Critical Reviews in Toxicology

http://informahealthcare.com/txc
ISSN: 1040-8444 (print), 1547-6898 (electronic)

Crit Rev Toxicol, 2014; 44(9): 791–822
 DOI: 10.3109/10408444.2014.937855

informa
healthcare

REVIEW ARTICLE

Weight-of-evidence evaluation of long-term ozone exposure and cardiovascular effects

Robyn L. Prueitt[1], Heather N. Lynch[1], Ke Zu[1], Sonja N. Sax[1], Ferdinand J. Venditti[2], and Julie E. Goodman[1]

[1]*Gradient, Cambridge, MA, USA and* [2]*Albany Medical College, Albany, NY, USA*

Abstract

We conducted a weight-of-evidence (WoE) analysis to assess whether the current body of research supports a causal relationship between long-term ozone exposure (defined by EPA as at least 30 days in duration) at ambient levels and cardiovascular (CV) effects. We used a novel WoE framework based on the United States Environmental Protection Agency's National Ambient Air Quality Standards causal framework for this analysis. Specifically, we critically evaluated and integrated the relevant epidemiology and experimental animal data and classified a causal determination based on categories proposed by the Institute of Medicine's 2008 report, *Improving the Presumptive Disability Decision-making Process for Veterans*. We found that the risks of CV effects are largely null across human and experimental animal studies. The few positive associations reported in studies of CV morbidity and mortality are very small in magnitude, mainly reported in single-pollutant models, and likely attributable to bias, chance, or confounding. The few positive effects in experimental animal studies were observed mainly in *ex vivo* studies at high exposures, and even the *in vivo* findings are not likely relevant to humans. The available data also do not support a biologically plausible mechanism for the effects of ozone on the CV system. Overall, the current WoE provides no convincing case for a causal relationship between long-term exposure to ambient ozone and adverse effects on the CV system in humans, but the limitations of the available studies preclude definitive conclusions regarding a lack of causation; thus, we categorize the strength of evidence for a causal relationship between long-term exposure to ozone and CV effects as "below equipoise."

Keywords

air pollution, cardiovascular system, causal framework, epidemiology, mode of action, risk assessment, systematic review, weight of evidence

History

Received 3 February 2014
Revised 22 May 2014
Accepted 19 June 2014
Published online 29 August 2014

Table of Contents

Address for correspondence: Julie E. Goodman, Gradient, 20 University Road, Suite 5, Cambridge, MA 02138, USA. E-mail: jgoodman@gradientcorp.com

Introduction

The United States Environmental Protection Agency (EPA) periodically reviews the literature regarding the effects of ozone and five other "criteria" air pollutants on health, as part of the process for setting health-based National Ambient Air Quality Standards (NAAQS). Ozone is not directly emitted into the air from specific sources, but is a secondary air pollutant that is formed by the photochemical reactions between precursor gases, primarily nitrogen oxides (NO_x) and volatile organic compounds (VOCs), in the presence of ultraviolet (UV) rays from the sun. Ozone formation and degradation are complex and depend on many factors, including the relative concentrations of precursor gases and meteorological factors (e.g. sunlight intensity and atmospheric mixing) (Figure 1). The chemical reactivity of VOCs can be quite variable, and the relative concentration of specific VOCs and NO_x is important for ozone formation because, under some conditions, formation of ozone is VOC-limited, whereas under other conditions, it is NO_x-limited (NRC 1991). Because NO_x is involved in both the formation and degradation of ozone, reducing NO_x may increase ozone concentrations under some conditions, and this must be considered when developing ozone control strategies. The critical role of UV intensity as a driver of ozone formation results in a distinct diurnal pattern for ambient ozone concentrations. Typically, ozone concentrations begin increasing as the sun rises, reach a peak near mid-day, and decrease markedly after sunset (see Figure AX3-42 in US EPA 2006). The net UV flux can be increased by the reflection of UV rays on snow cover, giving rise to unexpected, high ozone episodes in the winter (Carter and Seinfeld 2012). As a result of the variable factors influencing ozone formation and degradation, ambient ozone concentrations vary widely both spatially and temporally (USEPA 2013a).

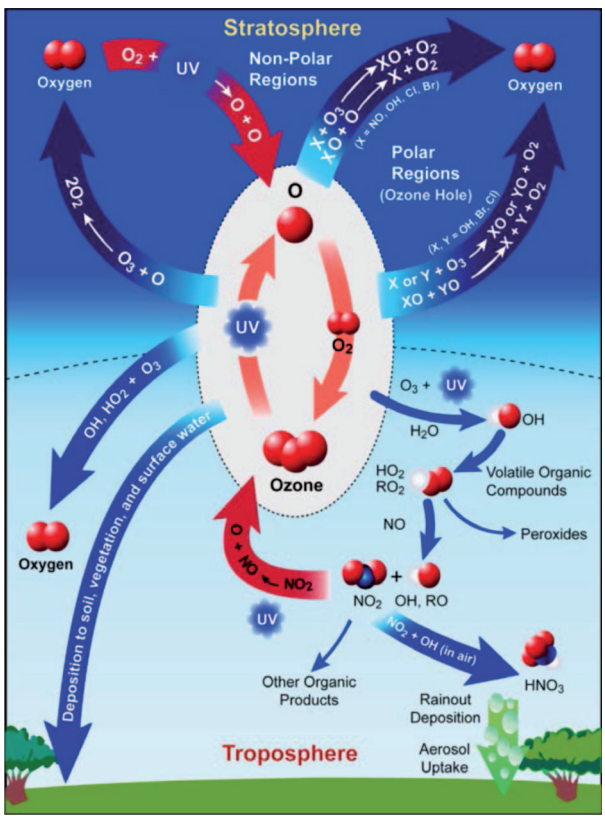

Figure 1. Overview of the photochemical processes influencing ozone formation. *Source:* US EPA, 2013a.

Each NAAQS has four elements: (1) an indicator (for photochemical oxidants, it is ozone), (2) an averaging time, (3) a numerical level or concentration, and (4) a statistical form (US EPA 2013a). Any discussion of a NAAQS for ozone is incomplete without stating the averaging time and statistical form of the standard. Until 1997, the primary (health-based) ozone standard was a daily 1-h maximum concentration of 120 parts per billion (ppb) that was not to be exceeded more than once a year. This standard was problematic because attainment of the 1-h standard could vary from year to year in a given area, depending primarily on meteorological conditions (NRC 1991). Therefore, in 1997, EPA determined that a longer averaging time for ozone (i.e. 8 h) would provide greater stability for meeting the standard. With the change in the averaging time from 1- to 8-h, the concentration of the standard was reduced from 120 ppb to 80 ppb (equivalent to 84 ppb using standard rounding conventions). In 2008, the ozone NAAQS was revised so that the annual fourth-highest daily maximum 8-h concentration of ozone, averaged over three years, should not exceed 75 parts per billion (ppb).

In the latest ozone review, EPA developed an Integrated Science Assessment (ISA) in which it evaluated the weight of evidence (WoE) for causal relationships between exposure to ozone and its effects on human health and welfare, based on the available scientific literature (US EPA 2013). The WoE framework EPA used for its evaluation is referred to as the "NAAQS causal framework."

Among the health effects evaluated in the ozone ISA, the EPA used the NAAQS causal framework to review the potential cardiovascular (CV) effects of long-term exposure (defined by the EPA as at least 30 days in duration) to ozone and concluded that the limited evidence is suggestive of a causal relationship (US EPA 2013). Goodman et al. (2013) described an updated causal framework (referred to herein as the "Goodman WoE framework") based on WoE best practices gleaned from a survey of more than 50 WoE frameworks, including the NAAQS causal framework (Rhomberg et al. 2013). In this analysis, we use the principles of the Goodman WoE framework (Goodman et al. 2013) to evaluate studies of long-term exposure to ozone and CV morbidity and mortality, to assess whether ozone may be a causal factor for adverse CV effects. We contrast our analysis to that conducted by the EPA in the ISA, and consider whether and how differences between the NAAQS causal framework and the Goodman WoE framework led to different conclusions. While this analysis focuses on effects

of long-term ozone exposure, we used the same methods to evaluate studies of short-term exposure to ozone and adverse CV effects in a companion analysis (Goodman et al. 2014).

Methods

We used the principles of the Goodman WoE framework (Goodman et al. 2013, Table 1), which consists of four phases, to evaluate the potential effects of long-term exposure to ozone on the CV system. We describe the methods of this analysis in our companion paper (Goodman et al. 2014) and summarize them briefly below.

In Phase 1, we defined the causal question and inclusion/exclusion criteria for selecting the studies to evaluate. We assumed that study identification in the ISA (US EPA 2013) was likely to be fairly comprehensive, as those processes involved literature searches by the EPA and input from expert scientists, advisory committees, and the public. Because of this, we initially used the ISA to identify studies for evaluation. We also conducted an independent literature search to identify additional relevant studies published between January 1, 2006 and November 4, 2013, as described in detail below in the Phase 1 section. We included the epidemiology and experimental animal studies of CV-related morbidity, mortality, and biomarkers in our analysis.

In Phase 2, we extracted the study characteristics and data into tables and assessed individual study quality and relevance. In addition to evaluating each study individually, we used a crude scoring method (described in detail in the Phase 2 section below) to compile information and roughly categorize each study based on quality. This method made it possible to assess study quality in a consistent manner across studies (by making sure that the same factors were considered for each study) and to get a general sense of study quality across metrics. The resulting scores were used to roughly divide studies into two categories, but were not used to rank studies. We want to emphasize that these scores were not thought of as checklists, but rather a way of getting an overall sense of quality, in addition to the detailed, qualitative assessment of quality that we conducted for each study. We considered several factors for each study type. For the epidemiology studies, we considered the study design, potential for selection bias and exposure and outcome misclassification, statistical approach, control for confounding, and sensitivity analyses. For the experimental animal studies, we considered the assignment to and size of experimental groups, use of appropriate controls, animal husbandry and housing conditions, exposure methods, outcome assessment, attrition bias, and statistical approach.

In Phase 3, we evaluated and integrated the data within and across realms of evidence (i.e. epidemiological, experimental animal). Within each realm, we assessed individual study results, as well as the consistency of results across studies for each endpoint, considering the strength of association, internal consistency, temporality, biological plausibility, exposure-response, and random error (chance) when feasible. For the evaluation and integration of data across realms of evidence, we considered the strength of association, consistency of association, coherence, biological plausibility, biological gradient, experimental evidence, temporality, specificity, confounding, bias, mechanistic evidence, and the adversity of effects. We compared alternative accounts of the evidence and formulated WoE conclusions, noting data gaps. In forming our conclusions, we assigned more weight (i.e. relied on to a greater extent) to studies that we considered to be of higher quality in Phase 2.

In Phase 4, we used the WoE conclusions from Phase 3 to categorize the potential causal relationship between long-term ozone exposure and CV effects. We relied on the categories of causal determination proposed in the Institute of Medicine (IOM) report *Improving the Presumptive Disability Decision-making Process for Veterans* (hereafter, "the IOM framework") (IOM 2008). The IOM framework is the basis for the NAAQS causal framework, and use of its four-level categorization scheme is consistent with WoE best practices (Goodman et al. 2013). We contrasted our conclusions with those of the EPA in the ISA, and we assessed how the differences between the NAAQS causal framework and the Goodman WoE framework affected the conclusions.

Table 1. Weight-of-evidence best practices from the Goodman WoE framework.

Phase	Steps
Phase 1	
Define the causal question and develop criteria for study selection	Frame the purpose of the evaluation and the causal questions to be evaluated, and define the criteria for selecting the studies relevant to the evaluation to ensure transparency
Phase 2	
Develop and apply criteria for review of individual studies	Conduct and present a systematic and consistent review of available studies relevant to the causal question. Evaluate the rigor and quality of individual study results using pre-defined criteria applied uniformly across studies
Phase 3	
Integrate and evaluate evidence	Make sound and defensible scientific judgments about the existence and nature of causative processes for the health outcome under consideration. This is one of the more challenging phases for any WoE framework; no matter how one lays out procedures and methods for synthesizing across studies, in the end, the question is about how studies in one setting (e.g. animal or *in vitro* assays) should affect our assessment of potential causality or risks in another (e.g. the general human population exposed environmentally)
Phase 4	
Draw conclusions based on inferences	Apply the results of the WoE evaluation from Phase 3 to make conclusions that can be used to inform regulatory decision-making. Although this phase is not risk management itself, it can be influenced by risk management considerations. In a regulatory setting, decisions about WoE categories ("known" causative agent, or "likely" causative agent, *etc*.) or findings about the science (sufficient evidence for a mode of action or to replace a default assumption for developing a toxicity value) are influenced in this stage by policy questions and regulatory consequences for those decisions and, ultimately, by policies and judgments about the sufficiency of evidence to support those decisions

Phase 1 – Explanation of the causal question and study selection

In this analysis, we addressed the question of whether long-term exposure to ozone is a causal factor for adverse, CV-related health effects in humans. To do this, we assessed the studies of arrhythmia, blood pressure, carotid artery intima-media thickness (CIMT), ischemic heart disease (IHD), cardiovascular disease (CVD), heart failure, stroke, and myocardial infarction (MI). We also assessed the experimental animal studies of heart rate, aortic staining for atherosclerosis, and ventricular function, as well as the epidemiology and animal studies of CV-related biomarkers. We identified studies of these endpoints from the literature reviewed in the most recent ISA (US EPA 2013) and we conducted an independent literature search to identify additional relevant studies.

We focused our literature search on papers published since the previous ozone review by the EPA in the 2006 Air Quality Criteria Document (AQCD) (US EPA 2006), as the EPA did in the ISA. We used PubMed and Scopus databases and included studies published between January 1, 2006, and November 4, 2013. The search terms for the PubMed database were ozone AND (cardiovascular diseases OR myocardial infarction OR heart attack OR blood pressure OR stroke OR hypertension OR hypotension OR prehypertension). The search terms for the Scopus database were (ABS (ozone) AND TITLE-ABS-KEY (cardiovascular OR heart OR cardiac OR "myocardial infarction" OR aneurysm OR hypertension OR "blood pressure" OR stroke OR cerebrovascular OR hypotension OR prehypertension)). We reviewed the reference lists of relevant reviews and key studies to identify additional studies.

We included epidemiology, experimental animal, and biomarker studies in our analysis. We included papers that evaluated ozone exposures of at least 30 days in duration, to be consistent with the EPA's definition of "long-term" in the ISA and to facilitate the comparison of our conclusions to those of the EPA. We excluded studies that evaluated exposures less than 30 days in duration. We also excluded experimental animal studies that reported results only for exposures to ozone combined with one or more other chemicals, because we were interested in effects attributable to ozone alone. While some of the studies identified by our literature search were cited in the ozone ISA, we identified several additional relevant studies that we included in our analysis (Table 2).

Phase 2 – Review of individual study quality and relevance

To evaluate the quality of the epidemiology and experimental animal studies selected for analysis in Phase 1, we categorized the study quality using a crude quantitative scoring method. Although changes in the levels of circulating biomarkers may be informative for elucidating potential mechanisms for the effects of ozone on the CV system, we did not consider them to be specific CV morbidity endpoints. Thus, we evaluated studies of CV-related biomarkers separately from those of CV morbidity endpoints.

Epidemiology studies

Several studies have evaluated the associations between long-term ozone exposure and various CV morbidity endpoints (e.g. arrhythmia, stroke, MI) and CV mortality. As we describe in our companion paper (Goodman et al. 2014), to aid in the critical review of individual studies, we used a scoring system that was based on several similar systems (Goodman et al. 2004, Rothman and Greenland 1998, Higgins et al. 2011, Nichols et al. 2013). This scoring method has not been validated externally, but it incorporates commonly recognized methodological issues in studies. We scored studies based on study design, potential for selection bias, potential for exposure and outcome misclassification, statistical approach, control for confounding, and sensitivity analyses. We did not score studies based on sample size because all studies had several hundred to tens of thousands of participants. We developed the scoring criteria before evaluating any studies, and two investigators independently used the criteria to develop scores for each study. If there was a discrepancy in scoring between the investigators, it was discussed, and if necessary, a third investigator was consulted to resolve any scoring issues.

Study design. We considered most prospective and retrospective cohort studies to be robust for making causal inferences and assigned them each a study design score of 1. We assigned cohort studies that estimated ozone exposures after information on health outcomes was collected (thus, temporality of effect was a concern) a score of 0. We assigned cross-sectional studies a score of 0 because within-subject variation can affect the validity of results.

Selection bias. We assigned a selection bias score of 1 to studies that had large, well-defined cohorts with little loss to follow-up ($\leq 20\%$) or that were based on a random sample of the underlying population. We assigned a score of 0 to studies where the inclusion of study participants might be related to ozone exposures or health outcomes (e.g. studies that restricted participants based on how close they lived to an air monitoring station, studies where response rates varied across communities, or studies that reported a loss to follow-up of

Table 2. Studies identified for inclusion in the analysis.

Realm of evidence	Number of studies identified from ISA	Number of additional studies identified from literature search	Total number of studies in analysis
Epidemiology – CV Morbidity	1	6	7
Epidemiology – CV Mortality	6	6	12
Epidemiology – Biomarkers	3	0	3
Experimental Animal – CV Morbidity	2	3	5
Experimental Animal – Biomarkers	2	3	5

greater than 20%). For studies that recruited participants from a single clinic or hospital, we considered that selection bias was likely; these cases may be related to socioeconomic factors that could also be correlated with ozone exposure or health outcomes. We assigned a selection bias score of − 1 to these studies.

Outcome assessment. We considered studies to be of higher quality if (where applicable) the outcomes were verified by a trained technician or a physician, multiple measurements were taken, and automatic devices were used. In these cases, we assigned studies an outcome assessment score of 1. When CV outcomes were obtained from hospital records or health registries and no independent review of medical records or death certificates was performed, we considered that outcome misclassification was likely and assigned a score of 0.

Exposure measurement. The potential for exposure measurement error is common in studies of ozone because most studies rely on measures of ambient concentrations that fluctuate over time. We assigned an exposure measurement score of 1 to studies that used personal monitors to measure individual exposures. For studies that relied on central-site monitors, we assigned a score of 0 to those that employed statistical modeling to account for spatial variation, or used measurements from local monitors (≤ 1 km to participants' residences), to estimate individual exposure. To some extent, the use of data from air monitoring stations in such close proximity to people's homes as surrogates for individual exposures may have reduced exposure assessment errors, but it could not completely account for individual mobility and indoor and workplace exposures. We assigned a score of − 1 to studies that estimated exposure using simple averages of measurements from several ambient monitors or measurements from a single ambient monitor for a large area.

Statistical modeling. We evaluated whether the appropriate statistical analyses were conducted based on the study design and research question. We considered Cox regression appropriate for long-term survival analysis because, under the assumption of proportional hazards, Cox regression does not specify baseline hazard and can account for time-varying exposure and covariates. We considered logistic regression appropriate for estimation of relative risks for binary outcomes if the prevalence/incidence of the outcomes was less than 10%. When the outcome was common (prevalence/incidence $\geq 10\%$), Poisson or log-binomial regression was considered appropriate for the estimation of relative risks. We considered generalized additive models to be appropriate statistical approaches to evaluate the associations between continuous predictors and outcomes, because generalized additive models incorporate nonlinear forms (smooth functions) of the predictors. Though considered appropriate in our study quality rating system, these statistical models are not without limitations. For example, a departure from the assumption of proportional hazard in Cox regression or over-fitting in generalized additive models may severely undermine the validity of the findings.

We also considered whether analyses were conducted using both single- and bi- or multi-pollutant models. Air pollutants tend to be highly correlated with each other and the outcome of interest, and this may be particularly true for CV effects (Barath et al. 2013). Analyses of several pollutants in the same statistical model may cause instability because of collinearity among the pollutants. However, when single-pollutant models indicate an association with a health effect that is in the same direction and of the same magnitude for both ozone and another pollutant that is correlated with ozone, unless bi- or multi-pollutant models are used, it is unclear whether the effect is due to ozone or whether ozone is a surrogate for effects from the other pollutant. Thus, we considered studies that used bi- or multi-pollutant models to be of higher quality than those that did not. For studies that used appropriate statistical models as defined above, we assigned a score of 1 to those that included bi- or multi-pollutant models and a score of 0 to those that only considered single-pollutant models. We assigned a score of − 1 to the studies that used inappropriate statistical models.

