Title

Long-term ozone exposure associated cause-specific mortality risks with adjusted metrics by cohort studies:

A systematic review and meta-analysis

Authors

Haitong Zhe Sun1,2, Pei Yu3, Changxin Lan4,5, Michelle Wan1, Sebastian Hickman1, Jayaprakash Murulitharan1, Huizhong Shen6, Le Yuan1, Yuming Guo3*, Alexander T. Archibald1,7*

Affiliations

1 Centre for Atmospheric Science, Yusuf Hamied Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, United Kingdom
2 Department of Earth Sciences, University of Cambridge, Cambridge CB2 3EQ, United Kingdom
3 School of Public Health and Preventive Medicine, Monash University, Melbourne, Victoria 3004, Australia
4 Institute of Reproductive and Child Health, Key Laboratory of Reproductive Health, National Health Commission of the People's Republic of China, Beijing 100191, China
5 Department of Epidemiology and Biostatistics, School of Public Health, Peking University, Beijing 100191, China
6 School of Environmental Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China
7 National Centre for Atmospheric Science, Cambridge CB2 1EW, United Kingdom

* Correspondence to: Alexander T. Archibald (ata27@cam.ac.uk) and Yuming Guo (yuming.guo@monash.edu)

main-text 6749 words + 3 tables + 6 figures

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
**ABSTRACT**

**BACKGROUND:** Long-term ozone (O$_3$) exposure could lead to a series of non-communicable diseases and increase the mortality risks. However, cohort-based studies were still rather rare, and inconsistent exposure metrics might impair the credibility of epidemiological evidence synthesis. To provide more accurate meta-estimation, this review updated the systematic review with inclusion of recent studies and summarised the quantitative associations between O$_3$ exposure and cause-specific mortality risks based on unified exposure metrics.

**METHODS:** Research articles reporting relative risks between incremental long-term O$_3$ exposure and causes of mortality covering all-cause, cardiovascular diseases, respiratory diseases, chronic obstructive pulmonary disease, pneumonia, ischaemic heart diseases, ischaemic stroke, congestive heart failure, cerebrovascular diseases, and lung cancer, estimated from cohort studies were identified through systematic searches in MEDLINE, Embase and Web of Science. Cross-metric conversion factors were estimated linearly by decadal of observations during 1990-2019. The Hunter-Schmidt random effect estimator was applied to pool the relative risks.

**RESULTS:** A total of 25 studies involving 226,453,067 participants (14 unique cohorts covering 99,855,611 participants) were included in the systematic review. After linearly adjusting the inconsistent O$_3$ exposure metrics into congruity, the pooled relative risks (RR) associated with every 10 nmol mol$^{-1}$ (ppbV) incremental O$_3$ exposure, by mean of warm-season daily maximum 8-hour average metric, was: 1.014 with 95% confidence interval (CI) ranging 1.009–1.019 for all-cause mortality; 1.025 (95% CI: 1.010–1.040) for respiratory mortality; 1.056 (95% CI: 1.029–1.084) for COPD mortality; 1.019 (95% CI: 1.004–1.035) for cardiovascular mortality; and 1.096 (95% CI: 1.065–1.129) for congestive heart failure mortality. Insignificant mortality risk associations were found for ischaemic heart disease, cerebrovascular diseases and lung cancer.

**DISCUSSION:** This review covered up-to-date studies, expanded the O$_3$-exposure associated mortality causes into wider range of categories, and firstly highlighted the issue of inconsistency in O$_3$ exposure metrics. Non-intercept linear regression-based cross-metric RR conversion was another innovation, but limitation lay in the observation reliance, indicating further calibration with more credible observations available. Large uncertainties in the multi-study pooled RRs would inspire more future studies to corroborate or contradict the results from this review.
CONCLUSION: Adjustment for exposure metrics laid more solid foundation for multi-study meta-analysis, and wider coverage of surface O\textsubscript{3} observations are anticipated to strengthen the cross-metric conversion in the future. Ever-growing numbers of epidemiological studies supported unneglectable cardiopulmonary hazards and all-cause mortality risks from long-term O\textsubscript{3} exposure. However, evidences on long-term O\textsubscript{3} exposure associated health effects were still scarce, and hence more relevant studies are encouraged to cover more population with regional diversity.

REGISTRATION: The review was registered in PROSPERO (CRD42021270637).

FUNDING: This study is mainly funded by UK Natural Environment Research Council, UK National Centre for Atmospheric Science, Australian Research Council and Australian National Health and Medical Research Council.

Keywords

Ozone; long-term exposure; mortality; cohort; metric; respiratory; cardiovascular

Highlights

1. Updated evidence for O\textsubscript{3}-mortality associations from 25 cohorts has been provided.
2. Adjusting various O\textsubscript{3} exposure metrics can provide more accurate risk estimations.
3. Long-term O\textsubscript{3}-exposure was associated with increased mortality from all-causes, respiratory disease, COPD, cardiovascular disease and congestive heart failure.
1 INTRODUCTION

