert, PE; Klein, M; Peel, JL; Sarnat, SE; Sarnat, JA. 2007. "Multipollutant modeling issues in a study of ambient air quality and emergency department visits in Atlanta." *J. Expo. Sci. Environ. Epidemiol.* 17:S29-S35.

Tunnicliffe, WS; Burge, PS; Ayres, JG. 1994. "Effect of domestic concentrations of nitrogen dioxide on airway responses to inhaled allergen in asthmatic patients." *Lancet* 344(8939-8940):1733-1736.

Turin, TC; Kita, Y; Rumana, N; Nakamura, Y; Ueda, K; Takashima, N; Sugihara, H; Morita, Y; Ichikawa, M; Hirose, K; Nitta, H; Okayama, A; Miura, K; Ueshima, H. 2012. "Short-term exposure to air pollution and incidence of stroke and acute myocardial Infarction in a Japanese population." *Neuroepidemiology* 38(2):84-92.

US EPA. 2008. "Integrated Science Assessment for Oxides of Nitrogen." National Center for Environmental Assessment (NCEA) - RTP Division, EPA/600/R-08/071, 260p., July.

US EPA. 2013. "Integrated Science Assessment for Oxides of Nitrogen–Health Criteria (First External Review Draft)." National Center for Environmental Assessment (NCEA), EPA/600/R-13/202. 878p., November.

Vencloviene, J; Grazuleviciene, R; Babarskiene, R; Dedele, A; Grazulevicius, T. 2011. "Short-term nitrogen dioxide exposure and geomagnetic activity interaction: Contribution to emergency hospitalization for acute coronary syndrome." *Int. J. Environ. Health Res.* 21(3):149-160.

Villeneuve, PJ; Chen, L; Rowe, BH; Coates, F. 2007. "Outdoor air pollution and emergency department visits for asthma among children and adults: A case-crossover study in northern Alberta, Canada." *Environ. Health* 6:40.

von Klot, S; Peters, A; Aalto, P; Bellander, T; Berglind, N; D'Ippoliti, D; Elosua, R; Hormann, A; Kulmala, M; Lanki, T; Lowel, H; Pekkanen, J; Picciotto, S; Sunyer, J; Forastiere, F. 2005. "Ambient air pollution is associated with increased risk of hospital cardiac readmissions of myocardial infarction survivors in five European cities." *Circulation* 112(20):3073-3079.

Wichmann, J; Ketzel, M; Ellermann, T; Loft, S. 2012. "Apparent temperature and acute myocardial infarction hospital admissions in Copenhagen, Denmark: A case-crossover study." *Environ. Health* 11:19.

Williams, R; Brook, R; Bard, R; Conner, T; Shin, H; Burnett, R. 2012. "Impact of personal and ambient-level exposures to nitrogen dioxide and particulate matter on cardiovascular function." *Int. J. Environ. Health Res.* 22(1):71-97.

Witten, A; Solomon, C; Abbritti, E; Arjomandi, M; Zhai, W; Kleinman, M; Balmes, J. 2005. "Effects of nitrogen dioxide on allergic airway responses in subjects with asthma." *J. Occup. Environ. Med.* 47(12):1250-1259.

Wong, CM; Vichit-Vadakan, N; Kan, H; Qian, Z. 2008. "Public Health and Air Pollution in Asia (PAPA): A multicity study of short-term effects of air pollution on mortality." *Environ. Health Perspect.* 116(9):1195-1202.

Zanobetti, A; Gold, DR; Stone, PH; Suh, HH; Schwartz, J; Coull, BA; Speizer, FE. 2010. "Reduction in heart rate variability with traffic and air pollution in patients with coronary artery disease." *Environ. Health Perspect.* 118(3):324-330.

Tables

Table 1.1 Causal Determinations for Short- and Long-term NO_2 Exposure in the 2008 and 2013 ISA

Health Effect Category	Causal Determination		Why 2013 Determination Is Not Supported
	2008 ISA for Oxides of Nitrogen	2013 Draft ISA for Oxides of Nitrogen	
Short-term			
Respiratory Effects	Sufficient to determine a likely causal relationship	Causal	Inadequate quality of epidemiology evidence Lack of coherence with controlled exposure or toxicology studies Uncertainty regarding whether NO_2 is a proxy for traffic-related pollution
Cardiovascular Effects	Inadequate to infer the presence or absence of a causal relationship	Likely causal	Inadequate quality of epidemiology studies Lack of coherence across studies Lack of confirmed MoAs
Total Mortality	Suggestive but not sufficient to infer a causal relationship	Likely causal	Inadequate quality of epidemiology studies Lack of confirmed MoAs
Long-term			
Respiratory Effects	Suggestive but not sufficient to infer a causal relationship	Likely causal	Inadequate quality of epidemiology studies Uncertainty regarding whether NO_2 is a proxy for traffic-related pollution
Cardiovascular Effects	Inadequate to infer the presence or absence of a causal relationship	Suggestive	Inadequate quality of epidemiology studies Lack of confirmed MoAs
Reproductive and Developmental Effects	Inadequate to infer the presence or absence of a causal relationship	Fertility, Reproduction, and Pregnancy: Suggestive	Inconsistent findings among studies with the same endpoints Lack of coherence across studies with different endpoints
		Birth Outcomes: Suggestive	Inconsistent findings regarding temporality of NO_2 exposure
		Postnatal Development: Suggestive	Inadequate quality of epidemiology studies Inconsistent findings among studies with the same endpoints Lack of confirmed MoAs
Total Mortality	Inadequate to infer the presence or absence of a causal relationship	Suggestive	Inadequate quality of epidemiology studies
Cancer	Inadequate to infer the presence or absence of a causal relationship	Suggestive	Inadequate quality of epidemiology studies

Note:
ISA = Integrated Science Assessment; MoA = mode of action; NO_2 = nitrogen dioxide.

Table 2.1 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 2.2 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 copollutant 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 (2013, Table II).

Table 2.3 EPA's Classification of Evidence for Potential "At-risk Factors"[1]

Classification	Health Effects
Adequate evidence	There is substantial, consistent evidence within a discipline to conclude that a factor results in a population or lifestage being at increased or decreased risk of air pollutant-related health effect(s) relative to some reference population or lifestage. Where applicable this includes coherence across disciplines. Evidence includes multiple high-quality studies.
Suggestive evidence	The collective evidence suggests that a factor results in a population or lifestage being at increased or decreased risk of an air pollutant-related health effect relative to some reference population or lifestage, but the evidence is limited due to some inconsistency within a discipline or, where applicable, a lack of coherence across disciplines.
Inadequate evidence	The collective evidence is inadequate to determine if a factor results in a population or lifestage being at increased or decreased risk of an air pollutant-related health effect relative to some reference population or lifestage. The available studies are of insufficient quantity, quality, consistency, and/or statistical power to permit a conclusion to be drawn.
Evidence of no effect	There is substantial, consistent evidence within a discipline to conclude that a factor does not result in a population or lifestage being at increased or decreased risk of air pollutant-related health effect(s) relative to some reference population or lifestage. Where applicable this includes coherence across disciplines. Evidence includes multiple high-quality studies.

Note:
(1) An "at-risk factor" is best described as an effect modifier, which is a technical term defined in epidemiology as a variable that differentially modifies the observed effect of a risk factor on disease status.
Source: US EPA (2013, Table III).