Control for confounders. We also assessed the degree to which studies considered key potential confounders. For cohort studies and individual-level cross-sectional studies, we assigned a score of − 1 for inadequate control for confounding if only age and sex were included. We assigned a score of 0 to studies that included age and sex, plus at least two of the following: body mass index (BMI), smoking, and socioeconomic status (SES). To obtain a score of 1, studies had to meet the criteria for a score of 0 and also had to have considered at least two of the following: diet, alcohol consumption, physical activity, second-hand smoking, family history of CVD conditions, pre-existing morbidities such as hypertension and diabetes, and use of medication such as antihypertensives and aspirin. For the ecological studies, we assigned a score of − 1 if no confounders were considered. We assigned a score of 0 if temporal trend (calendar time) and area SES indicators, such as employment rate and deprivation index, were controlled for, and we assigned a score of 1 if area smoking rate was additionally adjusted for. Although some studies with a score of 1 adjusted for a number of potential confounders, these were usually assessed only once at baseline and not updated during a long period of follow-up. Thus, residual confounding was still an issue for these studies.

Sensitivity analysis. Lastly, we considered whether analyses were carried out to assess the sensitivity of study findings to various assumptions. We considered studies that conducted sensitivity analyses (including an evaluation of the impact of restricting analyses to certain population groups, or the validation of the statistical or exposure model) to be the most robust; thus, we assigned these studies a score of 1. We assigned a sensitivity analysis score of 0 to studies that did not conduct any sensitivity analyses.

Overall quality score. We added the scores in each category to obtain an overall study quality score for each CV endpoint for each study; if a study evaluated more than one endpoint, it is possible that study has a different score for different endpoints. Based on the overall study quality score, we grouped the studies into two tiers: Tier I (with a score >0) and Tier II (with a score ≤ 0). We consider Tier I studies to be of a higher quality but, because these are crude rankings, we do not differentiate individual studies by their scores in our analysis; rather, we differentiate them by tiers.

Cardiovascular morbidity

We identified seven studies that examined the associations between long-term ozone exposure and CV morbidity endpoints. One study (Chuang et al. 2011) was from the ISA; we identified the other six studies (Atkinson et al. 2013, Beckerman et al. 2012, Breton et al. 2012, Dong et al. 2013a, b, Lipsett et al. 2011) from our literature search. These studies examined arrhythmia, blood pressure, CIMT, prevalence of IHD and CVD, heart failure, stroke, and MI (Table 3). We evaluated the quality of these studies by using the rating criteria described above.

Table 4 summarizes the study quality scores for the CV morbidity studies. We assigned a score of 1 to two large cohort studies (Lipsett et al. 2011, Atkinson et al. 2013) and a score of 0 to five cross-sectional studies (Chuang et al. 2011, Beckerman et al. 2012, Breton et al. 2012, Dong et al. 2013 a, b) for study design.

In our evaluation of the potential presence of selection bias, we assigned a score of 1 to a recent study conducted in Taiwan that included a random sample of participants in a large-scale national survey (Chuang et al. 2011), as we considered that selection bias was unlikely. Three studies had populations restricted to people who lived within a certain distance from an air monitoring station (Lipsett et al. 2011, Dong et al. 2013a, b), and one study included patients from 205 selected English clinical practices that were linked to hospital admission and mortality databases and was deemed to have high quality data (Atkinson et al. 2013). As it is unclear whether the selection of study populations from certain areas or selected clinical practices may be associated with exposure or outcome, we considered that selection bias in these studies was possible, and assigned a score of 0. A Canadian study recruited patients from a single pulmonary clinic (Beckerman et al. 2012), and a US study consisted of students from a single university (Breton et al. 2012). We assigned these two studies a score of − 1, as we considered that selection bias was likely to be present in the study populations.

Regarding the outcome assessment, we assigned a score of 1 to a British study that used a broadly defined endpoint, IHD, to avoid potential errors and misclassification (Beckerman et al. 2012). We assigned a score of 0 to studies in which outcome misclassification was likely to be present, including two studies that relied on databases linking participants to hospital admission records and mortality, with no independent review of medical records or death certificates (Lipsett et al. 2011, Atkinson et al. 2013); we also assigned a score of 0 to two studies in China that relied on self-reported diagnoses or medication use (Dong et al. 2013a, b). In addition, three studies investigated the effects of long-term ozone exposure on CV health using surrogate outcomes such as blood pressure and CIMT (Chuang et al. 2011, Breton et al. 2012, Dong et al. 2013a). These outcomes were measured manually in these studies and thus subject to random or systematic errors that could result in the misclassification of CV outcomes. Because of these limitations, we assigned a score of 0 to these three studies.

None of the studies used personal ozone exposure measurements and thus, none of the studies received a score of 1 for exposure measurement. We assigned a score of 0 to the studies that used inverse distance-weighted (IDW) or dispersion modeling to account for spatial variation (Lipsett et al. 2011, Beckerman et al. 2012, Breton et al. 2012, Atkinson et al. 2013). We also assigned a score of 0 to two Chinese studies that used measurements from local air monitoring stations (within 1 kilometer of participants' residences) (Dong et al. 2013a, b). We assigned a score of − 1 to one study in Taiwan that used area-level ozone concentrations as surrogates for personal exposures (Chuang et al. 2011).

For the evaluation of whether the data analyses employed adequate statistical approaches, we assigned a score of 1 to four studies that used appropriate statistical regression models (as defined above in our scoring criteria) and included co-pollutants, such as particulate matter PM [including fine (up to 2.5 micrometers in diameter) and coarse (up to 10 micrometers in diameter) PM, i.e. $PM_{2.5}$ and PM_{10}] or nitrogen dioxide (NO_2), in the models (Lipsett et al. 2011, Beckerman et al. 2012, Breton et al. 2012, Atkinson et al. 2013). We assigned a score of 0 to two studies that used appropriate regression models, but only considered single pollutant models (Chuang et al. 2011, Dong et al. 2013b). We assigned a score of − 1 to a Chinese study in which the authors inappropriately used logistic regression to evaluate hypertension because the prevalence was greater than 10% and the relative risk was not well estimated by the odds ratios yielded from logistic regression (Dong et al. 2013a).

Regarding whether the statistical analyses considered potential confounders (other than co-pollutants), we assigned a score of 1 to three studies with adequate adjustment for confounders, as defined above in our scoring criteria (Lipsett et al. 2011, Atkinson et al. 2013, Dong et al. 2013a, b), and a score of 0 to three studies with incomplete adjustment (Chuang et al. 2011, Beckerman et al. 2012). We assigned a score of − 1 to a study that examined the association between long-term ozone exposure and development of atherosclerosis that inappropriately adjusted for intermediates [i.e. low-density lipoprotein (LDL) cholesterol and high-density lipoprotein (HDL) cholesterol] in the causal pathway (Breton et al. 2012).

Lastly, we assigned a score of 1 to three studies with sensitivity analyses (Breton et al. 2012, Lipsett et al. 2011, Atkinson et al. 2013) and a score of 0 to those that did not conduct any sensitivity analyses (Chuang et al. 2011, Beckerman et al. 2012, Dong et al. 2013a, b).

Overall, we classified four studies of CV morbidity as Tier I and three as Tier II (Table 4).

Cardiovascular mortality

We identified six studies of long-term ozone exposure and CV mortality from the ISA (Abbey et al. 1999, Pope et al. 2002, Chen et al. 2005, Jerrett et al. 2009, Smith et al. 2009, Wang et al. 2009) and six studies from our literature search (Jerrett et al. 2005, Janke et al. 2009, Krewski et al. 2009, Lipsett et al. 2011, Spencer-Hwang et al. 2011, Carey et al. 2013) (Table 5). We did not include the study by Zanobetti and Schwartz (2011) that was cited in the ISA, as this study examined all-cause mortality, not CV-specific mortality, in individuals with pre-existing CV conditions. Although this study was mentioned in the CV mortality section of the ISA, the EPA did not rely on it for its causal determination.

DOI 10.3109/10408444.2014.937855

Table 3. Epidemiology studies of long-term ozone exposure and cardiovascular morbidity.

Study	Cohort	n	Ozone metric	Outcome	Copollutant(s) in model	Effect estimates	Increment in ozone	Point estimate	95% CI
Arrhythmia									
Atkinson et al. (2013)	Patients identified by the Clinical Practice Research Datalink	836 557 patients	Yearly average	Arrhythmia incidence	None	HR	3 μg/m^3	1.02	0.98–1.05
					PM_{10}			1.02	0.98–1.05
					NO_2			1.02	0.98–1.05
					SO_2			**1.03**	**1.00–1.06**
Blood pressure									
Chuang et al. (2011)	The Social Environmental and Biomarkers of Aging Study	1023 elderly	Yearly average	Systolic blood pressure	None	Increase in mmHg	8.95 ppb	**21.51**	**16.90–26.13**
				Diastolic blood pressure				**20.56**	**18.14–22.97**
Dong et al. (2013a)	Adults from 33 communities in Liaoning Province in northeast China	24 845 adults	3-year average	Systolic blood pressure	None	Increase in mmHg	22 μg/m^3	**0.73**	**0.35–1.11**
				Diastolic blood pressure					
				Systolic blood pressure in men					
				Diastolic blood pressure in men				**0.37**	**0.14–0.61**
				Systolic blood pressure in women				**1.05**	**0.52–1.58**
				Diastolic blood pressure in women				**0.58**	**0.24–0.91**
								0.04	− 0.50–0.58
								0.02	− 0.29–0.34
Hypertension									
Dong et al. (2013a)	Adults from 33 communities in Liaoning Province in northeast China	24 845 adults	3-year average	Hypertension prevalence	None	OR	22 μg/m^3	**1.13**	**1.06–1.20**
				Hypertension prevalence in men				**1.21**	**1.04–1.38**
				Hypertension prevalence in women				1.07	0.91–1.18
				Hypertension prevalence, age < 55 yrs				**1.13**	**1.07–1.20**
				Hypertension prevalence, 55 ≤ age < 65 yrs				1.02	0.92–1.14
				Hypertension prevalence, age ≥ 65 yrs				1.15	0.96–1.39
Carotid artery intima-media thickness									
Breton et al. (2012)	Students of the University of Southern Californifa	861 students	Early childhood (5-year average)	Carotid artery intima-media thickness (CIMT)	None	Absolute difference in μM	10 ppb	7.8	− 0.3–15.9
					NO_2			**10.0**	**1.4–18.6**
					PM_{10}			**8.5**	**0.2–16.9**
					$PM_{2.5}$			**9.1**	**0.9–17.4**
			Elementary school (6-year average)		None		9.3 ppb	**10.1**	**1.8–18.5**
					NO_2			8.8	− 0.1–17.7
					PM_{10}			**10.1**	**1.7–18.5**
					$PM_{2.5}$			**9.5**	**1.1–18.0**
			Lifetime average		None		7.9 ppb	7.5	− 0.8–15.8
					NO_2			7.0	− 1.6–15.6
					PM_{10}			8.0	− 0.4–16.3
					$PM_{2.5}$			7.0	− 1.3–15.3

(*Continued*)

Table 3. (*Continued*)

Study	Cohort	n	Ozone metric	Outcome	Copollutant(s) in model	Effect estimates	Increment in ozone	Point estimate	95% CI
Heart failure									
Atkinson et al. (2013)	Patients identified by the Clinical Practice Research Datalink	836 557 patients	Yearly average	Heart failure incidence	None	HR	3 μg/m^3	**0.94**	**0.90–0.98**
					PM_{10}			**0.95**	**0.91–0.99**
					NO_2			**0.95**	**0.91–0.99**
					SO_2			**0.95**	**0.91–0.99**
Ischemic heart disease									
Beckerman et al. (2012)	Patients from a pulmonary clinic in the Toronto Western Hospital, University Health Network	2360 patients	10-year average	Ischemic heart disease (IHD) prevalence	NO_2, $PM_{2.5}$	RR	NR*		
Cardiovascular disease									
Dong et al. (2013b)	Adults from 33 communities in Liaoning Province in northeast China	24 845 adults	3-year average	CVD prevalence	None	OR	22 μg/m^3	1.09	0.84–1.41
				CVD prevalence in men				1.11	0.78–1.59
				CVD prevalence in women				0.98	0.61–1.48
Stroke									
Lipsett et al. (2011)	The California Teachers Study Cohort	124 614 females	Monthly average	Stroke incidence	None	HR	11.00 ppb	1.02	0.95–1.08
					$PM_{2.5}$		10 ppb	0.92	0.83–1.02
Atkinson et al. (2013)	Patients identified by the Clinical Practice Research Datalink	836 557 patients	Yearly average	Stroke incidence	None	HR	3 μg/m^3	1.00	0.97–1.04
					PM_{10}			0.99	0.96–1.03
					NO_2			1.00	0.96–1.03
					SO_2			1.01	0.98–1.05
Dong et al. (2013b)	Adults from 33 communities in Liaoning Province in northeast China	24 845 adults	3-year average	Stroke prevalence	None	OR	22 μg/m^3	1.15	0.99–1.33
				Stroke prevalence in men				1.16	0.95–1.41
				Stroke prevalence in women				1.14	0.91–1.44
Myocardial infarction									
Lipsett et al. (2011)	The California Teachers Study Cohort	124 614 females	Monthly average	MI incidence	None	HR	11.02 ppb	1.03	0.95–1.11
					$PM_{2.5}$		10 ppb	1.06	0.94–1.19
Atkinson et al. (2013)	Patients identified by the Clinical Practice Research Datalink	836 557 patients	Yearly average	MI incidence	None	HR	3 μg/m^3	**0.96**	**0.93–1.00**
					PM_{10}			**0.95**	**0.92–0.98**
					NO_2			**0.95**	**0.92–0.98**
					SO_2			0.98	0.95–1.02

CI confidence interval, *CIMT* carotid artery intima-media thickness, *CVD* cardiovascular disease, *HR* hazard ratio, *IHD* ischemic heart disease, *MI* myocardial infarction, *n* number of people, NO_2 nitrogen dioxide, *NR* not reported, *OR* odds ratio, *PM* particulate matter, *ppb* parts per billion, *RR* relative risk, SO_2 sulfur dioxide

Bolded values are statistically significant.

*It was noted in the text that no associations were observed between ozone and prevalent IHD.

Table 4. Study quality – epidemiology studies of cardiovascular morbidity.*

Study	Study design	Selection bias	Outcome assessment	Exposure measurement	Statistical modeling	Control for confounders	Sensitivity analysis	Total score
Breton et al. (2012)	0	−1	0	0	1	−1	1	0
Chuang et al. (2011)	0	1	0	−1	0	0	0	0
Dong et al. (2013a)	0	0	0	0	−1	1	0	0
Beckerman et al. (2012)	0	−1	1	0	1	0	0	1
Dong et al. (2013b)	0	0	0	0	0	1	0	1
Atkinson et al. (2013)	1	0	0	0	1	1	1	4
Lipsett et al. (2011)	1	0	0	0	1	1	1	4

*The study quality scoring system used to determine these scores is described in the text.

Although some of the studies we included in this analysis evaluated mortality from all CV or cardiopulmonary conditions combined, several investigated specific CV mortality endpoints, such as MI or coronary heart disease (CHD) (Chen et al. 2005, Jerrett et al. 2005, 2009, Krewski et al. 2009, Lipsett et al. 2011, Spencer-Hwang et al. 2011). All of the studies analyzed mortality in healthy adults, with the exception of one study that evaluated the effect of ozone on mortality in kidney transplant recipients (Spencer-Hwang et al. 2011). Thus, these results are specific to this susceptible population. To evaluate studies of long-term ozone exposure and CV mortality, we used the same study quality scoring system described above for the studies of CV morbidity.

Table 6 summarizes the study quality scores for the studies of CV mortality. For study design, we assigned a score of 0 to the ecological and cross-sectional analyses and 1 to the cohort designs. For selection bias, because most of the studies were of large cohorts with high rates of verified follow-up, or were ecological/panel studies that include all deaths in a particular area, we considered that selection bias was unlikely and assigned a score of 1 to these studies (Jerrett et al. 2005, Janke et al. 2009, Wang et al. 2009, Lipsett et al. 2011). We assigned a score of 0 to eight studies because they had some potential for selection bias, most often due to the study inclusion criteria (Abbey et al. 1999, Pope et al. 2002, Chen et al. 2005, Jerrett et al. 2009, Krewski et al. 2009, Smith et al. 2009, Spencer-Hwang et al. 2011, Carey et al. 2013). For example, Spencer-Hwang et al. (2011) restricted the study population to patients who had their first kidney transplant and lived within 50 km of an air monitoring station; it is possible that this populations' ozone exposure and risk of mortality may have differed from those who were not included. Similarly, Carey et al. (2013) only included patients of a select group of general practitioners in England, and the selection may have been related to exposure or mortality risk.

For outcome assessment, we assigned a score of 1 to studies that used death certificates to identify deaths from CV disease (Abbey et al. 1999, Pope et al. 2002, Chen et al. 2005, Krewski et al. 2009, Smith et al. 2009, Carey et al. 2013). Other studies used registry data and did not verify the results using death certificates (Jerrett et al. 2005, Janke et al. 2009, Wang et al. 2009, Lipsett et al. 2011, Spencer-Hwang et al. 2011); thus, we assigned a score of 0 to these studies. Registry data can be unreliable; while some countries have complete death registries with the causes of death reported *via* physicians, other countries have incomplete or inaccurate registry systems (Khosravi et al. 2008). For example, China's surveillance system has no standard procedures or instruments for identifying and reporting causes of death (Wang et al. 2007). Jerrett et al. (2009) used death certificates to identify deaths through 1995 but used the National Death Index beginning in 1996, so we assigned a score of 0 for outcome assessment to this study.

The majority of the mortality studies used area ambient monitoring data from either a single monitor or a small number of monitors within each city. We assigned these studies a score of −1 for exposure measurement. Several studies interpolated the available air monitoring data or utilized exposure modeling, which allowed more precise exposure estimates; we assigned these studies a score of 0 (Abbey et al. 1999, Chen et al. 2005, Jerrett et al. 2005, Janke et al. 2009, Wang et al. 2009, Spencer-Hwang et al. 2011, Carey et al. 2013). None of the studies used personal monitoring data, so we did not assign a score of 1 to any of the studies.

All studies used appropriate statistical models [e.g. Cox regression models for the cohort studies and a generalized estimating equations model for the ecological, cross-sectional studies]; however, not all studies performed bi- or multi-pollutant analyses. Because certain co-pollutants may confound the ozone-mortality relationship, and numerous studies have reported an attenuation of mortality effect estimates for ozone when co-pollutants were included in statistical models (Pascal et al. 2012, Tao et al. 2012, Katsouyanni et al. 2009, Yang et al. 2012), we considered studies that included multi-pollutant analyses as more robust than those that did not. Among the 12 studies identified, we assigned a score of 1 to those studies that performed bi- or multi-pollutant analyses and a score of 0 to all others (Table 6).

All 12 studies controlled for at least some potential confounders. We assigned a score of 1 to studies with adequate adjustment for confounders, as defined above in our scoring criteria (Pope et al. 2002, Jerrett et al. 2005, 2009, Janke et al. 2009, Krewski et al. 2009, Smith et al. 2009, Lipsett et al. 2011), and a score of 0 to studies with incomplete adjustment for confounders (Abbey et al. 1999, Chen et al. 2005, Wang et al. 2009, Carey et al. 2013). We assigned the study by Spencer-Hwang et al. (2011) a score of −1, because the authors controlled for numerous kidney-related factors but not physical activity, BMI, diet, or other important CV disease risk factors.

Finally, all but two studies (Smith et al. 2009, Wang et al. 2009) performed at least some sensitivity analyses to assess the validity of statistical assumptions and/or the robustness of results. We assigned a score of 0 to the studies with no sensitivity analyses and a score of 1 to those with at least one analysis.