Atmospheric ozone (O$_3$) is a short-lived climate forcer.\(^1\) Besides warming the global atmosphere, O$_3$ in the stratosphere can abate the radiation hazards from ultraviolet rays onto organisms, while O$_3$ in the ambient air is of detrimental defects on ecosystem and human health,\(^2\)-\(^4\) and hence health effects caused from exposure to surface O$_3$ have become a serious public concern. Short-term (i.e. hours to days) exposure to high-level O$_3$ can cause a series of acute symptoms like asthma, respiratory tract infection, myocardial infarction, and cardiac arrest,\(^5\)-\(^8\) and long-term (i.e. over years) exposure can lead to chronic health outcomes including but not limited to preterm delivery, stroke, chronic obstructive pulmonary diseases, and cerebrovascular diseases.\(^9\)-\(^12\) Long-term ambient O$_3$ exposure was estimated to be responsible for over 0.36 million premature deaths globally in 2019 according to the Global Burden of Disease (GBD) report released by the Institute for Health Metrics and Evaluation (IHME).\(^13\)

Systematic reviews summarising the associations between the adverse health outcomes, and both the short-term and long-term O$_3$ exposures, have been performed in previous studies.\(^14\)-\(^16\) Studies on short-term O$_3$ exposure-induced morbidities are comparatively more abundant than the long-term O$_3$ exposure studies where the epidemiological evidences are less congruous. Some deficiencies are spotted in the two reviews for long-term O$_3$ exposure-associated mortality risk studies,\(^15\), \(^16\) the primary issue of which is the inconsistent use of various O$_3$ exposure metrics; however, no other reviews are found to remedy these flaws. As a secondary photolytic gaseous air pollutant, the warm-season and diurnal concentrations of surface O$_3$ will be much higher than cool-season and nocturnal concentrations,\(^17\), \(^18\) and thus the average and peak metrics of O$_3$ concentrations shall be of drastically different realistic implications.\(^19\) Under this circumstance, directly pooling the relative risks scaled in different metrics might lead to biases.

Atkinson et al. (2016) explored 6 types of mortality causes, but searched the literatures only till 2015;\(^16\) while Huangfu et al. (2019) updated the searches to 2018, but only 3 types of mortality causes were considered.\(^15\) We thus determine to update the review on the health effects of O$_3$ to include more categories of mortalities together with covering the most recent publications. Additionally, GBD estimations ascribed long-term O$_3$-exposure induced all-cause mortality for chronic obstructive pulmonary disease,\(^13\) which might lead to underestimations without considering other causes. It is reasonable to deduce that long-term O$_3$ exposure will exacerbate the mortality of certain diseases given that the short-term exposure increases the morbidity risks of the same diseases, and
thus scrutinising epidemiological evidences for multiple causes of mortality will provide more credible supports to fill in this gap.

The primary innovation of our updated review is our taking full advantage of global systemic stationary observations to explore the feasibility of adjusting the various exposure metrics, and pooling the multi-study risks with the unified exposure metric, the mean of warm-season daily maximum 8-hour average, in response to the up-to-date suggestions from the Lancet global environmental health collaboration. Through this updated systematic review and meta-analysis on long-term O₃ exposure associated cause-specific mortality risks, we aim to present and evaluate epidemiological evidences for 3 major questions not fully addressed by the previous 2 reviews, as (1) which mortality causes shall be ascribed to long-term O₃ exposure; (2) have the risk associations changed given the latest studies with more mature research design and methodologies; and (3) how to estimate the quantities of the risk association strengths by the suggested exposure metric. Both our methods and discoveries are expected to inspire future O₃-health studies, and support relevant policy-making to benefit the global population.

2. METHODS

2.1 Search strategy

We searched 3 research databases (MEDLINE, Embase, and Web of Science) from 1 September, 2015 till 1 February, 2022 to finish our systematic review and meta-analysis, updated from 2 previous reviews on long-term O₃ exposure-associated mortality.¹⁵, ¹⁶ Search terms also referred to these 2 previous systematic reviews with modifications to enhance the inclusion of potential relevant studies, as we combined the keywords relevant to the cause-specific mortalities (i.e. “mortality”, “death”, “premature death”, “all-cause”, “non-accidental”, “cardiopulmonary”, “respiratory”, “chronic obstructive pulmonary disease”, “pneumonia”, cardiovascular, “lung cancer”, “cerebrovascular”, “stroke”, “ischaemic heart disease”, “congestive heart failure”), the pollutant of research interest (i.e. “ozone”), and qualified epidemiological study types (i.e. “long-term”, “cohort study”, “prospective”, “retrospective”, “longitudinal study”). The detailed search strategies were listed in Table S1. Health outcomes considered in the systematic review were: mortality from (1) all causes (AC, ICD9: 001-799, ICD10: A00-R99); (2) all respiratory diseases (RESP, ICD9: 460-519, ICD10: J00-J98); (3) chronic obstructive
pulmonary diseases and allied conditions (COPD, ICD9: 490-496, ICD10: J19-J46); (4) all cardiovascular diseases (CVD, ICD9: 390-459, ICD10: I00-I98); (5) all cerebrovascular diseases (CEVD, ICD9: 430-438, ICD10: I60-I69); (6) ischaemic heart disease (IHD, ICD9: 410-414, ICD10: I20-I25); (7) congestive heart failure (CHF, ICD9: 428, ICD10: 150); (8) ischaemic stroke (ICD9: 434, ICD10: I61-I64); (9) pneumonia (ICD9: 480-487, ICD10: J12-J18); and (10) lung cancer (LC, ICD9: 162, ICD10: C33-C34).

2.2 Study eligibility criteria

As an updated systematic review, literatures identified in the previous 2 reviews underwent examination together with the newly retrieved ones. Studies were included during screening following the criteria as: (1) the epidemiological research should be conducted based on cohorts; (2) the exposure should include O₃ as an individual risk factor; (3) the health outcomes should be all-cause or cause-specific deaths at individual level; (4) studies provided hazard ratio (HR), risk ratio (RR) or odds ratio (OR) and their 95% confidence intervals (CIs) clearly and reported by every increase unit (e.g. 10-ppbV) of exposure concentrations, assuming linear risk relationship with adjusting key confounders; (6) the study should be published as an original research article in scholarly peer-reviewed journals in English. For articles from the same cohort, only one single study covering the widest populations and the longest follow-up period was reserved for meta-analysis, unless the subgroups of participants and study follow-up periods are clearly stated to be of mild overlapping; We followed the Preferred Reporting Items for Systematic Review and Meta-Analyses (PRISMA) guidelines to process the included studies on ambient O₃ exposure induced mortality.

2.3 Study selection and scrutinisation

All searched literatures were archived in Clarivate™ Analytics Endnote X9.3.1 reference manager software. Two literature review investigators (HZS and CL) conducted title and abstract pre-screening independently for all web-searched records and reviewed the full text for the pre-screened studies. Disagreements were resolved by discussions with a third reviewer (PY).
2.4 Data extraction

Details from each screened-out literatures were extracted and labelled for the purpose of meta-analysis, including (1) the authors with publication year as study labels of reference; (2) basic descriptive information of the study cohort embracing the cohort name, country, follow-up periods, numbers of cases and total participants, population genders and ethnics, exposure metrics, health outcomes, and major confounders; (3) the risk association effects preferably quantified in HR (and also RR/OR as substitute choices) per unit incremental exposure with 95% confidence interval (CI).

2.5 Study quality assessment

All screened-out studies underwent quality evaluation using the Quality Assessment Tool of Observational Cohort and Cross-Sectional Studies developed by National Institute of Health (NIH) (https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools), aiming to ensure the studies considered for meta-analysis are adequately reliable. The assessments were cross-validated by two authors (HZS and CL) independently, with the third author (PY) supervising any disagreements. Table S2 listed 14 assessment items assigned with 1 score for each, and the tallied scores were translated into a literature-specific rating of quality. Studies scoring full-mark 14 were categorised to be “Good”, while 10-13 as “Fair” and <10 as “Poor”.

Besides applying the quality assessment tool to determine which reviewed studies should be included for meta-analysis, checking the epidemiological evidence quality from the included literatures for each cause of mortality was finished through the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) system\(^{21,22}\) to yield rating bands ranging across “high”, “moderate”, “low”, and “very low”. This grading system by default rated “high-quality” for cohort studies as the starting point of evaluation, and the rate would be downgraded by five limitations as the existence of (1) risk of bias examined by the Quality Assessment Tool (Table S2), (2) imprecision (i.e. studies did not report the central risk estimations with confidence intervals), (3) inconsistency (i.e. the directions of the estimated risks were controversial across studies), (4) indirectness (i.e. studies did not include the desired population, exposure, or health outcomes), and (5) publication bias (i.e. researchers tended to publish studies with positive results); and upgraded by three strengths as reporting (1) exposure-response trend, (2) residual confounding (i.e. adjusting the confounders highlighted the risks), and (3) strong associations. Publication biases were graphically presented by funnel plots.\(^{23}\)
and statistically tested by trim-and-fill method. The review was registered in PROSPERO (CRD42021270637).

2.6 Exposure adjustment

2.6.1 Unit unification

There were two major units used to quantify the surface O₃ concentrations, nmol mol⁻¹ or parts per billion by volume mixing ratio (ppbV) more frequently used by atmospheric modelling researchers, and milligram per cubic metre by mass concentration (µg/m³) widely used by public health studies. These two units are interchangeable to each other based on the ideal gas law $PV = nRT$, if the air temperatures (T) and pressures (P) are given, as presented in eqs 1–4.

\[
1 \text{ ppbV } O_3 = \frac{1 \times 10^{-9} \text{ mol } O_3}{1 \text{ mol air}} \quad (\text{eq. 1})
\]

\[
1 \text{ mol air } \equiv \frac{RT}{P} \times 1 \text{ mol } (m^3) = \frac{8.314 \text{ Pa} \cdot m^3 \cdot K^{-1} \times T}{P} \quad (m^3) \quad (\text{eq. 2})
\]

\[
1 \times 10^{-9} \text{ mol } O_3 \times 47.997 \text{ g/mol}^{-1} = 47.997 \times 10^{-9} \text{ g } O_3 \quad (\text{eq. 3})
\]

\[
1 \text{ ppbV } O_3 = \frac{47.997 \times 10^{-9} \text{ g } \times 10^6 \text{ g } \cdot \text{ mol}^{-1}}{8.314 \text{ Pa} \cdot m^3 \cdot K^{-1} \times T/P \text{ m}^3 \text{ air}} = 5.773 \times 10^{-3} \times \frac{P(\text{Pa})}{T(K)} \text{ O}_3 \text{ air} \mu g \cdot m^{-3} \quad (\text{eq. 4})
\]

Assuming $T = 298.65 \text{ K } (25.5^\circ \text{C})$ and $P = 101.325 \text{ kPa}$, the ppbV-µg/m³ conversion factor could be approximated as 1 ppbV ~ 1.96 µg/m³. Though the surface air temperatures and pressures would vary across seasons, such simplification was still widely used in previous studies, being of more credibility for long-term surface O₃ studies averaging the surface air temperatures and pressures at longer periods. For example, even at very low temperature of 270 K, the conversion factor was 2.17, which corroborated the stability of linear conversion.