Table 5. Epidemiology studies of long-term ozone exposure and cardiovascular-related mortality.

Study	Cohort	N	Time period of Mortality analysis	Time period of Ozone data	Seasonal or all year	Ozone metric	Outcome	Copollutant(s) in model	Risk meas.	Unit of meas.	Estimate	95% CI
Cohort Studies												
Abbey et al. (1999)	AHSMOG cohort of non-smokers (CA)	6338	1982–1998	1982–1998	All year	Monthly avg.	Cardiopulmonary	None	RR (M)	12 ppb	1.08	0.91–1.29
									RR (F)		0.97	0.84–1.12
									RR (M)	> 100 ppb	1.06	0.88–1.29
									RR (F)		0.88	0.75–1.02
Pope et al. (2002)	ACS cohort, 134 metropolitan areas	500 000	1982–1998	1980–1998	All year	1-h max	Cardiopulmonary*	None	RR	NA	1.10	0.95–1.21
					July-Sept						1.11	0.99–1.20
Chen et al. (2005)	AHSMOG cohort (CA)	3239	1973–1998	1977–1998	All year	Monthly avg.	CHD	None	RR (M)	10 ppb	0.89	0.60–1.30
									RR (F)		0.97	0.68–1.38
Jerrett et al. (2005)	ACS, Los Angeles only	22 905	1982–2000	1999–2001	All year	Peak Max 8hr†	IHD	None	RR	NA	0.97	0.93–1.02
											0.98	0.95–1.02
						Peak Max 8hr†	Cardiopulmonary				**0.97**	**0.94–0.99**
											0.99	0.96–1.01
Janke et al. (2009)	354 local authorities in the UK	Avg.: 140 000	1998–2005	1998–2005	All year	8-h max	All circulatory	PM_{10}, NO_2, CO	RR‡	5.1 ppb	1.00	NR
							CHD				1.00	NR
							MI				0.99	NR
Jerrett et al. (2009)	ACS cohort, 86 metropolitan areas (US)	448 850	1982–2000	1977–2000; 1999–2000 (PM)	Summer only (April-Sept)	1-h max	Cardiopulmonary	None	RR¶	10 ppb	**1.016**	**1.01–1.02**
								PM			0.992	0.98–1.00
							CV	None			**1.014**	**1.01–1.02**
								PM			**0.983**	**0.97–0.99**
							IHD	None			**1.017**	**1.01–1.03**
								PM			**0.973**	**0.96–0.99**
Krewski et al. (2009)	ACS cohort, 116 metropolitan areas	1.2 mil	1982–2000	1980	All year	8-h max	Cardiopulmonary	None	HR	10 ppb	1.01	1.00–1.03
					Summer (April-Sept)						**1.03**	**1.02–1.04**
					All year			$PM_{2.5}$			0.99	0.96–1.01
					All year		IHD	None			1.01	0.98–1.03
					Summer (April-Sept)						1.01	0.99–1.02
					All year			$PM_{2.5}$			0.98	0.95–1.02
Smith et al. (2009)	ACS cohort, 66 cities	352 000	1982–2000	NR	Warm season (second and third quarters)	8-h max	Cardiopulmonary	None	RR	10 ppb	**1.03‡**	**1.01–1.05**
				2003–2005								
								Carbon and sulfate			**1.02‡**	**1.01–1.04**
								Carbon			1.02‡	1.00–1.04
Lipsett et al. (2011)	California Teachers Study (F only)	101 784	1997–2005	1996–2005	All year§	Monthly avg.	CV	None	HR	10 ppb	1.00	0.94–1.07
								$PM_{2.5}$			0.97	0.90–1.05
							IHD	None			1.07	0.97–1.17
								$PM_{2.5}$			0.99	0.88–1.11
							Cerebrovascular	None			1.02	0.89–1.16
								$PM_{2.5}$			0.97	0.82–1.14
					Summer-only		CV	None		22.96 ppb	1.02	0.96–1.07
							IHD				**1.09**	**1.01–1.19**
							Cerebrovascular				0.99	0.88–1.10

(*Continued*)

DOI 10.3109/10408444.2014.937855

Table 5. (*Continued*)

Study	Cohort	N	Time period of Mortality analysis	Time period of Ozone data	Seasonal or all year	Ozone metric	Outcome	Copollutant(s) in model	Risk meas.	Unit of meas.	Estimate	95% CI
Spencer-Hwang et al. (2011)	Cohort of kidney transplant recipients (US, 50 states)	32 239	1997–2003	1997–2003	All year	Monthly avg.	CHD	None	RR	10 ppb	**1.35**	**1.01–1.81**
								PM_{10}			**1.34**	**1.01–1.79**
Carey et al. (2013)	Cohort of GP patients in England	824 654	2003–2007	2002	All year	8-h max	Circulatory‖	None	HR	3.0 μg/m^3	**0.96**	**0.94–0.98**
Ecological studies												
Janke et al. (2009)	354 local authorities in the UK	Avg.: 140 000	1998–2005	1998–2005	All year	8-h max	All circulatory	PM_{10}, NO_2, CO	RR‡	5.1 ppb	1.00	NR
							CHD				1.00	NR
							MI				0.99	NR
Wang et al. (2009)	Cohort in Brisbane, Australia	887 955	1996–2004	1996–2004	All year	1-h max	Cardiorespiratory	None	RR	1 ppb	1.002	0.99–1.02
								NO_2 and SO_2			0.999	0.99–1.01

ACS American Cancer Society, *AHSMOG* Adventist Health Study of Smog, *avg*. average, *CA* California, *CHD* coronary heart disease, *CI* confidence interval, *CO* carbon monoxide, *CV* cardiovascular, *F* female, *GP* general practitioner, *HR* hazard ratio, *IHD* ischemic heart disease, *M* male, *max* maximum, *meas*. measure, *MI* myocardial infarction, *mil* million, *n* number of people, *NA* not available, NO_2 nitrogen dioxide, *NR* not reported, *PM* particulate matter, *ppb* parts per billion, *RR* relative risk, SO_2 sulfur dioxide, *UK* United Kingdom, *US* United States

Bolded values are statistically significant.

*RR estimated based on Figure 5 in Pope et al. (2002).

†Average of the four highest 8-h values.

‡Converted from % excess risk to an RR using the equation "e^(% risk) = R."

¶Results based on analysis of 86 cities for both single- and two-pollutant models.

§All-year data is from participants who had both ozone and $PM_{2.5}$ data during the exposure time period.

‖Circulatory = ICD-10 codes I00–I99.

Table 6. Study quality – epidemiology studies of cardiovascular-related mortality.*

Study	Study design	Selection bias	Outcome assessment	Exposure measurement	Statistical modeling	Control for confounders	Sensitivity analysis	Total score
Cohort Studies								
Smith et al. (2009)	0	0	1	−1	1	1	0	2
Spencer-Hwang et al. (2011)	1	0	0	0	1	−1	1	2
Abbey et al. (1999)	1	0	1	0	0	0	1	3
Carey et al. (2013)	1	0	1	0	0	0	1	3
Chen et al. (2005)	1	0	1	0	0	0	1	3
Jerrett et al. (2005)	0	1	0	0	0	1	1	3
Jerrett et al. (2009)	1	0	0	−1	1	1	1	3
Pope et al. (2002)	1	0	1	−1	0	1	1	3
Krewski et al. (2009)	1	0	1	−1	1	1	1	4
Lipsett et al. (2011)	1	1	0	0	1	1	1	5
Ecological studies								
Wang et al. (2009)	0	1	0	0	1	0	0	2
Janke et al. (2009)	0	1	0	0	1	1	1	4

*The study quality scoring system used to determine these scores is described in the text.

Overall, we classified all 12 studies of CV mortality as Tier I (Table 6).

Cardiovascular biomarkers

There is currently no known biologically plausible mechanism by which ambient ozone could cause CV effects. Ozone reacts on contact with the respiratory tract lining fluids, so it is not transported to extrapulmonary sites (Hatch et al. 1994, Medinsky 1996, Barath et al. 2013); however, it is possible that the products of ozone reaction may enter the circulation. Several studies have examined the circulating levels of biomarkers in humans and experimental animals after long-term exposure to ozone, to determine potential biological mechanisms for an association between ozone exposure and CV effects. The most commonly studied biomarkers are those associated with inflammation, coagulation, oxidative stress, blood lipids and glucose metabolism, and overall CV health. Each biomarker was examined in only one or two studies, so it is not feasible to compare the effects of ozone on specific biomarkers across studies in most cases; however, we evaluated the effects on biomarkers in the same biological pathway.

A limitation of biomarker studies is that some biomarkers may not be clinically relevant (i.e. associated with disease). Many of the biomarkers examined in the studies reviewed here are used as potential indicators of CV risk based on hypotheses suggested from small, retrospective studies assessing CVD risk factors, but they require further evaluation before they can be utilized in clinical (i.e. diagnostic) practice (Lewington et al. 2012, Ioannidis and Tzoulaki 2012, Sarwar et al. 2009). It is possible that many of the biomarkers are not indicative of adverse CV effects but instead are indicative of nonadverse biological changes (e.g. homeostatic changes) that may or may not lead to adverse effects in a given individual. Many factors related to study design can influence the measured concentrations of circulating biomarkers, including the time of day that blood is collected from participants (due to the influence of circadian patterns), their dietary intake patterns (such as after a high-fat meal), and their level of physical activity (Zhou et al. 2010).

We identified three epidemiology studies, each included in the ISA, that examined potential CV biomarkers (Chen et al. 2007, Chuang et al. 2011, Forbes et al. 2009) (Table 7). To evaluate these studies, we used the same study quality scoring system described above, but with different criteria for scoring outcome assessment. We assigned a score of 1 to studies that employed quality assurance/quality control procedures in sample storage and laboratory assays and reported high reliability of the results (intraclass correlation coefficient $>75\%$ or coefficient of variation $<10\%$), as we considered the influence of measurement errors to be small in such settings. We assigned a score of 0 to studies that did not receive a score of 1, as we considered that the measurement errors and subsequent outcome misclassifications were likely to be substantial.

Table 8 summarizes the study quality scores for the studies of CV biomarkers. For study design, we assigned a score of 0 to all three biomarker studies because of their cross-sectional design. For the potential presence of selection bias, we assigned a score of 1 to two studies that included representative samples of participants in large-scale national surveys (Forbes et al. 2009, Chuang et al. 2011), as we considered that selection bias was unlikely in such settings. We assigned a score of −1 to the study by Chen et al. (2007) that recruited participants from a single university, because selection bias was likely to be present in the study population.

For outcome assessment, one study employed strict storage and analytical procedures for blood samples, performed quality assurance/quality control analyses, and had low variability (coefficients of variation $<10\%$) in biomarker assays (Chen et al. 2007). Another study indicated good interlaboratory and intralaboratory reliability (intraclass correlations $>75\%$) (Chuang et al. 2011). We assigned a score of 1 to these two studies, as we considered measurement errors in the outcome assessment were likely to be small. Forbes et al. (2009) stored blood samples at room temperature and reported coefficients of variation $>10\%$. We assigned a score of 0 to this study, as we considered measurement errors in the outcome assessment were likely to be large.

For exposure measurement, we assigned a score of 0 to two studies that employed IDW and dispersion models to account for spatial variation and interpolate individual exposure (Chen et al. 2007, Forbes et al. 2009) and a score of −1 to the study by Chuang et al. (2011), which used averaged area-level ozone concentrations as surrogates for individual exposure.

For statistical modeling, we assigned a score of 0 to all three studies for using appropriate statistical regressions (as defined

previously above) but only single-pollutant models. We also assigned a score of 0 to all three studies for confounder adjustment, as all three had incomplete control for potential confounders (as defined above in our scoring criteria).

For the sensitivity analysis, we assigned a score of 1 to two studies that performed a variety of sensitivity analyses (Chen et al. 2007, Forbes et al. 2009) and a score of 0 to one study that did not perform any (Chuang et al. 2011).

Overall, we classified all three of the studies of CV biomarkers as Tier I (Table 8).

In our evaluation of study quality, we also considered additional factors that can influence biomarker measurements. The participants in all studies were healthy, although those in the study by Chuang et al. (2011) were at least 54 years of age. None of the studies noted whether blood was collected from participants at the same time of the day to limit diurnal variations in biomarker levels. Non-fasting blood samples were collected from participants in the study by Forbes et al. (2009), whereas the other two studies did not state whether fasting or non-fasting samples were used; however, fasting glucose was measured in the study by Chuang et al. (2011). All three studies were classified as Tier I, but we considered studies to be more reliable if they measured biomarkers in fasting blood samples.

Experimental animal studies

Cardiovascular morbidity

We identified two experimental animal studies of potential CV morbidity effects of long-term ozone exposure from the ISA (Chuang et al. 2009, Perepu et al. 2010) and three from our literature search (Perepu et al. 2012, Gordon et al. 2013, Sethi et al. 2012). These studies examined the effects on heart rate, blood pressure, atherosclerotic lesion development, and vascular function in rodents exposed to high ozone concentrations (500–800 ppb) (Table 9).

Many factors contribute to the quality of experimental animal studies. Recently, several investigators proposed guidelines for reporting animal data in primary studies to improve the quality of scientific publications (Macleod et al. 2009, Hooijmans et al. 2010, Kilkenny et al. 2010, van der Worp et al. 2010). We considered the information on study design and methods (including the number of animals, appropriate controls, and maintenance of the exposure concentration) proposed by these guidelines, as well as the use of appropriate statistical methods and replicability of observations, when assessing the quality of the ozone experimental animal studies. We used these factors to develop the criteria for our quantitative study quality scoring system, in order to classify the experimental animal studies as Tier I or Tier II. Table 10 summarizes the study quality scores for the experimental animal studies of CV morbidity.

Our criteria for exposure assignment were to assign a score of 1 to studies that explicitly stated that animals were randomized to treatment or control groups and a score of − 1 to those that did not. None of the experimental animal studies of CV morbidity specified how the treatment assignment was performed; thus, we assigned a score of − 1 to all five studies.

We defined an adequate experimental group as having at least ten animals of each sex and a clear description of different treatment groups, as per the EPA guidelines for 90-day inhalation toxicity studies (US EPA 1998), unless otherwise

Table 7. Epidemiology studies of long-term ozone exposure and cardiovascular-related biomarkers.

Study	Cohort	n	Ozone metric	Co-pollutant(s)	Effect estimates	Increment in ozone	Biomarker	Point estimate	95% CI
Chen et al. (2007)	First-year undergraduates at UC Berkeley with lifelong residence in either Los Angeles or San Francisco	120	Monthly average	None	Coefficient (β)	Lifetime exposure (ppb)	8-iso-PGF (pg/mL)	**0.023**	**0.006***
				$PM_{10-2.5}$, NO_2					NR†
				None			FRAP (pg/mL)	− 2.21	NR ($p = 0.45$)
				$PM_{10-2.5}$, NO_2					NR†
Chuang et al. (2011)	The Social Environmental and Biomarkers of Aging Study	1023 elderly	Yearly average	None	Increase in concentration	8.95 ppb	IL-6 (pg/mL)	0.14	− 0.18 to 0.46
							Neutrophils (%)	**13.74**	**11.50 to 15.99**
							Total cholesterol (mg/dL)	**56.47**	**47.26 to 65.69**
							HDL-cholesterol (mg/dL)	− 0.48	− 3.69 to 2.73
							Triglycerides (mg/dL)	2.13	− 18.10 to 22.36
							Fasting glucose (mg/dL)	**21.1**	**12.03 to 30.17**
							HbA1c (%)	**1.3**	**0.97 to 1.63**
Forbes et al. (2009)	Health Survey for England (White ethnic groups, ages 16 and older)	26 510	Yearly average	None	Percent change in concentration	1 μg/m^3 (0.5 ppb)	Fibrinogen	− 0.014	− 0.123 to 0.095
		17 566		None			CRP	− 0.155	− 0.874 to 0.569

CI confidence interval, *CRP* C-reactive protein, *FRAP* ferric reducing ability of plasma, *HbA1c* hemoglobin A1c, *HDL* high-density lipoprotein, *IL-6* interleukin-6, *n* number of people, NO_2 nitrogen dioxide, *NR* not reported, *PM* particulate matter, *PGF* prostaglandins-$F_{2\alpha}$, *ppb* parts per billion

Bolded values are statistically significant.

*Standard error.

†The authors noted that inclusion of the co-pollutants in the models did not change the magnitude of the association with ozone, but no results were shown.

Table 8. Study quality – epidemiology studies of cardiovascular-related biomarkers.*

Study	Study design	Selection bias	Outcome assessment	Exposure measurement	Statistical modeling	Control for confounders	Sensitivity analysis	Total score
Chen et al. (2007)	0	−1	1	0	0	0	1	1
Chuang et al. (2011)	0	1	1	−1	0	0	0	1
Forbes et al. (2009)	0	1	0	0	0	0	1	2

*The study quality scoring system used to determine these scores is described in the text.

justified (e.g. with a power calculation). We chose to base these criteria on the EPA's 90-day study guidelines because the studies reviewed here were of similar durations. We assigned a score of 1 when these conditions were met; otherwise, we assigned a score of −1 for this category. None of the experimental animal studies met this criterion, so we assigned them all a score of −1. Similarly, we evaluated studies on whether appropriate controls were employed. We assigned a score of 1 to studies that clearly indicated that treated animals were compared to a control group exposed to filtered air. If studies did not include a filtered air control group, or if controls were exposed to "room" or "ordinary" air (which could contain contaminants with the potential to alter results), we assigned a score of −1. All five studies used a filtered air control group, so we assigned them a score of 1.

The validity of chamber study results depends on the experimental conditions, maintenance of the chamber environment, the density of animals in each chamber (to minimize the effects of animal surface area or volume on exposure concentration), cleanliness, and control of potential animal stressors (e.g. temperature, humidity, air flow, atmospheric pressure, noise, and vibration) (Klaassen 2008, Dorato and Wolff 1991). We assigned a score of 1 to studies that explicitly mentioned the measures taken to ensure the accuracy and consistency of ozone exposure throughout the exposure period (e.g. continuous monitoring of ozone concentration), type of exposure method used (e.g. chamber or nose-only), and maintenance of adequate environmental conditions. We assigned a score of −1 if these parameters were not clearly mentioned in the study. We assigned all five studies a score of 1 for exposure environment.

Table 9. Experimental animal studies of long-term ozone exposure and cardiovascular morbidity.

Study	Endpoint	n	Species	Exposure design	Age at start of study	Ozone conc. (ppb)	Result	*p* value
Heart Rate								
Gordon et al. (2013)*	Heart rate (bpm)	7–8 exposed, 7–8 control	Rat	6 hr/day, 1 day/wk, 15 wks; chamber	4 mos	0	391.2 ± 12	>0.05
						800	379.8 ± 12	
					Senescent (20 mos)	0	379.0 ± 3.9	>0.05
						800	350.1 ± 11.4	
Cardiovascular Structure and Function								
Chuang et al. (2009)*	% oil red-O staining of aorta	≥5 exposed, ≥5 control	apoE −/− mouse	8 hr/day, 5 days/wk for 8 wks; chamber	6 wks	FA	0.70 ± 0.2	
						500	**1.50 ± 0.19**	**<0.05**
Sethi et al. (2012)*	LVDP (mmHg)	4 exposed, 4 control	Rat	8 hr/day for 56 days; chamber	Adult	FA	152.3 ± 3.8	
						800	**69.0 ± 15**	**<0.05**
Perepu et al. (2010)*	LVDP (mmHg)	6 exposed, 6 control	Rat; *ex vivo*†	8 hr/day for 56 days; chamber	Adult	FA	72.2 ± 4.9	
						800	**26.1 ± 1.1**	**<0.05**
	LVEDP (mmHg)					FA	19.0 ± 0.3	
						800	**39.1 ± 2.4**	**<0.05**
	+ dp/dT (mmHg/s)					FA	61.1 ± 6.1	
						800	**34.3 ± 4.0**	**<0.05**
	− dp/dT (mmHg/s)					FA	60.1 ± 5.0	
						800	**33.28 ± 2.0**	**<0.05**
Perepu et al. (2012)*	LVDP (mmHg)	6 exposed, 6 control	Rat	8 hr/day for 56 days; chamber	Adult	FA	151 ± 8.0	
						800	**71.3 ± 8.0**	**<0.05**
	LVEDP (mmHg)					FA	3.75 ± 0.7	
						800	**14.2 ± 1.6**	**<0.05**
	+ dp/dT (mmHg/s)					FA	6718 ± 354	
						800	**3845 ± 309**	**<0.05**
	− dp/dT (mmHg/s)					FA	5560 ± 423	
						800	**2955 ± 192**	**<0.05**

bpm beats per minute, *FA* filtered air, *LVDP* left ventricular developed pressure, *LVEDP* left ventricular end diastolic pressure, *n* number of animals, *ppb* parts per billion

Bolded values are statistically significant.