2.6.2 Metric unification

Surface O₃, as a secondary photochemistry pollutant involving photolysis of NO₂ to trigger chains of radical reactions, has concentrations that will vary significantly between
day and night-time, and between warm and cool seasons, as discussed by numbers of
studies.\textsuperscript{17, 28-31} Under this circumstance, various daily metrics to quantify the surface O\textsubscript{3} concentrations emerged due to series of considerations, which however brought in more
difficulties to assimilate epidemiological evidences. The previous reviews simply pooled
the reported risk association strengths without adjusting the diverse metrics,\textsuperscript{15, 16} which
we thought was a fatal defect requiring improvements.

We therefore designed to update the meta-analysis by unifying the exposure metrics
for pooled O\textsubscript{3} exposure-associated risks. As suggested by the U.S. EPA final report of Air
Quality Criteria for Ozone and Other Photochemical Oxidants,\textsuperscript{32} linear relationships
were assumed to estimate the cross-metric conversion factors using long-term reliable
observations as the Tropospheric Ozone Assessment Report (TOAR) archive\textsuperscript{19} and China
National Environmental Monitoring Centre (CNEMC, http://www.cnemc.cn/en/) in our
review, and correlation matrix was used to validate that the presumptions of linearity
were not violated. Both TOAR and CNEMC sites measured the surface O\textsubscript{3} by means of
the UV absorption technique with strict quality control so as to ensure the comparability
of the records across different countries and regions.\textsuperscript{33, 34} We considered 6 complex
metrics for mutual conversion as (1) annual mean of 24-hour daily average (ADA24), (2)
6-month warm season mean of 24-hour daily average (6mDA24), (3) annual mean of
daily maximum 8-hour average (ADMA8), (4) 6-month warm season mean of daily
maximum 8-hour average (6mDMA8), (5) annual mean of daily maximum 1-hour
average (ADMA1), and (6) 6-month warm season mean of daily maximum 1-hour
average (6mDMA1). Long-term averaging-based metric conversion could smooth the
temporal variations resulting from the seasonal and geographical solar radiation
variabilities. The linear conversion factors ($k$) were mathematically defined by eq 5, to
adjust the original metric into the target one with irreducible regression errors $\epsilon$.

$$C_{\text{Adjusted}} = k_{\text{Original} \rightarrow \text{Adjusted}} \times C_{\text{Original}} + \epsilon \quad \text{(eq. 5)}$$

### 2.7 Meta-analysis

We collectively named relative risks (RR) for HR/RR/OR throughout our meta-
analysis. All literature-reported RRs were converted into adjusted incremental risk ratios
with a 10-ppbV O\textsubscript{3}-exposure increase by target metric (i.e. 6mDMA8 in this study),
following eq 6 as shown below:

$$RR_{\text{Adjusted}} = e^{\frac{\ln RR_{\text{Original}}}{k_{\text{Original} \rightarrow \text{Adjusted}}}} \quad \text{(eq. 6)}$$
where $ln$ is the natural logarithm, $RR_{\text{original}}$ is the originally reported risk estimates scaled into 10-ppbV incremental exposure, and $k_{\text{original-Adjusted}}$ is the conversion factor for metric unification. Multi-study pooled risks with 95% confidence interval (CI) were calculated from the adjusted RR s by Hunter-Schmidt random effect meta-regression estimator to correct the potential errors and biases caused from the diversity of study population and methodologies.\textsuperscript{35}

We applied the Higgins $I^2$ to quantify the heterogeneity across studies. The Higgins statistics $I^2$ is defined as

$$I^2 = \frac{Q - df}{Q} \times 100\%$$  \hspace{1cm} (eq. 6)

where $Q$ is the Cochran’s non-parametric heterogeneity statistic assessing whether there are any cross-study differences in risks based on $\chi^2$ distribution and $df$ is the corresponding degrees of freedom.\textsuperscript{36} Low $I^2$ values indicate no important heterogeneity observed and high $I^2$ values, especially $>75\%$, indicate considerable heterogeneity.

Subgroup analyses were conducted by grouping the selected studies upon the gender, regions, $O_3$ exposure metrics, and methodological reliability of individual exposure assignment; together with the adjustment of ethnicity, body mass index (BMI), smoking history, lifestyle features, and exposure to PM$_{2.5}$ and NO$_2$. Subgroups should contain at least 3 studies. Leave-one-out sensitivity analyses were also accomplished to test the robustness of synthesised overall risks by meta-analysis. All meta-analyses were performed in R 4.1.1 with packages meta, metafor, and metainf.

The most widely recognised approach to construct the integrated exposure-response (IER)$^\text{37}$ relationships required sufficient epidemiological studies to comprehensively sample the population exposure levels. However, studies on long-term $O_3$ exposure health effects were relatively limited, under which circumstance we made methodological modifications to make better use of the variabilities in exposure levels by statistically imputing the exposure distributions for each study from the provided statistics (e.g. mean, standard deviation, and percentiles) for curve fitting as elaborated in Supplementary Text S1. Supplementary Text S2 described the detailed procedures of exposure distribution imputations with a demonstration provided in S3, through which high uncertainties were still observed in the fitted IER curves due to insufficient epidemiological evidences.
3. RESULTS

3.1 Study characters

From the 3 databases during September 2015 till February 2022, a total of 339 studies (77 from MEDLINE, 102 from Embase, and 160 from Web of Science) were searched; and together with 34 additional literatures added manually from the 2 previous systematic reviews,15, 16 373 studies underwent duplication censoring, deleting 101 duplicated studies. After detailed scrutinisation for 272 studies, a total of 25 studies concerning long-term O₃ exposure and multi-cause mortalities were finally enrolled for quality evaluation, meta-analysis and further discussions (Figure 1).38-62 Table 1 summarised the basic information of the 25 included studies sorted by the publication year and surname of the first author.