*Estimated from study figures using GetData Graph Digitizer.

†Rats were continuously exposed to ozone for 8 hr/day for 56 days; after sacrifice, hearts were removed and subjected to 30 min of global ischemia followed by 60 min of reperfusion. Endpoints were measured after ischemia/reperfusion.

Information on animal husbandry and adequate animal housing is also integral to a reliable animal study. We assigned a score of 1 to studies that reported on the source, age, housing, and feeding of animals, as well as the treatment conditions, acclimation period, and sacrifice methods (if applicable). If more than one of these details were missing or no information was provided, we assigned a score of − 1 for this category. Four studies provided adequate information on animal husbandry and housing conditions according to our criteria (Chuang et al. 2009, Perepu et al. 2010, 2012, Gordon et al. 2013) and we assigned a score of 1 to these studies. One study missed more than one of these details and we assigned a score of − 1 to this study (Sethi et al. 2012).

We also considered detailed descriptions of outcome measures, accompanied by consideration of the adequacy of the methods and/or the reproducibility of measures, as important study quality criteria. We assigned a score of 1 for outcome assessment if the authors provided details on how outcomes were measured and the reproducibility/adequacy of such measurements. For the animal studies that measured biomarkers, this included details on the assays and/or kits used to measure specific biomarkers, and the collection, handling, and storage of samples. We assigned a score of − 1 in the absence of any of these details. Four of the CV morbidity studies provided adequate information on outcome assessment and quality assurance according to our criteria (Chuang et al. 2009, Perepu et al. 2010, 2012, Gordon et al. 2013) and we assigned these studies a score of 1. One study did not provide adequate information and we assigned a score of − 1 for outcome assessment to this study (Sethi et al. 2012).

We also considered attrition bias, a type of selection bias caused by a differential loss of animals in exposed *versus* control groups. We assigned a score of 1 if the details regarding study-related deaths were provided, either explicitly or in such a way that they could be easily derived from study information (i.e. if a starting number of animals was given in conjunction with the number of animals in the final analyses). If no information was provided or was not discernible, we assigned a score of − 1 for attrition bias. The study by Gordon et al. (2013) provided mortality information and explained cause of death, so we assigned a score of 1 to this study. The other four studies did not provide information to determine whether any animals died during the study, so we assigned a score of − 1 to these studies.

For statistical analyses, we assigned a score of 1 if appropriate statistical methods were used and clearly denoted (e.g. t-tests accounting for multiple comparisons; analysis of variance (ANOVA) followed by the Newman-Keuls test). Specifically, to receive a score of 1, studies that compared results between exposed and control groups must have accounted for the effects of multiple comparisons, and studies in which effects in exposed animals were compared to their own baseline (pre-exposure) values must have employed methods to account for possible correlation between repeated measures (e.g. mixed effects models). In addition, our criteria for a score of 1 require the inclusion of standard errors/deviations, baseline/control results, and data for all relevant time points in the presentation of results. We assigned a score of − 1 if inappropriate statistical tests were used (e.g. t-tests with no post-hoc analyses) or if the data required for a score of 1 were missing. In all of the experimental animal studies of CV morbidity, the authors conducted adequate statistical methods according to our criteria, so we assigned a score of 1 to all studies.

Overall, we classified four experimental animal studies of CV morbidity as Tier I and one as Tier II (Table 10).

In addition to the study quality, we considered factors related to the relevance of the studies to humans, including the animal species and exposure levels used. All studies used rats except for the study by Chuang et al. (2009), which used mice. All studies were conducted *in vivo* with the exception of the study by Perepu et al. (2010); this study was an *ex vivo* study on isolated rat hearts that were subjected to 30 min of global ischemia followed by 1 h of reperfusion before CV endpoints were measured. The exposure duration across the studies was mainly subchronic. Chuang et al. (2009), Perepu et al. (2010), Perepu et al. (2012), and Sethi et al. (2012) exposed the animals for 56 days (~8 weeks); however, the latter three studies exposed the animals 7 days per week, while Chuang et al. (2009) exposed the animals for 5 days per week followed by 2 days of filtered air. Gordon et al. (2013) simulated long-term intermittent exposure, exposing rats for only 6 hours per day and 1 day per week for 15 weeks. The authors measured CV endpoints every other week, 1 day after exposure; thus, the results may not be relevant to the assessment of long-term effects of ozone, and rather, may reflect acute effects.

While all studies evaluated the effect of ozone on healthy adult rodents, Chuang et al. (2009) utilized apolipoprotein E (apoE) -/- C57Bl/6 mice. While the C57Bl/6 strain does not usually develop atherosclerotic lesions (the endpoint being investigated), the apoE −/− mice of this strain are susceptible to such lesions. In addition, Gordon et al. (2013) examined the effects of ozone in rats that were either 8 or 24 months old (senescent) at the end of the 4-month exposure period; the results of this study may be useful in determining the poten-

Table 10. Study quality – experimental animal studies of cardiovascular morbidity.*

	Study design								
Study	Exposure assignment	Appropriate control	Experimental groups	Experimental animals, housing and husbandry	Exposure environment	Outcome assessment	Attrition bias	Statistical methods	Total score
Sethi et al. (2012)	− 1	1	− 1	− 1	1	− 1	− 1	1	− 2
Chuang et al. (2009)	− 1	1	− 1	1	1	1	− 1	1	2
Perepu et al. (2010)	− 1	1	− 1	1	1	1	− 1	1	2
Perepu et al. (2012)	− 1	1	− 1	1	1	1	− 1	1	2
Gordon et al. (2013)	− 1	1	− 1	1	1	1	1	1	4

*The study quality scoring system used to determine these scores is described in the text.

tial for effects in sensitive populations. None of the studies measured CV effects at more than one exposure concentration, and the range of concentrations tested among the studies was small; thus, the studies preclude the assessment of dose-response relationships. The exposure concentrations in each study were an order of magnitude higher than the current ozone NAAQS of 75 ppb.

Healthy rats and mice undergo a significant reduction in heart rate (bradycardia), body temperature, blood pressure, and oxygen demand upon exposure to numerous airborne pollutants, including ozone (Watkinson et al. 2001). The decrease in heart rate and blood pressure reduces tissue perfusion, causes an increase in arrhythmic events (particularly type II atrioventricular node block and premature depolarization), and causes other downstream CV effects. This "hypothermic" response is not observed in humans, except after acute episodes of poisoning and drug overdose (Watkinson et al. 2003). However, studies have shown that the hypothermic response is no longer observed when the exposure lasts longer than 2 days (Watkinson et al. 2003); thus, it may not be applicable to long-term studies of ozone exposures in mice and rats. Of the five studies of CV morbidity reviewed here, only the study by Gordon et al. (2013) measured heart rate and blood pressure; therefore, it is difficult to assess how the hypothermic response may or may not have affected CV endpoints after long-term ozone exposure in rats and mice.

Other species differences, such as the varying nasal structures in rodents compared to humans and the fact that rodents breathe only through the nose, may also limit extrapolation of the results in rodents to humans. Further, rodents have a higher ventilation rate and higher ratio of body surface area/body volume and breathe more air, which would be expected to increase their internal dose of inhaled ozone (Hatch et al. 2013). Although anatomical differences cause rodents to remove a smaller fraction of the inhaled amount of ozone than humans (Miller 1995, Perepu et al. 2010), the high concentrations used in rodent studies may still limit the generalizability of results to humans. Because of differences in baseline heart rate, responses to ischemia (restricted blood supply) are different in humans and rodents (Perepu et al. 2010). In contrast to rodents, humans and other large mammals have higher cardiac muscle levels of inhibitory factor 1 (IF_1), which decreases the rate of tissue ATP depletion during ischemia, thereby delaying cell injury and death (Rouslin et al. 1995). In addition, CV diseases such as hypertension and heart failure are usually slow to develop in humans. Temporality of onset is difficult to replicate in animals, which often show a rapid development of symptoms under typical experimental conditions (Doggrell and Brown 1998); this may be why we were unable to identify any experimental animal studies of ozone and CV effects with chronic exposure durations.

Cardiovascular biomarkers

We identified five experimental animal studies that examined potential CV biomarkers in the circulation or cardiac tissue (Table 11). Two of these were included in the ISA (Perepu et al. 2010, Kodavanti et al. 2011) and three were identified from our literature search (Gordon et al. 2013, Perepu et al. 2012, Sethi et al. 2012). We assessed the quality of each study by considering the same evaluation criteria as for the experimental animal studies of CV effects discussed above, as well as other factors that can influence biomarker measurements. Table 12 summarizes the study quality scores for the experimental animal studies of CV biomarkers.

For each scoring category, we assigned the same scores as for CV morbidity to the studies by Perepu et al. (2010, 2012) and Gordon et al. (2013); thus, these studies were all classified as Tier I studies. The study by Sethi et al. (2012) was also scored the same across categories as for CV morbidity, with the exception of the outcome assessment category. The authors provided adequate information on the outcome assessment and quality assurance for the biomarker outcomes, according to our criteria, so we assigned a score of 1 for outcome assessment to this study. The change in scoring for this category increased the total score for this study, but is was still classified as a Tier II study. We assigned a score of 1 to all categories, except for exposure assignment, to the study by Kodavanti et al. (2011). Kodavanti et al. (2011) did not specify how the treatment assignment was performed, so we assigned a score of -1 for exposure assignment. This study met our criteria for the category of experimental group; used a filtered air control group for result comparison; provided adequate information regarding the exposure environment, animal husbandry and housing conditions, outcome assessment, quality assurance, and attrition bias; and conducted adequate statistical methods according to our criteria. Overall, we classified four experimental animal studies of CV biomarkers as Tier I and one as Tier II (Table 12).

Each experimental animal study of CV biomarkers was conducted using healthy adult rats, and the study by Gordon et al. (2013) also included senescent (20-month-old) rats. Each study was conducted *in vivo*, with the exception of the study by Perepu et al. (2010) which was an *ex vivo* study on isolated rat hearts subjected to ischemia and reperfusion before the biomarker levels were measured. We considered the *in vivo* studies to be more relevant than the *ex vivo* study. The exposure durations ranged from 8 hours per day for 56 days in the studies by Perepu et al. (2010, 2012) to intermittent exposures (1 day per week) over 16 (Kodavanti et al. 2011) or 17 weeks (Gordon et al. 2013). As the biomarkers were measured 1 or 2 days after the last exposure in the studies by Kodavanti et al. (2011) and Gordon et al. (2013), with a recovery period of 1 week between each exposure, the results of these studies may not be relevant to the assessment of the long-term effects of ozone. All studies evaluated only one exposure concentration: 500 ppb ozone in the study by Kodavanti et al. (2011) and 800 ppb ozone in the other three studies. The rats had access to food and water *ad libitum* throughout each study, and none of the studies reported whether blood was collected from all rats at a particular time of day.

Phase 3 – Integration and evaluation of evidence

Epidemiology studies

Cardiovascular morbidity

We classified four studies that examined the effects of long-term ozone exposure on CV morbidity outcomes as Tier I and three as Tier II. Table 3 presents the results of these studies.

Table 11. Experimental animal studies of cardiovascular-related biomarkers.

Study	Endpoint	n	Species	Exposure design	Age at start of study	Ozone conc. (ppb)	Result	*p* value
Perepu et al. (2010)*	TNF (pg/mg protein)	6 exposed, 6 control	Rat; *ex vivo*†	8 hr/day for 56 days; chamber	Adult	0	16 ± 2.5	NA
						800	**37 ± 2.3**	**< 0.05**
	IL-10 (pg/mg protein)					0	13 ± 1.4	NA
						800	**4.2 ± 0.4**	**< 0.05**
	SOD (U/mg protein)					0	54 ± 4.2	NA
						800	**25 ± 2.5**	**< 0.05**
	MDA (nmol/mg protein)					0	0.4 ± 0.02	NA
						800	**0.9 ± 0.08**	**< 0.05**
Perepu et al. (2012)*	TNF (pg/mg protein)	6 exposed, 6 control	Rat‡	8 hr/day for 56 days; chamber	Adult	0	8.8 ± 0.7	NA
						800	**25 ± 2.4**	**< 0.05**
	IL-10 (pg/mg protein)					0	18 ± 1.7	NA
						800	**11 ± 0.1**	**< 0.05**
	SOD (U/mg protein)					0	79 ± 4.2	NA
						800	**49 ± 1.7**	**< 0.05**
	Lipid peroxidation (nmol/mg protein)					0	4.5 ± 0.5	NA
						800	**10 ± 0.9**	**< 0.05**
Kodavanti et al. (2011)	White blood cells (per mL × 10^6 whole blood)	20 exposed, 20 control	Rat	5 hr/day, 1 day/wk, 16 wks; nose-only	Adult	0	1.65 ± 0.12	NA
						500	1.55 ± 0.07	> 0.05
	Lymphocytes (%)					0	88.05 ± 1.64	NA
						500	**67.45 ± 1.89**	**≤ 0.05**
	Fibrinogen (mg/dL)					0	204.5 ± 1.5	NA
						500	202.1 ± 3.0	> 0.05
	Total cholesterol (mg/dL)					0	97.2 ± 1.8	NA
						500	99.4 ± 1.8	> 0.05
	HDL (mg/dL)					0	27.5 ± 0.8	NA
						500	29.4 ± 0.8	> 0.05
	LDL (mg/dL)					0	11.4 ± 0.4	NA
						500	11.0 ± 0.4	> 0.05
	Triglycerides (mg/dL)					0	46.7 ± 3.3	NA
						500	50.1 ± 4.6	> 0.05
	Glucose (mg/dL)					0	194.7 ± 7.5	NA
						500	183.5 ± 3.4	> 0.05
Sethi et al. (2012)	TNF (pg/mg protein)	6 exposed, 6 control	Rat	8 hr/day for 56 days; chamber	Adult	0	11.8 ± 1.1	NA
						800	26.8 ± 2.1	**< 0.05**
	SOD (U/mg protein)					0	99.6 ± 5.1	NA
						800	55.1 ± 3.4	**< 0.05**
Gordon et al. (2013)	White blood cells	7-10 exposed, 7-10 control	Rat	6 hr/day, 1 day/wk, 17 wks; chamber	Adult (4 mos.) and senescent (20 mos.)	800	NC	> 0.05
	Lymphocytes						NC	> 0.05
	CRP						NC	> 0.05
	FGF-basic						**Decreased**	**0.00053**
	GCP2						**Decreased**	**0.037**
	IL-11						**Decreased**	**0.042**
	MCP-3						NC	> 0.05
	MDC-2						NC	> 0.05
	MIP-1alpha						NC	> 0.05
	MIP-1beta						NC	> 0.05
	MIP2						**Decreased**	**0.038**
	MIP-3beta						NC	> 0.05
	Myoglobin						NC	> 0.05
	RANTES						NC	> 0.05
	VEGF						NC	> 0.05

CRP C-reactive protein, *FGF* fibroblast growth factor, *GCP2* granulocyte chemotactic protein, *HDL* high-density lipoprotein, *IL* interleukin, *LDL* low-density lipoprotein, *MCP* monocyte chemotactic protein, *MDA* malondialdehyde, *MDC* macrophage-derived chemokine, *MIP* macrophage inflammatory protein, *n* number of animals, *NA* not available, *NC* no change, *ppb* parts per billion, *RANTES* regulated on activation, normal T cell expressed and secreted, *SOD* superoxide dismutase, *TNF* tumor necrosis factor, *VEGF* vascular endothelial growth factor

Bolded values are statistically significant.

*Estimated from study figures using GetData Graph Digitizer.

†Rats were continuously exposed to ozone for 8 hr/day for 56 days; after sacrifice, hearts were removed and subjected to 30 min of global ischemia followed by 60 min of reperfusion. Biomarkers were measured after ischemia/reperfusion.

‡Rats were continuously exposed to ozone for 8 hr/day for 56 days; after sacrifice, hearts were removed and biomarkers measured.

Blood pressure. Two Tier II cross-sectional studies evaluated the effects of long-term ozone exposure on blood pressure (Chuang et al. 2011, Dong et al. 2013a). One study in an elderly Taiwanese population (Chuang et al. 2011) reported that an interquartile range (IQR) increment of 8.95 ppb in a 1-year average ozone exposure was associated with increases of 21.51 mmHg (95% CI: 16.90 to 26.13) and 20.56 mmHg (95% CI: 18.14 to 22.97) in systolic and diastolic blood pressure, respectively. This study was conducted in a random sample of participants of a large-scale nation-wide survey; the presence of selection bias was thus unlikely. However, it is likely that ozone and blood pressure measurement error were present, as the reported increases in blood pressure were so large for such a small increase in ozone concentration that they seem implausible. Uncontrolled confounders such as physical activity, occupation, and pre-existing conditions were also likely to have impacted the results. Furthermore, the statistically significant associations between ozone and blood pressure were yielded from single-pollutant models. It is unclear whether the observed increases in blood pressure were attributable to a true causal effect of ozone or were the result of confounding by uncontrolled co-pollutants, such as PM_{10}, $PM_{2.5}$, and NO_2, which showed similarly strong effects on blood pressure in single-pollutant models.

A study in China (Dong et al. 2013a) reported that an increment of 22 μg/m^3 (11.2 ppb) in 3-year average ozone levels was associated with clinically insignificant increases of 0.73 mmHg (95% CI: 0.35 to 1.11) and 0.37 mmHg (95% CI: 0.14 to 0.61) in systolic and diastolic blood pressure, respectively. Stratified analyses by sex showed that the effects of ozone on blood pressure were only present in men, but not in women. This study also reported an OR of 1.13 (95% CI: 1.06 to 1.20) for prevalent hypertension associated with an increment of 22 μg/m^3 in 3-year average ozone levels. Subgroup analyses showed that the effect of ozone was stronger in men (OR = 1.21, 95% CI: 1.04 to 1.38) than in women (OR = 1.07, 95% CI: 0.91 to 1.18). When stratified by age, long-term ozone exposure was positively associated with the prevalence of hypertension for individuals younger than 55 years of age (OR = 1.13, 95% CI: 1.07 to 1.20) or those older than 64 years of age (OR = 1.15, 95% CI: 0.96 to 1.39), but not for individuals between 55 and 64 years of age (OR = 1.02, 95% CI: 0.92 to 1.14). This cross-sectional study restricted the population to those who lived within 1 km of an air monitoring station; thus, the presence of selection bias was possible. Because blood pressure was measured manually and hypertension was defined based on these blood pressure measurements, errors in outcome assessment were possible. In addition, the prevalence of hypertension occurred in greater than a third of this population. The authors inappropriately employed logistic regression for analyses of hypertension prevalence and ozone levels, as it overestimated the relative risks. Moreover, the effect estimates for ozone were not adjusted for co-pollutants. PM_{10} and SO_2 showed similarly strong effects on blood pressure in this population. Age appeared to be an effect modifier for the association between ozone and hypertension; however, there is no biologically plausible mechanism to explain why the effects of ozone would only be present in people younger than 55 or older than 64 years of age, but absent in people between 55 and 64 years of age. In contrast to the study by Chuang et al. (2011), the reported increases in blood pressure in this study are more plausible.

Although Chuang et al. (2011) and Dong et al. (2013a) both reported positive associations between long-term ozone exposure and blood pressure, the increases in blood pressure associated with similar increments in ozone levels differed by an order of magnitude in these two studies. As these are both Tier II studies, it is unclear whether the reported results are reliable.

Carotid artery intima-media thickness. A Tier II cross-sectional study (Breton et al. 2012) investigated whether early life and lifetime ozone exposures affected CIMT, a marker for atherosclerosis. An increase of two standard deviations in early childhood ozone exposure or elementary school ozone exposure was associated with an increase of 7.8 μM (95% CI: − 0.3 to 15.9) or 10.1 μM (95% CI: 1.8 to 18.5) in CIMT, respectively. Adjusting for co-pollutants such as NO_2, PM_{10}, or $PM_{2.5}$ yielded similar results. The effect of lifetime ozone exposure on CIMT was weaker when compared to ozone exposures in early life and not statistically significant. This study recruited students from a single university; thus, the presence of selection bias was likely. CIMT was measured by a single specialist using ultrasound, which likely introduced random and systematic measurement errors in outcome assessment. Early life and lifetime ozone exposure were spatially interpolated using inverse distance-squared weighting (IDW2) models based on a lifetime residential history provided by the participants; thus, substantial errors in exposure measurement and interpolation are possible. Furthermore, although this study met our scoring criteria for an adequate statistical analysis, the authors inappropriately adjusted for LDL-cholesterol and HDL-cholesterol (two intermediate factors in the causal pathway of atherosclerosis) in the statistical regression models, making the results problematic and difficult to interpret.

Table 12. Study quality – experimental animal studies of cardiovascular-related biomarkers.*

	Study design								
Study	Exposure assignment	Compared to appropriate control	Experimental groups	Experimental animals, housing, and husbandry	Exposure environment	Outcome assessment	Attrition bias	Statistical analysis	Total score
Sethi et al. (2012)	− 1	1	− 1	− 1	1	1	− 1	1	0
Perepu et al. (2010)	− 1	1	− 1	1	1	1	− 1	1	2
Perepu et al. (2012)	− 1	1	− 1	1	1	1	− 1	1	2
Gordon et al. (2013)	− 1	1	− 1	1	1	1	1	1	4
Kodavanti et al. (2011)	− 1	1	1	1	1	1	1	1	6

*The study quality scoring system used to determine these scores is described in the text.

Stroke. Three Tier I studies examined the effects of long-term ozone exposure on the occurrence of stroke (Lipsett et al. 2011, Atkinson et al. 2013, Dong et al. 2013b). Two of these were cohort studies that reported null associations between ozone and the incidence of stroke (Lipsett et al. 2011, Atkinson et al. 2013). Because both studies included a large number of participants and stroke cases, the null findings are likely not due to insufficient power. As discussed above, possible selection bias, errors in outcome and exposure measurement, and residual confounding also cannot be completely ruled out. The third study was a cross-sectional study in China (Dong et al. 2013b) that reported a positive, but not statistically significant, association between 3-year average ozone exposure and prevalence of stroke [odds ratio (OR) = 1.15, 95% CI: 0.99 to 1.33]. The study population was restricted to those who lived within 1 km of an air monitoring station; the presence of selection bias was possible in this study. In addition, prevalence of stroke was assessed by self-reported diagnoses, without confirmation by independent chart review. Proxy interviews were conducted for deceased individuals. Recall and information bias were likely to be present in these settings.

Together, the two cohort studies with greater statistical power reported null findings regarding long-term ozone exposure and stroke, while one cross-sectional study reported a non-statistically significant, positive association between the increment in ozone and the prevalence of stroke.

Myocardial infarction. The association between long-term exposure to ozone and the incidence of MI was evaluated in two Tier I studies (Lipsett et al. 2011, Atkinson et al. 2013), and their findings were inconsistent.

The first study is a prospective cohort study, conducted on female teachers in California for over a period of almost 10 years (Lipsett et al. 2011). The authors reported a null association between monthly ozone exposure and development of MI [hazard ratio (HR) = 1.03, 95% confidence interval (CI): 0.95 to 1.11], even after adjustment for $PM_{2.5}$ (HR = 1.06, 95% CI: 0.94 to 1.19). The study population was a well-defined cohort, with minimal loss to follow-up. Both MI hospitalization and mortality were ascertained to capture all incident MI cases. The population was restricted to those who lived within 20 km of an air monitoring station, thus the presence of selection bias was possible. An exhaustive list of potential confounders and risk factors for MI was assessed at baseline and controlled for in the analyses. However, information on these factors was not updated during the follow-up, so residual confounding cannot be ruled out. There were more than 100 000 participants and more than 1300 incident cases of MI; therefore, insufficient power was not likely to account for the lack of statistical significance.

The second study is a retrospective cohort study in Great Britain (Atkinson et al. 2013), that reported that an increment of 3 $\mu g/m^3$ (1.5 ppb) in a yearly average ozone exposure at baseline was associated with a 4% decrease in the risk of developing MI (95% CI: 0.93 to 1.00). This negative association persisted after controlling for co-pollutants such as PM_{10} and NO_2, but it was no longer statistically significant after adjustment for sulfur dioxide (SO_2). This study included more than 800 000 patients from 205 general (family) practices. The selection of the practices was based on whether clinical data were of good quality and linked hospital admission and mortality data were available, making the presence of selection bias possible. Various confounders, including co-pollutants, were controlled for in the analyses. Similar to the study by Lipsett et al. (2011), other potential confounders were only assessed at baseline, and certain potential confounders, such as physical activity, were not considered.

The effect estimates of ozone exposure from these two studies were small and not consistently in one direction (>1 or <1). Possible selection bias, errors in outcome and exposure measurement, and residual confounding further contribute to the uncertainty of the results.

Other cardiovascular disease endpoints. The CV morbidity endpoints of arrhythmia, heart failure, CVD, and IHD were examined in only one Tier I study each. Atkinson et al. (2013) evaluated the effects of long-term ozone exposure on the incidence of arrhythmia. The risk estimates for ozone yielded from single-pollutant models were close to the null value and not statistically significant. Further adjustment for co-pollutants such as PM_{10} and NO_2 in bi-pollutant models did not change the results. In a bi-pollutant model with adjustment for SO_2, an IQR increment of 3 $\mu g/m^3$ (1.5 ppb) in the yearly average ozone level was associated with an OR of 1.03 (95% CI: 1.00 to 1.06). This study employed a retrospective cohort design and included a large number of participants; it was of higher quality among the epidemiological studies we evaluated. However, possible selection bias, measurement errors in outcome and exposure assessment, and residual confounding contribute to the uncertainty of the findings. In addition, because the risk estimates for ozone were generally close to 1 and not statistically significant, they were likely to be chance findings.

In addition to arrhythmia, Atkinson et al. (2013) examined the effects of ozone on the incidence of heart failure. An increment of 3 $\mu g/m^3$ (1.5 ppb) in baseline yearly ozone level was associated with a decreased incidence of heart failure (HR = 0.94, 95% CI: 0.90 to 0.98). Further adjustments for co-pollutants such as PM_{10}, NO_2, and SO_2 yielded similar statistically significant, negative estimates (HR = 0.95, 95% CI: 0.91 to 0.99). The risk estimates, though statistically significant, were very close to 1 and may be due to chance.

A Tier I cross-sectional study in China (Dong et al. 2013b) reported null associations between 3-year average ozone levels and CVD prevalence (OR = 1.09, 95% CI: 0.84 to 1.41). Stratified analyses by sex also yielded null results for both men and women. As discussed above, possible selection bias and measurement errors in outcome and exposure assessment contributed to uncertainty in the results.

A Tier I cross-sectional study in Canada examined the effects of long-term exposure to air pollutants on IHD prevalence among respiratory patients (Beckerman et al. 2012). No association between 10-year average ozone levels and IHD prevalence was observed in this population. The patients included in the study were identified from a single pulmonary clinic in Toronto, Canada, which likely introduced selection bias. Possible errors in exposure measurement and incomplete control for potential confounders further contributed to the uncertainty of the results. Because IHD was broadly defined, outcome misclassification was unlikely even without an independent chart review.

Conclusions. Overall, the Tier I studies reported null or negative associations between long-term ozone exposure and CV morbidity. Studies we judged to be of lower quality (Tier II) yielded more positive associations; however, a number of methodological limitations in these studies severely undermined the validity of the findings. When endpoints such as blood pressure and CIMT were used, there were more positive associations with ozone exposure compared to using disease endpoints such as MI and stroke. It remains unknown whether the clinically modest changes observed in these endpoints, if due to ozone, would persist and manifest into clinically significant diseases or symptoms. Given the largely null, and occasionally conflicting, findings in the studies of ozone and CV morbidity, factors such as biological plausibility, exposure-response, and temporality are less relevant to assess the causal inference from these findings. In addition, the effects of ozone on most CV morbidity endpoints were only evaluated in one or two studies, and the findings have not been replicated and confirmed in different populations. We conclude that the current body of epidemiology evidence is limited, but the largely null findings in the Tier I studies do not support a causal relationship between long-term exposure to ozone and CV morbidity.

Cardiovascular mortality

We classified all 12 epidemiology studies of CV mortality as Tier I studies. Table 5 presents the results of these studies. In the following sections, we summarize the results of these studies based on study design and consider the strength of association, internal consistency, temporality, and outcome assessment.

Cohort studies. Ten of the Tier I studies we identified are cohort studies (Abbey et al. 1999, Pope et al. 2002, Chen et al. 2005, Jerrett et al. 2005, 2009, Krewski et al. 2009, Smith et al. 2009, Lipsett et al. 2011, Spencer-Hwang et al. 2011, Carey et al. 2013). The results of these studies are described below.

Adventist Health and Smog Study Cohort. Two studies investigated CV mortality in the Adventist Health and Smog Study (AHSMOG) cohort of 6338 non-smoking Seventh-day Adventists living in the areas of San Francisco, Los Angeles, and San Diego (Abbey et al. 1999, Chen et al. 2005). This cohort had a low loss to follow-up with respect to vital status ($<0.01\%$). Both sets of investigators estimated exposure by interpolating data from 348 fixed-site monitors to zip code centroids according to home and work location histories of study participants, but exposure measurement error was still possible. Both studies performed a sensitivity analysis testing the robustness of the proportional hazards model. The results of these analyses indicated that the assumptions of proportionality were met, increasing confidence in the statistical analyses of both the studies. Both studies restricted analysis to zip code centroids within 50 km of a monitoring station, so selection bias was possible. Each study also only adjusted for a limited set of confounders.

The two studies reported similar results. Abbey et al. (1999) evaluated mortality from all cardiopulmonary causes and reported that there was a slight, but non-statistically significant increase in mortality in men and a decrease in mortality in women. These results were consistent across the main analysis and the analysis restricted to ozone concentrations >100 ppb. Chen et al. (2005) reported decreased, but not statistically significant, risks of mortality from CHD in analyses stratified by sex.

American Cancer Society Cohort. Several studies examined the ozone-mortality associations in the American Cancer Society (ACS) cohort, which consists of about 500 000 adults in 134 metropolitan areas in the US (Pope et al. 2002, Jerrett et al. 2005, 2009, Krewski et al. 2009, Smith et al. 2009). In each of these studies, the authors evaluated 20 variables with 44 terms for individual characteristics that could confound or modify the associations between pollution and death. These factors included median household income, diet, and physical activity, air conditioning in the home, outdoor temperature, occupational exposures, BMI, and education. Ascertainment of deaths in this cohort was 98% complete through 1988 and 93% complete for 1988, and the ACS verified all deaths using death certificates. All of the ACS cohort studies used appropriate statistical methods (i.e. Cox proportional hazards modeling). Although the ACS cohort includes data for more than 1.2 million people across the US, all but one (Jerrett et al. 2005) of the available ACS studies restricted study participants to metropolitan areas where pollution monitors were available; thus, selection bias was possible in these studies. By contrast, Jerrett et al. (2005) interpolated available monitoring data across the entire study area (the LA region), rather than excluding areas with fewer monitors. While all of the following studies used data from the ACS cohort and were of a similar overall quality, there were some methodological differences with regard to selection criteria, time period of the study, outcome assessment, and exposure measurement across studies.

Pope et al. (2002) evaluated all cardiopulmonary deaths in the cohort. The authors conducted several in-depth sensitivity analyses to account for spatial trends and alternative pollution indices and reported that the results remained stable in these alternative models, which increases confidence in the overall results of the study. Unlike some of the other ACS studies, the authors only used single-pollutant models and reported slight, non-statistically significant increases in cardiopulmonary mortality [relative risk (RR) = 1.10 and 1.11, respectively, for all year and seasonal analyses]. The association between ozone and all-cause mortality, which is not as subject to outcome misclassification error, was null in this study. Results for all-cause mortality should be somewhat consistent with those for CV mortality, as they are in this study, because often, the ultimate cause of death is recorded as being CV-related (i.e. the heart stops beating when a patient is near death), even though this may not be the underlying cause of death.

Jerrett et al. (2005) assessed mortality from cardiopulmonary conditions and IHD based on long-term maximum 8-h ozone and peak daily ozone concentrations. This study differs from the other ACS cohort studies because selection bias was unlikely; the analysis focused on Los Angeles, with no exclusion criteria based on residence near an ozone monitor. Exposure misclassification was possible, but may have been minimized because the authors employed interpolation of monitoring data. This allowed exposures to be estimated on

a smaller scale compared with studies that estimated exposure using data from monitors far away from the participants' homes. An important limitation of this study is that the mortality data were collected before the ozone data; thus, the ozone data used in the analyses were unlikely to reflect the exposures incurred at the time of death. The risks for IHD or cardiopulmonary mortality were all less than null, and the decrease was statistically significant (RR = 0.97, 95% CI: 0.94 to 0.99) in the analysis of peak ozone exposure and cardiopulmonary mortality. Similar findings were reported for all-cause mortality.

Jerrett et al. (2009) evaluated CV mortality risks in single- and bi-pollutant models with $PM_{2.5}$ and limited their analyses to the summer season. The authors analyzed cardiopulmonary, all CV, and IHD mortality separately. Unlike the study by Jerrett et al. (2005), the ozone and mortality data analyzed in this study were from similar time periods, establishing temporality between exposure and outcome. Exposure measurement error was likely, however, due to the area-level ozone data. The authors reported that there was a slight increase in risk (RRs approximately 1.01) in all three categories of mortality in the single-pollutant models; however, the associations were negative (and mainly statistically significant) in the bi-pollutant models with $PM_{2.5}$. In addition, all-cause mortality was not associated with ozone exposure. Overall, the authors were unable to detect any effect of long-term ozone exposure on CV mortality in this study of over 448 000 individuals spanning 18 years.

Krewski et al. (2009) also analyzed cardiopulmonary and IHD mortality in the ACS cohort. Although exposure measurement error was likely due to the use of area-level ozone data, the strengths of this study were the adequate statistical analyses with multi-pollutant models, adequate control of confounding, and the availability of sensitivity analyses. A small but statistically significant association between ozone and cardiopulmonary mortality was observed in the summer-only analysis in a single-pollutant model, but not in the all-year analysis and not for IHD mortality. As in the study by Jerrett et al. (2009), the results were null (RRs < 1) when the all-year analyses were adjusted for $PM_{2.5}$. All-cause mortality was also associated with ozone exposure in the summer-only analyses in single-pollutant models.

The study by Smith et al. (2009) evaluated all cardiopulmonary mortality in the ACS cohort. The most significant limitation of this study was that mortality and exposure data were not temporally related; the ozone data were collected after the mortality data (2003–2005 and 1982–2000, respectively). It is likely that ozone levels were lower in the metropolitan areas of the ACS studies in 2003–2005 compared to the 1980s and 1990s; thus, the risk estimates for mortality may be biased toward the null in this study. In addition, unlike all of the other ACS studies, no sensitivity analyses were conducted. Smith et al. (2009) reported a small but statistically significant increase in cardiopulmonary mortality in the warm season in single- and multi-pollutant models with carbon and sulfate. The results were not statistically significant in the bi-pollutant model with carbon only. In addition, all-cause mortality was not associated with ozone exposure in single or multi-pollutant models. The lack of temporality reduces confidence in the results of this study.

Other Cohorts. Lipsett et al. (2011) investigated the association between ozone and deaths from all CV conditions, IHD, and cerebrovascular conditions in the California Teacher's cohort (women only). The strengths of this study were a high follow-up rate (virtually 100%), unlikely selection bias, adequate statistical methods with bi-pollutant models, and adequate control of confounding, including adjustment for multiple CV health-related lifestyle factors. Exposure measurement error was possible, but not likely, as monthly individual exposure estimates were created via spatial linkage of the geocoded residential addresses to nearby monitoring stations. The majority of results, in all-year and seasonal analyses with and without control for the effects of $PM_{2.5,}$ were negative or null. The one exception was a statistically significant increase in IHD mortality in the summer-only analysis (RR = 1.09, 95% CI: 1.01 to 1.19). For all-cause mortality, the authors reported no association in the summer-only analysis.

Spencer-Hwang et al. (2011) evaluated the risk of death from CHD in non-smoking patients who underwent kidney transplants and lived within 50 km of an air pollutant monitor. This study was considered to have inadequate control of confounders; although the authors controlled for medication use and various transplant-related variables, they failed to adjust for physical activity, diet, alcohol use, BMI, and other important lifestyle factors that may have affected CHD risk. Selection bias was possible due to the inclusion criteria. The strengths of the study include adequate statistical analyses with bi-pollutant models and inclusion of sensitivity analyses. The authors reported statistically significant associations for a 10 ppb increase in ozone and CHD in single- and bi-pollutant models that included PM_{10}. All-cause mortality associations were not significantly increased, however, in both single- and multi-pollutant analyses.

The most recent cohort study by Carey et al. (2013) analyzed a large population in England. Exposure measurement error was possible, but the authors reduced the potential for this error by using air dispersion models of ozone and other pollutants. The dispersion models incorporated available monitoring data as well as modeled concentrations and considered the effects of physical and chemical processes in the atmosphere. This allowed ozone to be estimated for each 1 km^2, which should improve the accuracy of exposure measurements. The authors also performed numerous sensitivity analyses, including one that tested the assumption that circulatory deaths were under-reported on death certificates (a possible issue reported in the literature). The authors reported that disease misclassification was unlikely, which increases confidence in the results of this study. The participants were recruited from a list of patients registered with a general practitioner, so selection bias was possible. In addition, there was incomplete control of confounders and the authors did not include co-pollutants in their analyses. The results of this study indicated that there was a statistically significant decrease in deaths from all CV conditions (HR = 0.96, 95% CI: 0.94 to 0.98) with every 3.0 $\mu g/m^3$ (1.5 ppb) increase in ozone.

Ecological studies. Based on our scoring system, two cross-sectional, ecological studies were ranked as Tier I studies (Janke et al. 2009, Wang et al. 2009). Although these stud-

ies are considered less robust by design because they cannot address causation, they can be potentially useful for identifying areas requiring further study.

Janke et al. (2009) collected mortality data from all circulatory disease, CHD, and MI for 354 local authorities in the UK between 1998 and 2005. The authors conducted analyses with several co-pollutants and collected ozone and mortality data during the same period. However, the authors obtained mortality data from a registry and did not verify the cause of death using death certificates, so outcome misclassification was likely. In addition, the authors modeled exposure levels, conducted sensitivity analyses to validate the exposure modeling, and adequately controlled for potential confounders. This study reported no association between ozone exposure and mortality from any category of CV disease.

Wang et al. (2009) investigated the association between long-term exposures to gaseous air pollutants (NO_2, ozone, SO_2) and cardiorespiratory mortality in Brisbane, Australia. Although exposure measurement error was possible in this study, the authors used geographical information system techniques to map the spatial patterns of the gas concentrations and to assign exposure estimates at the statistical local area level, which allowed for more accurate exposure assessment. The study suffered from numerous limitations, however, including a lack of verification of deaths using death certificates, which likely resulted in outcome misclassification; a lack of sensitivity analyses; and incomplete adjustment for confounders. The authors reported null associations with mortality in the single-pollutant analysis (RR = 1.002, 95% CI: 0.99 to 1.02) and after adjustment for NO_2 and SO_2 (RR = 1.00, 95% CI: 0.99 to 1.01).