3.2 Metrics and exposure assignments

Our updated systematic review stressed more on the exposure metrics and methodologies to obtain O₃ exposure, as summarised in Table 2. Abbey et al. (1999),38 Jerrett et al. (2013)46 and Lipsett et al. (2011)43 did not state the metric they used clearly, but based on comparisons between the reported surface O₃ concentrations and TOAR observational archives, we reasonably assumed ADA24 for the first study, and ADMA8 for the rest two. Details of the metric matching were given in Supplementary Text (S4). Lipfert et al. (2006)39 used the highest 95th percentile by hourly resolved O₃ concentrations as the peak exposure metric, which was only used in this one singular study, and hence approximated to DMA1. Krewski et al. (2009)41 and Smith et al. (2009)42 were both studies on ACS CPS II, and thus the same exposure assignment methodologies and metrics were assumed as Jerrett et al. (2009).40 Likewise, Cakmak et al. (2018)53 and Weichenthal et al. (2017)52 were assumed to inherit Crouse et al. (2015)48 as all these 3 studies were on CANCHEC. Warm season was defined as 6 months from April to September in terms of the northern hemisphere by default, but we made no exceptions to 3 studies as Zanobetti et al. (2011)44 using May to September, and Crouse et al. (2015)48 and Paul et al. (2020)57 using May to October, due to limited number of studies searched.

Across all included studies, multiple methods were applied to obtain gap-free surface O₃ concentrations for individual-level exposure assignment. The most basic way was the
nearest neighbour matching between participant residential locations and in situ observation sites, which were more frequently used in earlier studies. A comparatively more complicated way was statistical spatial interpolation, by inverse distance weighting or ordinary kriging. Full spatial coverage products, such as satellite-based remote-sensing and chemistry transport models, were used in some studies by supervised-learning-based data fusion techniques including but not limited to universal kriging embedded land use regression, Bayesian hierarchical model and ensemble learning to enhance the spatial extrapolation accuracy, which were evaluated to be of higher credibility than the basic ones described previously. All basic interpolation methods using merely the observations were rated as “Low”; applying chemical transport model simulations without calibration from the observations as “Moderate”; linearly coupling the observations with simulations as “Good”; and multi-source data assimilation by means of more sophisticated approaches as “High”. To sum up, 8 studies were rated “High”, 5 were “Good”, 2 were “Moderate”, and 10 were “Low”. Methodological progresses with time were evident as manifested in Table 2, prefiguring an explosion of population-based environmental health studies in the age of big data.

Based on the TOAR and CNEMC in situ observations, the cross-metric non-intercept linear conversion factors were estimated with regression accuracies given in Figure 2. Synthesised from the recent relevant studies, the 6mDMA8 metric was more recommended to highlight the peak exposure; and therefore, we chose to convert the originally reported RRs uniformly into the 6mDMA8 scale as standard. The O₃ exposure levels by the original and unified metric were listed in Supplementary Text S1. Demonstrations for the conversion interpretation and procedures were presented in Supplementary Text S5, respectively.

3.3 Meta-analysis results

We conducted meta-analyses for long-term O₃ exposure-associated into 10 categories of mortalities as (1) all causes (AC), (2) all respiratory diseases (RESP), (3) chronic obstructive pulmonary diseases and allied conditions (COPD), (4) all cardiovascular diseases (CVD), (5) all cerebrovascular diseases (CEVD), (6) ischaemic heart disease (IHD), (7) congestive heart failure (CHF), and (8) lung cancer (LC), with the exposure metrics adjusted into 6mDMA8.
3.3.1 All-cause mortality

A total of 23 studies were included into O₃ exposure-associated all-cause mortality meta-analysis, pooling the overall risk into RR = 1.014 (95% CI: 1.009–1.019, I²: 97.8%) with every 10-ppbV incremental exposure by 6mDMA8 as presented in Figure 3. Subgroup meta-analysis by originally reported metrics concluded the significances of risks vary across metrics, as high-concentration highlighted metrics like 6mDMA8 were of the highest positive risk (RR = 1.022, 95% CI: 1.014–1.031) while the smoothed metric ADA24 reported negative association (RR = 0.980, 95% CI: 0.960–1.001), as shown in Figure S1. Grouped by study regions, significant discrepancies of the risk pattern was found (Figure S2), as the studies in North America revealed positive associations as RR = 1.019 (95% CI: 1.014–1.024), while researches on European populations showed reversed risks as RR = 0.910 (95% CI: 0.827–1.001), though not significant. The cross-region divergences did not necessarily indicate differences in population vulnerability, as (1) less and younger study population, (2) shorter follow-up durations, and (3) use of smoothed exposure metrics for studies in Europe could all potentially obscure the potential risk associations. Subgroup analysis manifested that high inter-study heterogeneities originated from metric inconsistency, methodological reliability of individual exposure assignment, and population variabilities, as encapsulated in Table S3. The funnel plot was visually symmetrical (Figure S3), and studies reporting risks below the pooled value were even slightly more, indicating no detected severe potential publication biases.

No significant inter-gender differences were observed based on the limited studies reporting gender-specific risk association strengths. Further subgroup analyses were unfeasible due to the lack of reporting in the literature. Alternatively, grouped RRs were estimated based on whether the original researches had adjusted the confounding effects from ethnicity, body mass index (BMI), smoking history, lifestyle features, exposure levels of PM₂.₅ and NO₂, and no inter-group divergences were observed (Table S3).

3.3.2 Respiratory mortality

Meta-analysis for O₃ exposure-associated all respiratory mortality includes 16 studies, pooling which gave the overall RR = 1.025 (95% CI: 1.010–1.040, I²: 83.9%) for every 10-ppbV incremental O₃ exposure by 6mDMA8 (Figure 4). Based on sub-group meta-analysis for different metrics (Figure S4), peak metrics showed more significant increasing risks than ADA24, where most of the heterogeneities were from (I² = 87.0%). Cross-metric divergences were generally higher than intra-metric discrepancies. Studies
on North America populations showed better homogeneity in positive risks (RR = 1.029, 95% CI: 1.011–1.047, I² = 71.1%, Figure S5) than the European cohorts, pooling from which the overall risks were congruously insignificant (RR = 0.941, 95% CI: 0.856–1.036, I² = 91.2%). For O₃-COPD mortality association, the pooled RR was 1.056 (95% CI: 1.029–1.084, I² = 94.5%) for 10-ppbV incremental O₃ exposure by 6mDMA8 from 7 studies. No apparent positive publication biases were detected for both respiratory and COPD mortalities from the funnel plot (Figure S3).

### 3.3.3 Cardiovascular mortality

A total of 15 studies were included to pool the overall O₃ exposure-induced CVD mortality risks as RR = 1.019 (95% CI: 1.004–1.035, I² = 97.7%) for each 10-ppbV additional O₃ exposure by 6mDMA8 (Figure 5). Given the fact that the lower bound of uncertainty interval was so close to the null hypothesis (i.e. RR = 1), the positive risk association found in this review could be controversial, and thus would require more studies to support or refute the finding. Heterogeneities (I² > 79.2%) were observed through all 3 metric-grouped studies as presented in Figure S6. Positive risk associations were found on 10 North American cohorts (RR = 1.036, 95% CI: 1.019–1.053) while oppositely for 5 European cohorts (RR = 0.934, 95% CI: 0.865–1.008), as shown in Figure S7. There was no need to be concerned with the publication bias, and no more inter-group divergences were spotted except for grouping by exposure assignment methodological credibility (Table S3). The pooled risk for the congestive heart failure-induced mortality from 4 studies was RR = 1.074 (95% CI: 1.054–1.093, I² = 85.8%) with every 10-ppbV incremental O₃ exposure by 6mDMA8.

### 3.3.4 Other mortality causes

The other cause-specific mortality risks attributable to long-term O₃ exposure were not statistically significant (Figure 6), as IHD mortality risk pooled from 10 studies was RR = 1.012 (95% CI: 0.987–1.039, I² = 98.7%), CEVD mortality risk pooled from 6 studies was RR = 0.993 (95% CI: 0.979–1.008, I² = 80.6%), and LC mortality risk pooled from 13 studies was RR = 0.966 (95% CI: 0.926–1.007, I² = 84.2%). For all 8 studied mortality causes, we also provided pooled risks by 3 more widely used metrics (i.e. 6mDA24, ADMA8, and ADA24) besides 6mDMA8, as listed in Table 3 for reference.
3.4 Study assessment

All 25 studies included into our final meta-analysis were rated to be above “Fair” (14 as “Fair” and 11 as “Good”) by the Quality Assessment Tool for Observational Cohort Studies, as listed in Table S4. All studies well met 10 out of 14 assessment items, while 9 studies did not sufficiently clarify the participant exclusion criteria; 2 re-analysis study reports did not clearly state the O₃ exposures; 14, 42 2 studies were of relatively insufficient follow-up durations (e.g. less than 5 years) to observe the outcomes resulting from long-term exposure; 45, 63 and 10 studies were of methodological deficiencies in individual exposure assignment, 38-46, 55 most of which were conducted before 2013 when data assimilation techniques were not maturely developed to fuse observations with other full spatial coverage products such as satellite-based remote sensing and atmospheric mechanistic simulations. The satisfactory assessment results overall indicated inconspicuous risks of bias, laying the reliable foundation for meta-analyses.

Tables S5 displayed GRADE epidemiological evidence assessment results for each mortality cause from all involved corresponding studies. In brief, the overall judgements for all-cause, respiratory, cardiovascular, ischaemic heart disease, congestive heart failure, and lung cancer mortality risks were “High”, while the rating for the rest 2 cause-specific mortality risks (COPD and cerebrovascular diseases) were both “Moderate”. Inconsistence of the risk directions (i.e. positive or negative associations) was the most common reason for downgrading, except for the CHF-induced mortality. There were 6 additional risks, as an assessment upgrading item for the pooled RRs of all-cause, respiratory and cardiovascular mortality. Cakmak et al. (2018) spotted higher RRs after adjusting the confounder compared to the crude values, 53 which gave prominence to the positive risk associations and thus correspondingly upgraded the rating for all-cause, ischaemic heart disease, and lung cancer mortalities. No substantial positive publication biases were found based on the collected evidences.