Conclusions. Of the 10 Tier I cohort studies, five reported statistically significant increases in CV mortality in at least one analysis (Jerrett et al. 2009, Smith et al. 2009, Krewski et al. 2009, Lipsett et al. 2011, Spencer-Hwang et al. 2011, Carey et al. 2013). Each of these studies reported point estimates close to unity (i.e. 1), indicating that the ozone-mortality association was weak. Furthermore, in all but one of these studies, the results were not in the same direction (>1 or <1) across analyses; i.e. the authors reported both increases and decreases in mortality associated with ozone. The study by Spencer-Hwang et al. (2011) was the only one of these five studies in which the ozone-CV mortality association was consistent and remained statistically significant after the inclusion of co-pollutants. The other five Tier I cohort studies reported null or negative results. In particular, the ACS cohort studies represent some of the largest and most robust epidemiological studies of air pollution health effects. The fact that these studies found primarily null effects for CV-related and all-cause mortality, particularly in multi-pollutant analyses, calls the likelihood of a causal relationship between ozone and mortality into question. In addition, while limited in their usefulness for causation analysis, both Tier I ecological studies reported no change in CV mortality risk with ozone exposure. Although much uncertainty exists in the ozone-CV mortality studies because of possible exposure measurement error, confounding, and other issues, the totality of the evidence does not support a causal relationship between long-term ozone exposure and CV mortality.

Cardiovascular biomarkers

We classified all three epidemiology studies of long-term ozone exposure and CV biomarkers as Tier I. Table 7 presents the results of these studies.

Two epidemiology studies examined biomarkers associated with inflammation. These were proinflammatory biomarkers that should increase in concentration after exposure if ozone induces systemic inflammation. In the study by Chuang et al. (2011), levels of IL-6 were unchanged and neutrophil counts were statistically significantly increased with each IQR increase in yearly average ozone concentration in single-pollutant models. In the study by Forbes et al. (2009), levels of C-reactive protein (CRP) in non-fasting blood samples were not affected by ozone, as they were slightly decreased, but not statistically significantly, with each 1 $\mu g/m^3$ (0.5 ppb) increase in annual average ozone exposure.

One epidemiology study (Forbes et al. 2009) examined a biomarker of coagulation. This prothrombotic marker, fibrinogen, would be expected to increase if ozone produces an adverse effect on the CV system, as it is a marker for cardiovascular disease. Fibrinogen is also an acute phase protein, so it is used as a marker of inflammation. In the study by Forbes et al. (2009), there were no statistically significant changes in levels of fibrinogen in non-fasting blood samples associated with each 0.5 ppb increase in the annual average ozone exposure.

One epidemiology study examined oxidative stress biomarkers. Chen et al. (2007) reported that the oxidative stress marker 8-iso-prostaglandins-$F_{2\alpha}$ (8-iso-PGF) was increased in healthy young adults with increasing lifetime exposure to ozone. In the same study, the antioxidant capacity marker FRAP (ferric reducing ability of plasma) was reduced, but this result was not statistically significant. The authors noted that inclusion of $PM_{10-2.5}$ and NO_2 in the models did not alter the associations with ozone.

One epidemiology study examined biomarkers of lipid and glucose metabolism. Chuang et al. (2011) reported no statistically significant effects of ozone on HDL cholesterol or triglycerides, whereas levels of total cholesterol, fasting glucose, and hemoglobin A1c (HbA1c) increased with each IQR increase in the yearly average ozone concentration in single-pollutant models. HbA1c is used to monitor the degree of control over glucose metabolism, and increases in HbA1c are associated with the risk of developing arterial plaques (Chuang et al. 2010).

Together, the studies of changes in levels of CV biomarkers after long-term exposure to ozone are limited, in that each biomarker was only assessed in one study. Among groups of biomarkers (e.g. inflammation, oxidative stress, lipid and glucose metabolism), there were no consistent effects after exposure, as ozone did not induce statistically significant changes in most biomarkers. There is little biological plausibility for the few statistically significant effects, as there are many other factors besides ozone that contribute to changes in biomarker levels, and it is unclear whether the majority of the biomarkers are clinically relevant.

Experimental animal studies

Cardiovascular morbidity

We classified four of the experimental animal studies of morbidity endpoints as Tier I and one as Tier II (Sethi et al. 2012). Table 9 presents the results of these studies. Below, we briefly describe the implications and limitations of these studies, considering the issues of internal consistency as well as the clinical significance and relevance of the results for each of the individual studies.

Only one experimental animal study investigated the effects of long-term ozone exposure on heart rate and blood pressure (Gordon et al. 2013). The authors exposed 4- and 20-month-old (senescent) rats intermittently (1 day/week) to 800 ppb ozone over 15 weeks. They reported no changes in heart rate measured every other week, 1 day after ozone exposure, compared to control rats exposed to filtered air. The authors did not present the data, but they also reported that intermittent ozone exposure over 15 weeks had no effect on systolic and diastolic blood pressure. As the measurements for these endpoints were taken 1 day after the exposure, with a recovery period of 1 week between exposures, the results may not be relevant to the assessment of the long-term effects of ozone.

The remaining four studies investigated the effect of ozone exposure on other measures of cardiac structure and function. Chuang et al. (2009) measured red-O staining of the aorta, which is indicative of atherosclerotic lesions, in groups of male mice. The apoE −/− mice used in this study were sensitive to atherosclerotic lesions, unlike normal C57B16 mice. The authors reported that 8 weeks of exposure to 500 ppb ozone resulted in a statistically significant increase in the percent of aortic surface staining (0.7% in the controls *versus* 1.50% in exposed mice).

Sethi et al. (2012), Perepu et al. (2010), and Perepu et al. (2012) each investigated the effect of ozone exposure on left ventricular function in male rats, as measured by left ventricular pressure and the rate of ventricular pressure development. All three studies reported that exposure to 800 ppb ozone for 56 days resulted in statistically significant decreases in all measures of left ventricular function either *in vivo* (Perepu et al. 2012, Sethi et al. 2012) or in isolated rat hearts subjected to ischemia and reperfusion (Perepu et al. 2010). We considered the *ex vivo* study by Perepu et al. (2010) to be less relevant than the *in vivo* studies. The study by Sethi et al. (2012) is the only experimental animal study of CV morbidity that we classified as Tier II, and it did not provide information on animal husbandry/housing and outcome assessment as required by our study scoring criteria. The invasive methods used for measuring ventricular function in the *in vivo* studies (Perepu et al. 2012, Sethi et al. 2012) may have affected the results. These methods have disadvantages, such as the potential for aortic damage (Eskesen et al. 2012) and the need for anesthesia. In both studies, the rats were anesthetized with ketamine and xylazine, which have been shown to decrease left ventricular function in rats and humans (Droogmans et al. 2008, Jakobsen et al. 2010). Furthermore, it is unclear whether the results of either study represent clinically significant changes, as abnormalities in the mechanical function of the ventricle can occur in the presence or absence of the clinical syndrome of heart failure (Zile and Brutsaert 2002).

Overall, each CV morbidity endpoint was examined in only one or a few experimental animal studies. The Tier I study of heart rate and blood pressure reported no effects of ozone exposure, but because of the exposure regimen, the results may not be relevant to long-term exposures. Another Tier I study reported a small increase in aortic surface staining for atherosclerotic lesions in mice that are sensitive to such lesions. Two Tier I studies and one Tier II study from the same group of investigators reported decreases in ventricular function after ozone exposure; however, one of the Tier I studies was an *ex vivo* study that is less relevant to humans, and both *in vivo* studies used invasive methods and a type of anesthesia that may have affected the results. Together, the evidence from experimental animal studies of CV morbidity is limited but may indicate effects on atherosclerotic lesions and ventricular function that should be confirmed by other investigators, before they can be considered to support a causal relationship between long-term ozone exposure and CV morbidity.

Cardiovascular biomarkers

We classified four experimental animal studies of long-term exposure to ozone and CV biomarkers as Tier I and one as Tier II. The results of these studies are presented in Table 11.

Biomarkers of inflammation. All five rat studies examined biomarkers associated with inflammation. The majority were proinflammatory biomarkers that should increase during inflammation, but one study examined an anti-inflammatory marker, interleukin (IL)-10, that would be expected to decrease during inflammation. One study assessed the effects of ozone exposure on the levels of CRP. Gordon et al. (2013) reported a slight increase in CRP levels in rats exposed to 800 ppb ozone for 6 hours per week over 17 weeks, but this result was not statistically significant. White blood cell (WBC) and lymphocyte counts were examined in two studies, with mixed results. WBC counts were non-statistically significantly decreased in the study by Kodavanti et al. (2011) and non-statistically significantly increased in the study by Gordon et al. (2013). Lymphocyte counts were decreased ($p < 0.05$) in the study by Kodavanti et al. (2011) but were slightly, though non-statistically significantly, increased in the study by Gordon et al. (2013).

The only other biomarkers of inflammation examined in more than one study were the proinflammatory marker tumor necrosis factor (TNF) and the anti-inflammatory marker IL-10. These biomarkers were studied in the heart tissue of rats exposed to 800 ppb ozone for 8 hours per day over 56 days by the same group of investigators (Perepu et al. 2010, 2012, Sethi et al. 2012). Levels of TNF were increased in the hearts of exposed rats that were subjected to ischemic injury and reperfusion in an *ex vivo* study (Perepu et al. 2010) as well as in uninjured hearts from *in vivo* studies (Perepu et al. 2012, Sethi et al. 2012). The levels of IL-10 were decreased after exposure in the *ex vivo* study (Perepu et al. 2010) as well as in one of the *in vivo* studies (Perepu et al. 2012). The clinical relevance of the findings from the *ex vivo* study is unclear.

Gordon et al. (2013) measured a large number of serum biomarkers in rats exposed to 800 ppb ozone for 6 hours per week over 17 weeks. Of the inflammatory biomarkers measured, three were statistically significantly decreased

[granulocyte chemotactic protein-2 (GCP-2), IL-11, and macrophage inflammatory protein (MIP2)], whereas all others were unchanged. The three biomarkers with decreased levels were proinflammatory markers that should increase if ozone induces systemic inflammation. However, the biomarkers measured in this study were general cytokines and chemokines that are not necessarily associated with risk of CVD.

Overall, the studies of inflammatory biomarkers do not indicate any consistent changes associated with ozone exposure. The majority of biomarkers examined were either not affected by ozone, or the reported effects were in the opposite direction for an adverse effect, as the levels of many of the proinflammatory biomarkers were decreased, rather than increased, after ozone exposure. The only consistent results were in the studies conducted in rat hearts by the same group of investigators, in which TNF levels were increased and IL-10 levels were decreased. We considered the *ex vivo* study by Perepu et al. (2010) to be less relevant than other studies, and the results are not consistent with the majority of the other studies reviewed here that indicate no adverse effects of long-term ozone exposure on inflammation.

Biomarkers of oxidative stress. Three rat studies (Perepu et al. 2010, 2012, Sethi et al. 2012) examined oxidative stress biomarkers. Perepu et al. (2010, 2012) reported an increase in malondialdehyde (MDA), an indicator of lipid peroxidation, in ischemic-injured and uninjured heart tissue of rats exposed to 800 ppb ozone for 8 hours per day over 56 days. Under the same experimental conditions, these authors, as well as Sethi et al. (2012), also reported a decrease in the antioxidant enzyme superoxide dismutase (SOD) in ischemic-injured and uninjured rat heart tissue. The *ex vivo* study by Perepu et al. (2010) is less relevant than the *in vivo* studies (Perepu et al. 2012, Sethi et al. 2012).

Biomarkers of coagulation. Only one experimental animal study examined a biomarker of coagulation. Kodavanti et al. (2011) reported a slight decrease in fibrinogen levels in rats exposed to 500 ppb ozone for 5 hours per week over 16 weeks, but these results were not statistically significant. As noted above, the fibrinogen levels would be expected to increase if ozone produces an adverse effect on the CV system.

Biomarkers of lipid and glucose metabolism. One experimental animal study examined biomarkers of lipid and glucose metabolism. There were no changes in the levels of total cholesterol, HDL or LDL cholesterol, triglycerides, or glucose after exposure to 500 ppb ozone for 5 hours per week over 16 weeks (Kodavanti et al. 2011).

Biomarkers of overall cardiovascular health. Biomarkers of overall CV health were examined in the experimental animal study by Gordon et al. (2013). There were no changes in the serum levels of macrophage-derived chemokine, myoglobin, or the angiogenesis marker vascular endothelial growth factor (VEGF) in rats exposed to 800 ppb ozone for 6 hours per week over 17 weeks, whereas the levels of fibroblast growth factor (FGF) basic, which is implicated in the pathogenesis of arteriosclerosis, were decreased. This decrease was in the opposite direction of an adverse effect for this biomarker.

Conclusions. Overall, there is no consistent evidence for the effects of long-term ozone exposure on the levels of biomarkers of inflammation, coagulation, oxidative stress, lipid and glucose metabolism, or overall CV health from experimental animal studies. The few statistically significant associations were mainly reported in studies of rat hearts after exposure to high levels of ozone and, thus, may not be relevant to humans. In addition, when more than one study examined the same biomarker, there was not always consistency in the direction of the reported effects (regardless of statistical significance), and many were not consistent with adverse effects on the CV system. The reported effects were often very small changes in biomarker concentrations that may be indicative of homeostatic processes (Goodman et al. 2010). Exposure-response relationships for the few reported effects are difficult to discern, as the experimental animal studies each examined only one high concentration of ozone.

Evaluation of evidence across realms

Below, we integrate and evaluate the data across all realms of evidence regarding long-term ozone exposure and CV morbidity and mortality, so that the interpretation of each realm of evidence informs the interpretation of the others. For this evaluation, we considered several aspects to aid in our judgments regarding the WoE. These include the Bradford Hill criteria (as they are commonly referred to) of strength of association, consistency of associations, coherence, biological plausibility, biological gradient (exposure-response), temporality, specificity, and experimental evidence (Hill 1965). The Bradford Hill criteria are not meant to be specific rules to follow and, although they were developed mainly for the interpretation of epidemiological results, we modified them for use in evaluating studies from different realms. In addition to the Bradford Hill criteria, we also considered confounding and bias among the studies, the adversity of reported effects, potential mechanisms, and alternative explanations of the evidence.

For each aspect of the evaluation below, we considered both study quality and relevance. Regarding study quality, although we classified the majority of studies as Tier I, these studies all suffer from some methodological limitations that can affect the interpretation of their results. Despite this, we judged that they are of higher quality, and therefore, are likely more reliable for supporting decisions regarding causation than the Tier II studies. Because of this, we assigned more weight to the Tier I studies (i.e. we relied on them to a greater extent in forming conclusions regarding causality). Regarding study relevance, we considered whether individual study results were relevant to humans at ambient ozone exposures.

Strength of association

When studies report risks that are large and precise, it increases the confidence that an association is causal and not likely attributable to chance, bias, error, or other factors. Although there is no "bright line" above which risks can be considered strong, generally risk estimates indicating less than a two-fold change are considered weak (Taubes 1995).

In the few studies that reported statistically significant effects on CV morbidity outcomes, the magnitudes of the effects were generally quite small (i.e. well below two-fold).

For example, in all Tier I studies, both the positive and negative effect estimates were close to the null value; 1.21 was the highest reported positive risk estimate and 0.94 was the lowest reported negative risk estimate. In addition, the reported effects on biomarker concentrations among the epidemiological and experimental animal studies were very small and may be indicative of homeostatic processes.

The reported associations between ozone and CV-related mortality were also very small, with the magnitudes of effect ranging from 1.014–1.35 for the statistically significant positive associations and 0.96–0.983 for the statistically significant negative associations. These estimates are very weak and within the range of magnitude reported to have a high likelihood of being attributable to confounding (Boffetta et al. 2008).

In several of the studies of morbidity and mortality, the effects were in the opposite direction of adversity and often of similar magnitude as those findings that indicate adverse effects. Because it is unlikely that ozone causes beneficial effects in some studies and harmful effects in others, this indicates that even the statistically significant positive associations may not be indicative of causation.

Overall, because the vast majority of positive results for effects of long-term ozone exposure on CV outcomes were small in magnitude, and some were even below unity, they do not support a causal relationship.

Consistency of associations

The strength of an inference of causality is greater when there is a consistent pattern of effects observed across several independent studies. Because the vast majority of CV morbidity endpoints were examined in only one or two studies each, the consistency of each specific endpoint cannot be discerned with any confidence. Instead, one can determine the consistency of CV morbidity effects as a whole across the epidemiology or experimental animal studies. In contrast, the higher number of studies of CV mortality allows for the determination of consistency across studies of associations for this endpoint.

The Tier I studies of CV morbidity examined the disease endpoints of MI, stroke, arrhythmia, CVD, IHD, and heart failure. Of these endpoints, only MI and stroke were examined in more than one study, and the results of these studies consistently indicate no effects of long-term ozone exposure on either endpoint. In the two studies that examined MI incidence, one reported null results in both single- and bi-pollutant models with $PM_{2.5}$ (Lipsett et al. 2011), whereas the other reported negative associations that were statistically significant in single-pollutant as well as bi-pollutant models with PM_{10} and NO_2 (Atkinson et al. 2013). The three studies that examined the occurrence of stroke all reported null results (Lipsett et al. 2011, Atkinson et al. 2013, Dong et al. 2013b). The studies that examined CVD (Dong et al. 2013b) and IHD (Beckerman et al. 2012) also reported no associations between long-term ozone exposure and these endpoints, whereas the study by Atkinson et al. (2013) reported a decreased incidence of heart failure in single- and bi-pollutant models with PM_{10}, NO_2, and SO_2. For the endpoint of arrhythmia, Atkinson et al. (2013) reported no effects of ozone in single-pollutant or bi-pollutant models with PM_{10}, NO_2, and SO_2. As a whole, the Tier I studies reported only null or negative associations between long-term ozone exposure and CV morbidity.

The Tier II studies of CV morbidity all examined the surrogate endpoints of blood pressure and CIMT. Dong et al. (2013a) reported an increased prevalence of hypertension that was higher in men and individuals less than 55 years of age and very small increases in systolic and diastolic blood pressure in the entire cohort as well as in men, but not women. Chuang et al. (2011) reported increases in systolic and diastolic blood pressure in an elderly population that were more than an order of magnitude higher than the increases reported by Dong et al. (2013a). Only one study examined changes in CIMT, reporting increases in this endpoint with early childhood and elementary school exposure to ozone in most of the bi-pollutant models and no changes in CIMT with lifetime average exposure (Breton et al. 2012). We consider these Tier II studies to be of lower quality than the Tier I studies, and it is possible that the study limitations led to their positive results.

The Tier I cohort studies of CV-related mortality reported inconsistent results: three studies reported no changes in CV mortality across all analyses, two studies reported statistically significant decreases in at least one analysis, and five studies reported small but statistically significant increases in CV mortality in at least one analysis. In three of the five studies that reported increased CV mortality, the positive associations were reported in single-pollutant models and all were null in bi-pollutant models (Jerrett et al. 2009, Krewski et al. 2009, Lipsett et al. 2011). The other two studies reported statistically significant associations in bi- or multi-pollutant models (Smith et al. 2009, Spencer-Hwang et al. 2011). None of these five cohort studies reported statistically significant associations between ozone and all-cause mortality except for the study by Krewski et al. (2009). The two ecological Tier I studies reported no effects of ozone on CV mortality (Janke et al. 2009, Wang et al. 2009).

Studies of CV morbidity endpoints in experimental animals reported no statistically significant effects of long-term ozone exposure on heart rate and blood pressure in one study (Gordon et al. 2013), and an increase in aortic surface staining for atherosclerosis in another study (Chuang et al. 2009). Two Tier I studies and one Tier II study from the same group of investigators examined left ventricular function, and all three reported statistically significant decreases in all measures of this endpoint in *in vivo* (Perepu et al. 2012, Sethi et al. 2012) and *ex vivo* (Perepu et al. 2010) study designs.