3.5 Sensitivity analysis

Leave-one-out sensitivity analyses showed stable risk estimates as summarised in Table S6, except for the lung cancer mortality risks after eliminating Kazemiparkouhi et al. (2019), the only study reporting positive risk association, 55 while the rest 11 studies concluded insignificant risks or even protective effects. Since the metric harmonisation in our study was an innovative attempt, we provided both metric-adjusted and unadjusted
crude results for reference as presented in Table 3. The crude results were pooled from the originally reported relative risk values only unified into per 10-ppbV incremental exposure, without being transformed into any metrics for congruity. Along with the meta-analyses on all qualified studies, the relative risks were also pooled by keeping only one latest study with the largest population for each separate cohort, as summarised in Table S7. Under this circumstance, the pooled unit incremental mortality risks with every 10-ppbV incremental O₃ exposure by 6mMDA8 metric were RR = 1.008 (95% CI: 1.006–1.009, I² = 82.6%) for all causes, RR = 1.034 (95% CI: 1.017–1.050, I² = 81.7%) for all respiratory diseases, RR = 1.060 (95% CI: 1.040–1.080, I² = 90.2%) for COPD, RR = 1.032 (95% CI: 1.010–1.055, I² = 98.2%) for all cardiovascular diseases, RR = 1.008 (95% CI: 0.973–1.045, I² = 99.2%) for ischaemic heart disease, and RR = 0.966 (95% CI: 0.931–1.002, I² = 83.8%) for lung cancer. Studies for mortality risks of cerebrovascular diseases and congestive heart failure were respectively conducted on different cohorts, and hence such supplementary analysis was unnecessary.

4. DISCUSSION

4.1 Improvements as an updated review

This work improves on 2 previous high-quality reviews¹⁵,¹⁶ by covering up-to-date peer-reviewed studies, and expanding the O₃-exposure associated causes of mortality into wider range of categories. It is the first systematic review of the association between long-term O₃ exposure and cause-specific mortality highlighting the issue of inconsistent use of exposure metrics to our best knowledge. Since tropospheric O₃ is a photochemical pollutant which largely depends on solar radiation, the surface O₃ concentrations can vary drastically between day and night, as well as warmer and cooler seasons. We pointed out that a 10-ppbV increase in annual daily 24-hour average concentration (ADA24) is more constrained in magnitude than a 10-ppbV increase in warm-season daily 8-hour maximum average concentration (6mDMA8) owing to the wider variability in the range of the latter metric. Taking the observations by TOAR and CNEMC in situ monitoring networks during 1990-2019 as an example, the surface O₃ concentrations were 27.6 ± 6.1 (IQR: 24.1–31.0) ppbV by ADA24, while correspondingly 53.1 ± 10.6 (IQR: 47.7–61.4) ppbV by 6mDMA8, which indicated a 10-ppbV change fell below the IQR by the 6mDMA8, but could exceed the IQR using the ADA24 metric. This was why we believe
adjusting the exposure metrics was necessary for O\textsubscript{3} exposure-attributable health risk meta-analysis.

We also put forward a feasible approach to mutually convert the O\textsubscript{3} exposure concentrations and corresponding risk strengths in various metrics by non-intercept linear projections following the operational suggestions from EPA\textsuperscript{,32} but update the linear conversion factors based on global in situ surface O\textsubscript{3} observations during 1990-2019. The methodological innovation took advantage of multi-dimensional information from the original studies, which could inspire further observation collections and researches for corroborations and improvements.

### 4.2 Metric relevant issues

Although linear coefficients were applied onto the cross-metric conversions, irreducible noises still existed given the high root mean squared errors (RMSE) as shown in Figure 2, which exposed the limitation of risk strength adjustment into the same exposure metric by simple linear conversion, as the actual cross-metric relationships could be way more complicated. However, there was no other way but using the linear conversion coefficients as surrogates to unify the RRs by different metric reported in original studies, and thus to avoid uncertainties brought by the conversion of metrics, using a promissory consistent exposure metric or estimating the unit excess RRs in multiple metrics would be highly advocated in future long-term O\textsubscript{3}-exposure epidemiology studies.

Such linear conversion of risk associations could be validated by Kazemiparkouhi et al. (2020)\textsuperscript{,55} where multiple metrics were applied to estimate the mortality risks. For COPD mortality, the RR was 1.072 (95% CI: 1.067–1.077) by 6mDMA\textsubscript{1} for every 10-ppb\textsubscript{V} additional exposure; and after converting into 6mDMA\textsubscript{8} metric using the linear coefficient 0.831 (Figure 2), the estimated RR was 1.087 (95% CI: 1.081–1.093), close to the literature reported 1.084 (95% CI: 1.079–1.089),\textsuperscript{55} which justified our linear conversion method. Converting Cross-metric linear conversions would not change the risk association direction, but using different exposure metrics when estimating the O\textsubscript{3}-exposure attributable mortality risks could potentially cause discrepancies. For an instance, Kazemiparkouhi et al. (2020) concluded excess hazards of long-term O\textsubscript{3} exposure on all-cause mortality using 6mDMA\textsubscript{1} and 6mDMA\textsubscript{8} as quantitative metrics, but 6mDA\textsubscript{24} led to a specious prevention effect (RR = 0.990, 95% CI: 0.988–0.991), which should be attributed to the existence of a theoretical exposure safety level for O\textsubscript{3} below which no negative health effects should occur. Under this circumstance, lower-level metrics (e.g. ADA\textsubscript{24}) by averaging the peak O\textsubscript{3} exposures might obscure the
effective doses above the threshold, and also reduce the signal-to-noise ratios, so that
were of lower credibility in recognising hazardous population exposures than higher-level
metrics (e.g. 6mDMA8).