Among the three Tier I studies of CV-related biomarkers, none of the biomarkers were examined in more than one study. Further, there were no consistent effects on biomarkers in the same biological pathways. For example, although neutrophils were increased in the study by Chuang et al. (2011), other markers of inflammation (IL-6 and CRP) were not changed (Chuang et al. 2011, Forbes et al. 2009). In addition, the oxidative stress marker 8-iso-PGF was increased, but the antioxidant capacity marker FRAP was not altered in the same study (Chen et al. 2007). Further, for biomarkers of lipid and glucose metabolism, Chuang et al. (2011) reported increases in total cholesterol, fasting glucose, and HbA1c, but no effects of ozone on levels of HDL cholesterol or triglycerides. The animal studies of biomarkers also reported inconsistent results across specific biomarkers and biological pathways,

with statistically significant associations for some markers of inflammation and oxidative stress mainly reported in studies of rat heart tissue from the same group of investigators (Perepu et al. 2010, 2012, Sethi et al. 2012).

Overall, there are no consistent associations between long-term exposure to ozone and CV endpoints across studies in each realm, with the exception of a consistent decrease in left ventricular function and similar changes in biomarkers of inflammation and oxidative stress, in three rat studies conducted by the same group of investigators. There was also a consistent lack of associations across all Tier I studies of CV morbidity endpoints. If ozone were a causal factor for CV effects, one would expect to see consistent associations across studies in each realm, subject to the degree of variability that would be expected based on each study's design and statistical power.

Coherence

An inference of causality from one realm of evidence is stronger when other lines of evidence support a causal interpretation of the association. As CV morbidity was examined in both epidemiology and experimental animal studies, we evaluate coherence across species for morbidity endpoints. Blood pressure was examined in two Tier II studies (Chuang et al. 2011, Dong et al. 2013a) and one Tier I study in rats (Gordon et al. 2013). While the two studies reported increases in blood pressure associated with long-term ozone exposure, there were no effects of ozone on this endpoint in rats exposed to 800 ppb ozone. Because we consider the two studies to be of lower quality (Tier II), it is unclear whether the reported results on blood pressure are reliable; the limitations of these studies may have led to the positive results. Among the other morbidity endpoints studied, the null or negative associations with specific disease endpoints reported in all Tier I studies are not coherent with the reported effects on cardiac function in the *in vivo* and *ex vivo* studies in rats exposed to much higher concentrations of ozone (Perepu et al. 2010, 2012, Sethi et al. 2012).

There is no consistent evidence for effects among the epidemiology and experimental animal studies of ozone and CV-related biomarkers. Effects on biomarkers of inflammation were largely null across species, with one Tier I epidemiological study indicating an increase in neutrophil counts (Chuang et al. 2011) and two Tier I studies in rats indicating effects in the opposite direction of adversity: a decrease in lymphocyte counts (Kodavanti et al. 2011) and decreased levels of the proinflammatory markers GCP-2, IL-11, and MIP2 (Gordon et al. 2013). There were no changes in the levels of fibrinogen, a biomarker of coagulation also used as a marker of inflammation, in a Tier I epidemiology study and a Tier I experimental animal study. Data regarding biomarkers of oxidative stress were limited, although one Tier I epidemiological study indicated an increase in the levels of 8-iso-PGF, but not FRAP; three studies in rats (two Tier I and one Tier II) indicated decreased levels of the antioxidant enzyme SOD in heart tissue and two Tier I studies reported increased levels of the lipid peroxidation indicator MDA in rat heart tissue. For biomarkers of lipid and glucose metabolism, the results were not consistent across species. Total cholesterol and glucose levels were increased with long-term ozone exposure in the epidemiology study by Chuang et al. (2011), but these changes were not observed in rats; in fact, glucose levels were decreased, although not statistically significantly, in the rat study by Kodavanti et al. (2011). The results for HDL and LDL cholesterol in rats also do not support effects of ozone on lipid and glucose metabolism, as there were small, but not statistically significant, changes in these markers that were in the opposite direction of an adverse effect (Kodavanti et al. 2011).

Overall, there is no coherence among the effects reported in the epidemiology and experimental animal studies. However, the differences in exposure parameters and varying methods of exposure and outcome measurement among these studies may not provide a basis for expecting coherence between human and experimental animal results.

Biological plausibility

An inference of causality is strengthened if data are available that demonstrate a biologically plausible mode of action (MoA) for the observed effects. There is currently no known MoA by which long-term exposure to ambient ozone could cause adverse effects on the CV system. The lack of evidence for effects of long-term exposure to ozone on various CV morbidity endpoints leaves few (if any) potential mechanisms by which ozone could contribute to CV disease or death. Ozone is an irritant gas that reacts on contact with respiratory tract lining fluids and is not transported to extrapulmonary sites. One proposed MoA is that ozone indirectly alters CV function by the generation of secondary oxidized lipoproteins in the respiratory tract that can enter the systemic circulation (Barath et al. 2013). Although the data are limited, the epidemiology and experimental animal studies reviewed here indicate that long-term ozone exposure does not significantly alter biomarkers of inflammation or oxidative stress. Of the few studies that reported statistically significant effects on biomarkers, the changes were of small magnitude, and in animal studies they were either only reported in studies of rat heart tissue from the same group of investigators, including two *in vivo* studies (one Tier I and one Tier II) and one Tier I *ex vivo* study that may not be relevant to humans, or were in the opposite direction of an adverse effect. Thus, these effects likely do not represent biologically plausible mechanisms for the effects of long-term ozone exposure because there are many other factors that can contribute to small changes in the levels of these biomarkers. In addition, there is no evidence that small changes in biomarker levels at the high exposures used in the rat studies indicate that the same changes would occur at lower exposures and be indicative of adverse effects.

Another proposed MoA is that of changes in cardiac autonomic control (Barath et al. 2013). If this occurs, one would expect to see decreases in heart rate variability (HRV) and changes in vascular tone with ozone exposure. We did not identify any studies that examined the effects of long-term ozone exposure on HRV.

Overall, the limited evidence does not provide significant biological support for either of the proposed MoAs for CV effects of long-term ozone exposure discussed above. While these and other proposed mechanisms may be biologically

plausible, the lack of evidence for CV effects of long-term ozone exposure in the current scientific literature does not support the existence of any specific MoA.

Biological gradient

A well-characterized exposure-response relationship greatly strengthens an inference of causality. The exposure-response relationship for the CV effects of long-term ozone exposure is difficult to discern, however. None of the experimental animal studies of CV morbidity examined more than one exposure level and the two studies that examined the same endpoints used the same exposure level, so an exposure-response relationship across studies cannot be evaluated. Exposure-response relationships for CV morbidity outcomes from studies are not likely, as all of the Tier I studies reported null or negative results. Similarly, 10 of the 12 Tier I studies of CV mortality reviewed here reported either null associations or small positive associations in single-pollutant models that were null after controlling for co-pollutants. Among the studies of biomarkers, some biomarker levels increased with increasing increments of long-term ozone exposure, but these were most often very small changes that may be indicative of homeostatic processes.

Overall, the lack of positive associations with CV effects in the majority of studies, and the limited number of exposure concentrations used in the experimental animal studies, preclude any meaningful characterization of a potential exposure-response relationship for long-term exposure to ozone and CV effects.

Temporality

An inference of causality is strengthened by evidence for a temporal sequence between the exposure to a pollutant and an observation of the effect. Temporality is achieved by the nature of experimental animal toxicology studies and is also achieved by the studies reviewed here, given the long-term exposures. An exception to this is that two of the studies of CV mortality used mortality data collected from time periods before ozone exposure data were measured, so the exposure data were unlikely to reflect the exposures incurred before or at the time of death.

Specificity

Evidence that links a specific effect to a specific exposure can strengthen a causal inference; however, any given effect may have multiple causes. None of the CV outcomes examined in this analysis are specific to ozone. Each has other risk factors with a greater likelihood of contributing to CV effects. The classical risk factors for CVD include a family history of CVD, elevated cholesterol, elevated blood pressure, diabetes, high BMI, and current smoking; based on population-attributable risk calculations, approximately 75–85% of CHD in the US is attributed to these factors (Lloyd-Jones et al. 2010). Although one Tier I epidemiology study of biomarkers indicated increased cholesterol levels and two Tier II studies indicated increases in blood pressure with long-term ozone exposure, these outcomes also have multiple, well-known risk factors (e.g. diet, physical activity, family history) that are more likely contributors than the ozone exposure. We included specific criteria in our scoring system regarding whether the studies adequately controlled for potential risk factors, as discussed further below.

Experimental evidence

A causal inference can be strengthened through the availability of "natural experiments," where a change in exposure results in a change of occurrence of an effect. We did not identify any natural experiments regarding long-term ozone exposure and CV effects.

Confounding

Because cofounders can be partially or fully responsible for the observed associations between an exposure and a health outcome, it is imperative that they are considered in studies. Lifestyle factors such as smoking and physical activity affect the risks of CV morbidity and mortality (USPHS 2014, CDC 1999), and are conceivably correlated with ozone exposures. In addition, many co-pollutants, particularly PM or certain PM species (e.g. sulfate), have been shown to confound associations between ozone and CV effects (e.g. Katsouyanni et al. 2009, Franklin and Schwartz 2008). In our study quality rating system, we scored studies that accounted for potential confounders, including analyses of co-pollutants, higher than studies that did not. Even in the studies that adjusted for the most variables, confounding from factors that were not considered in the analyses can contribute to uncertainty in the findings. Among the five studies of CV mortality that accounted for co-pollutants and reported statistically significant effects, the results were no longer significant or were significantly reduced in three of the studies when confounding factors were controlled for. It is unlikely that confounders had a major impact on the experimental animal studies.

Bias

The main sources of bias in ozone studies are selection bias, exposure measurement error, and outcome misclassification. Selection bias and outcome misclassification were better controlled for in some studies but may have been an important source of bias in others, for which the magnitude and direction of the bias was difficult to discern. For example, studies that evaluated CV-related mortality could have been subject to outcome misclassification because a CV-related endpoint may not have been the underlying cause of death. As noted above for several studies, the lack of coherence between the results for CV-related mortality and all-cause mortality may call into question the findings for CV-related deaths.

Exposure measurement error was a possible source of bias across all studies (Rhomberg et al. 2011). Several studies used ambient measurements close to participants' residences, which had lower potential for exposure measurement error than studies that relied on area-wide exposures, but they still cannot completely account for factors such as individual mobility and exposures indoors and from the workplace. Other studies used central-site monitors as surrogates of personal exposure, and had the highest risk of bias from exposure measurement error. The magnitude and direction of this potential bias likely differed across studies, as personal-ambient ozone correlations differ as a function of factors specific to the individual,

season, and city, including time-activity patterns, building characteristics, and ventilation practices, which vary by location and season. These sources of bias did not impact findings in experimental animal studies.

Adversity

The causal question of this analysis is to establish whether long-term exposure to ozone can cause adverse effects on the CV system, so it is important to evaluate whether the reported effects of ozone on CV morbidity outcomes may be homeostatic processes. In general, effects that are adaptive, compensatory, transient, or reversible are less likely to be adverse; if they are precursors to apical effects or are severe or irreversible, they are more likely to be adverse (Goodman et al. 2010).

Many of the studies examined inherently adverse disease outcomes such as MI, stroke, and CVD; these studies reported mainly null or negative associations with long-term ozone exposure. Other studies examined surrogate endpoints such as blood pressure, CIMT, and biomarker levels. The reported increases in blood pressure associated with similar increments in ozone level varied among the two Tier II studies that examined this endpoint. One study (Dong et al. 2013a) reported very small increases (e.g. 0.73 mmHg systolic and 0.37 mmHg diastolic) that may not be clinically relevant, as, based on our professional judgment, these are within the normal variability of blood pressure measurements. The other study (Chuang et al. 2011) reported increases that were more than an order of magnitude higher (21.51 mmHg systolic and 20.56 mmHg diastolic); based on our professional judgment, these increases seem implausible, as such increases are only observed in individuals taking multiple anti-hypertensive medications. Another Tier II study (Breton et al. 2012) reported small increases in CIMT that are of uncertain clinical relevance. It is not known if the reported effects on these surrogate endpoints, if attributable to ozone, would manifest into clinically significant disease outcomes. Similarly, the reported effects on biomarkers are also small in magnitude and may not be clinically relevant, as they may be more likely indicative of homeostatic processes.

Evaluation of alternative accounts

We evaluated alternative accounts of the observations from all realms of evidence regarding long-term ozone exposure and CV effects. One account is that ozone is a causal factor for adverse CV effects. This account is supported by a small number of statistically significant, positive associations of ozone with certain CV outcomes, but it requires one to dismiss the many alternative explanations for these associations despite their plausibility (e.g. that the few positive associations reported in studies of CV morbidity and mortality are due to chance or confounding). This account requires that one focuses only on the few positive findings, regardless of their small magnitude, clinical significance, or lack of confirmation in other studies of similar or higher quality. This account also requires an explanation for the lack of coherence between the animal and human data for effects on CV morbidity endpoints or biomarkers; for example, the null or negative associations reported for specific disease endpoints in all of the Tier I studies are not coherent with the reported effects on cardiac function in the *in vivo* and *ex vivo* studies in rats.

Further, this account requires that one accepts the existence of an exposure-response relationship, even though there was a lack of a consistently observed exposure-response among the studies and none of the animal studies examined more than one exposure level (making exposure-response impossible to measure). It also requires that one relies on the data from biomarker studies to contribute to the understanding of a biologically plausible MoA, even though a mechanism for potential CV effects of ozone is unknown and many other factors besides ozone contribute to the small and potentially homeostatic changes in the levels of some of the biomarkers examined. To accept this account as true, one must accept that long-term exposure to ambient levels of ozone induces adverse effects on various measures of CV function, even though the body of evidence reviewed above does not support this.

An alternative account is that ozone is not a causal factor for adverse effects on the CV system and that the few positive associations observed in some of the studies across realms of evidence are attributable to other factors. This account requires one to accept that the biological support for a MoA for CV effects of ozone is not adequate. This account also requires one to accept that the few positive effects at high exposures in animal studies are not relevant to humans, particularly at ambient ozone concentrations. This account is supported by the totality of the data across realms of evidence, which provides plausibility for the few associations observed in some studies to be deemed false positive results that are likely attributable to bias, chance, or confounding (especially given that most of these associations are reported in the Tier II studies). If this account is true, a causal relationship between long-term ozone exposure and adverse CV effects would be understood as not likely in humans, based on the currently available evidence, and the few positive associations would be attributed to alternative explanations.

When assessing the WoE in support of the competing accounts, it is clear that the first account requires more *ad hoc* assumptions and is not adequately supported by the totality of the currently available data. Thus, the WoE for this account is weak compared to the more substantial WoE supporting the alternative account of a lack of a causal relationship between long-term ozone exposure and adverse CV effects.

Phase 4 – Drawing conclusions based on inferences

We applied the WoE conclusions from Phase 3 to categorize the causal relationship between long-term ambient ozone exposure and adverse CV effects. We relied on the four-level categorization of the strength of the overall evidence for or against a causal relationship from exposure to effect, proposed by IOM (2008):

1. *Sufficient*: The evidence is sufficient to conclude that a causal relationship exists.
2. *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.
3. *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 formed judgment.
4. *Against*: The evidence suggests the lack of a causal relationship.

Our WoE conclusions indicate that the evidence does not support a causal relationship, so there is no question that the evidence for a causal association falls "below equipoise." The question is whether the database of available scientific literature is adequate to conclude that the evidence is sufficient to suggest a lack of a causal relationship. Several of the CV morbidity and biomarker endpoints were examined in only one or two studies each, limiting confirmation of the results in other settings. The epidemiological studies are limited, in that chance, bias, and confounding cannot be ruled out with confidence. Exposure measurement error may have impacted the findings across all epidemiology studies, as they used central-site monitors as surrogates for personal ozone exposure. Although some studies adjusted for many potential confounding factors, residual confounding and confounding from other factors not considered in the analyses may also contribute to uncertainty in the findings. The experimental animal studies are of limited relevance to humans, given the high ozone exposure concentrations used. The biomarker studies did not have consistent results and are of uncertain relevance, given that many other factors besides ozone contribute to changes in these markers, and their clinical significance is unclear. Because of these limitations and the small number of studies for each endpoint, the overall database for the potential CV effects of long-term ozone exposure does not provide definitive evidence regarding a lack of causation.

When considering the study limitations discussed above, we conclude that the currently available evidence as a whole is not sufficient to make a scientifically formed judgment regarding the lack of a causal relationship, and we categorize the strength of evidence for a causal relationship between long-term exposure to ozone and CV effects as "below equipoise."

Framework comparison

In the ISA, the EPA evaluated the relationship between long-term ozone exposure and CV effects using the NAAQS causal framework, which includes a five-level categorization scheme for the strength of the overall evidence for causality: Causal relationship; Likely to be a causal relationship; Suggestive of a causal relationship; Inadequate to infer a causal relationship; and Not likely to be a causal relationship (US EPA 2013). The EPA concluded that the body of evidence for long-term ozone exposure and CV effects is Suggestive of a causal relationship. To meet this classification, the EPA had to consider that "at least one high-quality epidemiological study shows an association with a given health outcome but the results of other studies are inconsistent" or that "a well-conducted toxicological study…shows effects in animal species" (US EPA 2013).

Below, we compare our current evaluation, based on the Goodman WoE framework (Goodman et al. 2013), to the EPA's evaluation in the ISA, and we discuss the possible reasons for the discrepancy in causal conclusions between the two evaluations.

The EPA identified a limited number of relevant studies for its assessment, evaluating only one epidemiology study of CV morbidity (Chuang et al. 2011), three epidemiology studies that examined CV biomarkers (Chen et al. 2007, Chuang et al. 2011, Forbes et al. 2009), and one epidemiology study of CV mortality (Jerrett et al. 2009). Although five other studies of CV mortality were identified in the evaluation of total mortality in the ISA, these were not included in the EPA's assessment of causality for CV effects. The EPA also evaluated two experimental animal studies of CV morbidity (Chuang et al. 2009, Perepu et al. 2010) and two experimental animal studies of CV biomarkers (Perepu et al. 2010, Kodavanti et al. 2011). The EPA concluded that the epidemiology evidence for CV morbidity effects and CV mortality is limited, and that further research is needed to understand whether there are any effects of long-term ozone exposure on CV outcomes in humans. The EPA also concluded that there was evidence of effects in experimental animal studies, so the generally limited body of evidence is suggestive of a causal relationship. By contrast, we identified seven epidemiology studies of CV morbidity, three epidemiology studies of CV biomarkers, 12 epidemiology studies of CV mortality, five experimental animal studies of CV morbidity, and five experimental animal studies of CV biomarkers. The majority of studies we identified that were not in the assessment of CV effects in the ISA were published prior to 2013 and could have been included in the ISA. If we had limited our analysis to the studies assessed in the ISA – which included only morbidity studies of surrogate, rather than adverse (disease) endpoints and only one Tier I study of CV mortality – we still would have concluded that the evidence does not support a causal relationship but is not sufficient to make a scientifically formed judgment regarding a lack of causality and that it should be categorized as "below equipoise."

The EPA did not evaluate the results of each study in the context of its strengths and limitations, nor did it conduct a systematic evaluation of studies such that those of higher quality received more weight in the analysis. In addition, the EPA did not discuss the studies in the context of the modified Bradford Hill aspects noted in the NAAQS causal framework. We found that a consideration of these factors led to the conclusion that 1) the higher quality (Tier I) epidemiology studies indicate only null or negative associations for specific CV disease endpoints, 2) the vast majority of reported effects in the studies in each realm are very small in magnitude, and 3) the reported effects are not supported by exposure-response relationships or a biologically plausible MoA.

The EPA also did not consider the overall body of evidence across realms. The EPA presented its findings for each realm independently, with no evaluation of whether the results were coherent across outcomes. The EPA concluded that the epidemiology evidence for CV morbidity effects and CV mortality is limited, and that there is evidence of effects in experimental animal studies, so the generally limited body of evidence is suggestive of a causal relationship. We found an overall lack of effects in the epidemiology literature, including biomarker studies, and that experimental animal studies were not likely to be relevant to humans and did not lend support to any of the proposed human MoAs. This is supported by the general lack of effects on CV morbidity endpoints in the epidemiology studies, particularly the higher quality (Tier I) studies.

The EPA also did not fully consider the adversity of outcomes in the studies it evaluated. For example, the EPA did not evaluate whether any changes in biomarker levels were large enough to represent or lead to adverse effects, or whether these changes were more likely to be homeostatic. We found

that the biological changes reported were not likely to be on an adverse pathway and thus unlikely to be indicative of or lead to adverse effects.

Although the NAAQS causal framework includes the weighing of alternative views on controversial issues, in the ISA, the EPA evaluated only whether evidence supports a causal relationship – it did not consider whether the data support an alternative explanation. In contrast, the Goodman WoE framework is designed to evaluate not just causality, but also whether the data support alternative explanations. We found that the WoE better supports a lack of a causal relationship between long-term ambient ozone exposure and adverse CV effects than a causal relationship.

In contrast to the four-level categorization scheme in the IOM framework, the NAAQS causal framework includes a five-level categorization scheme for causal determinations. As noted above, the classification of evidence as Suggestive of a causal relationship requires either one high-quality epidemiology study to show an association or one well-conducted study in experimental animals to show the effects, and the EPA based its Suggestive categorization for the CV effects of long-term exposure to ozone mainly on experimental animal studies. Even if these studies are of high quality, they are unlikely to be relevant to humans because they were not conducted with relevant ozone exposures, and one of the studies was an *ex vivo* study conducted in rat hearts after an ischemia-reperfusion injury. Despite this, the preponderance of evidence across realms does not support a causal relationship between long-term exposure to ozone and adverse CV effects. Thus, even if the data met the EPA criteria for the Suggestive category, as a whole, the data clearly are not suggestive of an association. The IOM framework is more appropriate for making conclusions regarding causation, particularly because it does not allow for a single study – which may be inconsistent with the preponderance of other data – to provide evidence that a causal relationship is at least as likely as not.

Conclusions

The current WoE provides no convincing case for a causal relationship between long-term ambient ozone exposure and its adverse effects on the CV system, and the few positive associations reported in some epidemiology studies are most likely attributable to alternative explanations such as bias, chance, or confounding. Because of the limitations of the available studies, however, they do not provide definitive evidence regarding a lack of causation. We conclude that the evidence as a whole is not sufficient to make a scientifically formed judgment regarding a lack of a causality but, based on the available evidence, we categorize the strength of evidence for a causal relationship between long-term exposure to ozone and CV effects as "below equipoise."

Declaration of interest

The authors are employed by Gradient, a private environmental consulting firm, and Albany Medical College, a private medical school. The Gradient staff have strong expertise in assessing human, experimental animal, and mechanistic data in WoE analyses (as is evident in recent evaluations conducted for bisphenol A, naphthalene, formaldehyde, chlorpyrifos, methanol, styrene, nickel, and toluene diisocyanate) and have presented several of these analyses to regulatory bodies. In addition, Gradient staff, including the authors of this paper, have carefully evaluated the science underlying EPA's review of various NAAQS and offered both oral and written testimony to EPA. Gradient has also addressed issues on systematic review and integration of evidence for a number of clients. The authors conducted the work reported in this paper during the normal course of employment, with financial support provided by the Texas Commission on Environmental Quality (TCEQ). The authors have the sole responsibility for the writing, content, and conclusions in this paper. The conclusions are not necessarily those of the TCEQ.

References

Abbey DE, Nishino N, McDonnell WF, Burchette RJ, Knutsen SF, Lawrence Beeson W, Yang JX. (1999). Long-term inhalable particles and other air pollutants related to mortality in nonsmokers. Am J Respir Crit Care Med, 159, 373–82.

Atkinson RW, Carey IM, Kent AJ, van Staa TP, Anderson HR, Cook DG. (2013). Long-term exposure to outdoor air pollution and incidence of cardiovascular diseases. Epidemiology, 24, 44–53.

Barath S, Langrish JP, Lundback M, Bosson JA, Goudie C, Newby DE, et al. (2013). Short-term exposure to ozone does not impair vascular function or affect heart rate variability in healthy young men. Toxicol Sci, 135, 292–9.

Beckerman BS, Jerrett M, Finkelstein M, Kanaroglou P, Brook JR, Arain MA, et al. (2012). The association between chronic exposure to traffic-related air pollution and ischemic heart disease. J Toxicol Environ Health A, 75, 402–11.

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, 988–95.

Breton CV, Wang X, Mack WJ, Berhane K, Lopez M, Islam TS, et al. (2012). Childhood air pollutant exposure and carotid artery intima-media thickness in young adults. Circulation, 126, 1614–20.

Carey IM, Atkinson RW, Kent AJ, van Staa T, Cook DG, Anderson HR. (2013). Mortality associations with long-term exposure to outdoor air pollution in a national English cohort. Am J Respir Crit Care Med, 187, 1226–33.

Carter WPL, Seinfeld JH. (2012). Winter ozone formation and VOC incremental reactivities in the Upper Green River Basin of Wyoming. Atmos Environ, 50, 255–66.

Centers for Disease Control (CDC). (1999). Physical Activity and Health: A Report of the Surgeon General, 1996. [Online] Available at http://www.cdc.gov/nccdphp/sgr/sgr.htm [Accessed on 21 May 2014].

Chen C, Arjomandi M, Balmes J, Tager I, Holland N. (2007). Effects of chronic and acute ozone exposure on lipid peroxidation and antioxidant capacity in healthy young adults. Environ Health Perspect, 115, 1732–7.

Chen LH, Knutsen SF, Shavlik D, Beeson WL, Petersen F, Ghamsary M, Abbey D. (2005). The association between fatal coronary heart disease and ambient particulate air pollution: Are females at greater risk? Environ Health Perspect, 113, 1723–9.

Chuang GC, Yang Z, Westbrook DG, Pompilius M, Ballinger CA, White CR, et al. (2009). Pulmonary ozone exposure induces vascular dysfunction, mitochondrial damage, and atherogenesis. Am J Physiol Lung Cell Mol Physiol, 297, L209–L216.

Chuang KJ, Yan YH, Cheng TJ. (2010). Effect of air pollution on blood pressure, blood lipids, and blood sugar: A population-based approach. J Occup Environ Med, 52, 258–62.

Chuang KJ, Yan YH, Chiu SY, Cheng TJ. (2011). Long-term air pollution exposure and risk factors for cardiovascular diseases among the elderly in Taiwan. Occup Environ Med, 68, 64–8.

Doggrell SA, Brown L. (1998). Rat models of hypertension, cardiac hypertrophy and failure. Cardiovasc Res, 39, 89–105.

Dong GH, Qian Z, Wang J, Chen W, Ma W, Trevathan E, et al. (2013b). Associations between ambient air pollution and prevalence of stroke and cardiovascular diseases in 33 Chinese communities. Atmos Environ, 77, 968–73.

Dong GH, Qian ZM, Xaverius PK, Trevathan E, Maalouf S, Parker J, et al. (2013a). Association between long-term air

pollution and increased blood pressure and hypertension in China. Hypertension, 61, 578–84.
Dorato MA, Wolff RK. (1991). Inhalation exposure technology, dosimetry, and regulatory issues. Toxicol Pathol, 19(4 Part 1), 373–83.
Droogmans S, Lauwers R, Cosyns B, Roosens B, Franken PR, Weytjens C, et al. (2008). Impact of anesthesia on valvular function in normal rats during echocardiography. Ultrasound Med Biol, 34, 1564–72.
Eskesen K, Olsen NT, Dimaano VL, Pinheiro A, Sogaard P, Fritz-Hansen T, et al. (2012). New approach to intracardiac hemodynamic measurements in small animals: Echo-guided percutaneous apical puncture. J Ultrasound Med, 31, 1233–8.
Forbes LJ, Patel MD, Rudnicka AR, Cook DG, Bush T, Stedman JR, et al. (2009). Chronic exposure to outdoor air pollution and markers of systemic inflammation. Epidemiology, 20, 245–53.
Franklin M, Schwartz J. (2008). The impact of secondary particles on the association between ambient ozone and mortality. Environ Health Perspect, 116, 453–8.
Goodman JE, Dodge DG, Bailey LA. (2010). A framework for assessing causality and adverse effects in humans with a case study of sulfur dioxide. Regul Toxicol Pharmacol, 58, 308–22.
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, 829–49.
Goodman JE, Prueitt RL, Sax SN, Lynch HN, Zu K, Lemay JC, et al. (2014). Weight-of-evidence evaluation of short-term cardiovascular effects of ozone. Crit Rev Toxicol, 44, 725–90.
Goodman M, Teta MJ, Hessel PA, Garabrant DH, Craven VA, Scrafford CG, Kelsh MA. (2004). Mesothelioma and lung cancer among motor vehicle mechanics: A meta-analysis. Ann Occup Hyg, 48, 309–26.
Gordon CJ, Jarema KA, Lehmann JR, Ledbetter AD, Schladweiler MC, Schmid JE, et al. (2013). Susceptibility of adult and senescent Brown Norway rats to repeated ozone exposure: An assessment of behavior, serum biochemistry and cardiopulmonary function. Inhal Toxicol, 25, 141–59.
Hatch GE, McKee J, Brown J, McDonnell W, Seal E, Soukup J, et al. (2013). Biomarkers of dose and effect of inhaled ozone in resting versus exercising human subjects: Comparison with resting rats. Biomark Insights, 8, 53–67.
Hatch GE, Slade R, Harris LP, McDonnell WF, Devlin RB, Koren HS, et al. (1994). Ozone dose and effect in humans and rats. A comparison using oxygen,-18 labeling and bronchoalveolar lavage. Am J Respir Crit Care Med, 150, 676–83.
Higgins JPT, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD, et al.; Cochrane Bias Methods Group, Cochrane Statistical Methods Group. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ, 343, d5928.
Hill AB. (1965). The environment and disease: Association or causation? Proc R Soc Med, 58, 295–300.
Hooijmans CR, Leenaars M, Ritskes-Hoitinga M. (2010). A gold standard publication checklist to improve the quality of animal studies, to fully integrate the Three Rs, and to make systematic reviews more feasible. Altern Lab Anim, 38, 167–82.
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. pp. 781. Available at http://books.nap.edu/openbook.php?record_id = 11908&
Ioannidis JP, Tzoulaki I. (2012). Minimal and null predictive effects for the most popular blood biomarkers of cardiovascular disease. Circ Res, 110, 658–62.
Jakobsen CJ, Torp P, Vester AE, Folkersen L, Thougaard A, Sloth E. (2010). Ketamine reduce left ventricular systolic and diastolic function in patients with ischaemic heart disease. Acta Anaesthesiol Scand, 54.
Janke K, Propper C, Henderson J. (2009). Do current levels of air pollution kill? The impact of air pollution on population mortality in England. Health Econ, 18, 1031–55.
Jerrett M, Burnett RT, Ma R, Pope CA, Krewski D, Newbold KB, et al. (2005). Spatial analysis of air pollution and mortality in Los Angeles. Epidemiology, 16, 727–36.
Jerrett M, Burnett RT, Pope CA, Ito K, Thurston G, Krewski D, et al. (2009). Long-term ozone exposure and mortality. N Engl J Med, 360, 1085–95.
Katsouyanni K, Samet JM, Anderson HR, Atkinson R, Le Tertre A, Medina S, et al. (2009). Air Pollution and Health: A European and North American Approach (APHENA). HEI Research Report 142. Health Effects Institute, Boston, MA . pp.132.
Khosravi A, Rao C, Naghavi M, Taylor R, Jafari N, Lopez AD. (2008). Impact of misclassification on measures of cardiovascular disease mortality in the Islamic Republic of Iran: A cross-sectional study. Bull World Health Org, 86, 688–96.
Kilkenny C, Browne WJ, Cuthill IC, Emerson M, Altman DG. (2010). Improving bioscience research reporting: The ARRIVE guidelines for reporting animal research. PLoS Biol, 8, e1000412.
Klaassen CD, ed. (2008). Casarett and Doull's Toxicology: The Basic Science of Poisons, 7th edition. McGraw-Hill Companies, Inc, pp. 1309.
Kodavanti UP, Thomas R, Ledbetter AD, Schladweiler MC, Shannahan JH, Wallenborn JG, et al. (2011). Vascular and cardiac impairments in rats inhaling ozone and diesel exhaust particles. Environ Health Perspect, 119, 312–18.
Krewski D, Jerrett M, Burnett RT, Ma R, Hughes E, Shi Y, et al. (2009). Extended follow-up and spatial analysis of the American Cancer Society Study Linking Particulate Air Pollution and Mortality. HEI Research Report 140. Health Effects Institute, Boston, MA. pp. 154.
Lewington S, Bragg F, Clarke R. (2012). A review on metaanalysis of biomarkers: Promises and pitfalls. Clin Chem, 58, 1192–204.
Lipsett MJ, Ostro BD, Reynolds P, Goldberg D, Hertz A, Jerrett M, et al. (2011). Long-term exposure to air pollution and cardiorespiratory disease in the California teachers study cohort. Am J Respir Crit Care Med, 184, 828–35.
Lloyd-Jones DM, Hong Y, Labarthe D, Mozaffarian D, Appel LJ, Van Horn L, et al.; American Heart Association Strategic Planning Task Force and Statistics Committee. (2010). Defining and setting national goals for cardiovascular health promotion and disease reduction: The American Heart Association's strategic Impact Goal through 2020 and beyond. Circulation, 121, 586–613.
Macleod MR, Fisher M, O'Collins V, Sena ES, Dirnagl U, Bath PM, et al. (2009). Reprint: Good laboratory practice: Preventing introduction of bias at the bench. J Cereb Blood Flow Metab, 29, 221–3.
Medinsky MA. (1996). Determinants of gas and vapor uptake in the respiratory tract. CIIT Activities, 16, 1–6.
Miller FJ. (1995). Uptake and fate of ozone in the respiratory tract. Toxicol Lett, 82–83, 277–85.
National Research Council (NRC), Committee on Tropospheric Ozone Formation and Measurement. (1991). Rethinking the Ozone Problem in Urban and Regional Air Pollution. Washington, DC: National Academy Press, pp. 523.
Nichols JL, Owens EO, Dutton SJ, Luben TJ. (2013). Systematic review of the effects of black carbon on cardiovascular disease among individuals with pre-existing disease. Int. J. Public Health, 58, 707–24.
Pascal M, Wagner V, Chatignoux E, Falq G, Corso M, Blanchard M, et al. (2012). Ozone and short-term mortality in nine French cities: Influence of temperature and season. Atmos Environ, 62, 566–72.
Perepu RS, Dostal DE, Garcia C, Kennedy RH, Sethi R. (2012). Cardiac dysfunction subsequent to chronic ozone exposure in rats. Mol Cell Biochem, 360, 339–45.
Perepu RS, Garcia C, Dostal D, Sethi R. (2010). Enhanced death signaling in ozone-exposed ischemic-reperfused hearts. Mol Cell Biochem, 336, 55–64.
Pope CA, Burnett RT, Thun MJ, Calle EE, Krewski D, Ito K, Thurston GD. (2002). Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. JAMA, 287, 1132–41.
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, 651–71.
Rhomberg LR, Goodman JE, Bailey LA, Prueitt RL, Beck NB, Bevan C, et al. (2013). A survey of frameworks for best practices in weight-of-evidence analyses. Crit Rev Toxicol, 43, 753–84.
Rothman KJ, Greenland S. (1998). Modern Epidemiology (Second Edition). Philadelphia: Lippincott Williams & Wilkins, pp. 738.
Rouslin W, Broge CW, Guerrieri F, Capozza G. (1995). ATPase activity, IF1 content, and proton conductivity of ESMP from control and ischemic slow and fast heart-rate hearts. J Bioenerg Biomembr, 27, 459–66.
Sarwar N, Thompson AJ, Di Angelantonio E. (2009). Markers of inflammation and risk of coronary heart disease. Dis Markers, 26, 217–25.
Sethi R, Manchanda S, Perepu RS, Kumar A, Garcia C, Kennedy RH, et al. (2012). Differential expression of caveolin-1 and caveolin-3: Potential marker for cardiac toxicity subsequent to chronic ozone inhalation. Mol Cell Biochem, 369, 9–15.

Smith KR, Jerrett M, Anderson HR, Burnett RT, Stone V, Derwent R, et al. (2009). Public health benefits of strategies to reduce greenhouse-gas emissions: Health implications of short-lived greenhouse pollutants. Lancet, 374, 2091–103.

Spencer-Hwang R, Knutsen SF, Soret S, Ghamsary M, Beeson WL, Oda K, et al. (2011). Ambient air pollutants and risk of fatal coronary heart disease among kidney transplant recipients. Am J Kidney Dis, 58, 608–16.

Tao Y, Huang W, Huang X, Zhong L, Lu SE, Li Y, et al. (2012). Estimated acute effects of ambient ozone and nitrogen dioxide on mortality in the Pearl River Delta of southern China. Environ Health Perspect, 120, 393–8.

Taubes G. (1995). Epidemiology faces its limits. Science, 269, 164–9.

US EPA (1998). Health Effects Test Guidelines: OPPTS 870.3465–90-Day Inhalation Toxicity. Office of Prevention, Pesticides and Toxic Substances, EPA, 712-C-98-204, pp. 17, August.

US EPA. (2006). Air Quality Criteria for Ozone and Related Photochemical Oxidants. Volume 1. Research Triangle Park, NC: Office of Research and Development, National Center for Environmental Assessment (NCEA), EPA, 600/R-05/004aF, pp. 821, February.

US EPA. (2013). Integrated Science Assessment for Ozone and Related Photochemical Oxidants (Final). National Center for Environmental Assessment (NCEA), EP, A/600/R–10/076F, pp. 1251, February.

US Public Health Service (USPHS). 2014. The Health Consequences of Smoking—50 Years of Progress: A Report of the Surgeon General. Office of the Surgeon General. pp. 1440. http://www.surgeongeneral.gov/library/reports/50-years-of-progress/

van der Worp HB, Howells DW, Sena ES, Porritt MJ, Rewell S, O'Collins V, Macleod MR. (2010). Can animal models of disease reliably inform human studies? PLoS Med, 7, e1000245.

Wang L, Yang G, Jiemin M, Rao C, Wan X, Dubrovsky G, Lopez AD. (2007). Evaluation of the quality of cause of death statistics in rural China using verbal autopsies. J Epidemiol Community Health, 61, 519–26.

Wang XY, Hu W, Tong S. (2009). Long-term exposure to gaseous air pollutants and cardio-respiratory mortality in Brisbane, Australia. Geospat Health, 3, 257–63.

Watkinson WP, Campen MJ, Nolan JP, Costa DL. (2001). Cardiovascular and systemic responses to inhaled pollutants in rodents: Effects of ozone and particulate matter. Environ Health Perspect, 109(Suppl 4), 539–46.

Watkinson WP, Campen MJ, Wichers LB, Nolan JP, Costa DL. (2003). Cardiac and thermoregulatory responses to inhaled pollutants in healthy and compromised rodents: Modulation via interaction with environmental factors. Environ Res, 92, 35–47.

Yang C, Yang H, Guo S, Wang Z, Xu X, Duan X, Kan H. (2012). Alternative ozone metrics and daily mortality in Suzhou: The China Air Pollution and Health Effects Study (CAPES). Sci Total Environ, 426, 83–9.

Zanobetti A, Schwartz J. (2011). Ozone and survival in four cohorts with potentially predisposing diseases. Am J Respir Crit Care Med, 184, 836–41.

Zhou X, Fragala MS, McElhaney JE, Kuchel GA. (2010). Conceptual and methodological issues relevant to cytokine and inflammatory marker measurements in clinical research. Curr Opin Clin Nutr Metab Care, 13, 541–7.

Zile MR, Brutsaert DL. (2002). New concepts in diastolic dysfunction and diastolic heart failure: Part I. Diagnosis, prognosis, and measurements of diastolic function. Circulation, 105, 1387–93.