Data mining techniques were able to realise high-accuracy predictions of surface \( \text{O}_3 \)
concentrations, but errors were never avoidable. Carey et al. (2013) used a basic IDW
spatial interpolation approach to obtain the surface \( \text{O}_3 \) concentrations where the \( R^2 \) were
0.24–0.76,\(^{45}\) while years later Di et al. (2017) applied an ensemble learning approach,
achieving \( R^2 = 0.80, \) \( \text{RMSE} = 2.91 \) ppbV.\(^{51}\) Carey et al. (2013) reported the IQR of \( \text{O}_3 \)
exposure concentrations as 3.0 ppbV, which was comparable to the RMSE of Di et al.
(2017).\(^{51}\) Besides, lower \( R^2 \) could be accompanied with higher prediction errors, which
might have concealed the highest and lowest quartiles, and led to failures in
distinguishing the population-level exposures. This concern had been reflected in our
subgroup meta-analysis by exposure metrics, that lower-level metrics were more inclined
to report insignificant risks, which also cast sceptics on the reliability of studies covering
narrow exposure variabilities. We therefore are in favour of the Lancet suggestions to use
peak metrics to quantify the long-term \( \text{O}_3 \) exposure such as 6mDMA8, and also speak
highly of the state-of-the-art data techniques to reduce errors in \( \text{O}_3 \) concentration
prediction, so as to make a distinction between the high- and low-exposure populations.

### 4.3 Pathogenesis supports

Atkinson et al. (2016) concluded insignificant pooled risks for long-term \( \text{O}_3 \)
exposure associated all-cause and respiratory mortality,\(^ {16} \) which contradicted our results.
It should mainly be ascribed to the heterogeneity between the more recent studies and
earlier ones. The majority of studies collected in Atkinson et al. (2016) applied primitive
statistical methods (i.e. nearest neighbourhood matching, IDW and ordinary kriging
interpolation) for individual exposure assignment, which might have weakened the
individual-level exposure distinguishment. In addition, some studies using ADA24 as the
exposure metric could have also obscured the significance of associations.\(^ {43, 45, 46} \) In
contrast, studies after 2016 more frequently applied advanced numerical simulation
models and data assimilation techniques to increase the precision of population exposure
assessment; and most of them used 6mDMA8 metric to foreground the high exposures.\(^ {50,}
52, 53, 55-57\) These recent studies stuck out the significant \( \text{O}_3 \)-mortality associations.

From another aspect, pathological mechanisms had been at least partially ascertained
by laboratorial experiments. The inhaled \( \text{O}_3 \) could constrict the muscles in the airways
leading to shortness of breath, and damage the lining with inflammation.\(^ {64} \) Long-term \( \text{O}_3 \)
exposure could increase the oxidative stress in the cardiovascular system,\(^6^5\) and cause progressive thickening of the carotid arteries to restrict cerebral blood supply.\(^6^6\) Additionally, strong associations had already been found between short-term O\(_3\) exposure and a variety of cardiopulmonary symptoms as reported by a number of observational epidemiological studies,\(^6\) which also supported the long-term exposure effects, as it was unreasonable to presume no incremental risks by long-term exposure given the verified significant short-term effects. We thus were inclined to approve of the opinion that long-term O\(_3\) exposure would increase mortality risks in agreement with GBD report.\(^2^0\)

To alleviate the population health loss resulting from O\(_3\) exposure, the U.S. EPA appealed for optimisations in real-time accessibility of air quality index, with which residents could be able to avoid unnecessary high pollution exposure (https://www.epa.gov/ground-level-ozone-pollution/health-effects-ozone-pollution). Appropriate diets and supplements including carotenoids, vitamin D and vitamin E were recognised to be preventive against air pollution induced respiratory damages, which was a practical protective measure for the vulnerable.\(^6^7\)

### 4.4 Concentration-response relationship

Few studies had examined the concentration-response curves between long-term O\(_3\) exposure and mortality, and thus the threshold exposure level (also known as theoretical minimum risk exposure level, TMREL) below which no adverse health effects would be assumed to occur was still controversial. For all-cause mortality, Di et al. (2017) reported a safe exposure level as 30 ppbV by 6mDA24 metric (approximately as 49.9 ppbV by 6mDMA8),\(^5^1\) while Shi et al. (2022) suggested a lower level as 40 ppbV by 6mDMA8, both estimated from the Medicare beneficiary cohort.\(^5^8\) For respiratory mortality, Jerrett et al. (2009) tested the concentration-response relationships and estimated the threshold level as 60 ppbV by 6mDMA1 (49.9 ppbV by 6mDMA8),\(^4^0\) while Lim et al. (2019) failed to identify a significant threshold level.\(^5^6\) For cardiovascular mortality, Lim et al. (2019) showed no apparent health hazards below 45 ppbV by 6mDMA8,\(^5^6\) and Paul et al. (2020) prescribed a threshold level around 35 ppbV by 6mDMA8 metric for diabetic patients.\(^5^7\) These evidence-based threshold exposure levels were all no higher than the current standards, as 70 ppbV for daily maximum 8-hour exposure under NAAQS (The National Ambient Air Quality Standards regulated by the U.S. EPA)\(^6^8\) and 50 ppbV by warm-season DMA8 under WHO global air quality guidelines.\(^6^9\) However, whether the standard guidelines should be revised to be more strictly would require more further studies.
To synthesise epidemiological evidences, Burnett et al. (2014) developed an integrated exposure-response (IER) function-based curve-fitting method to pool the risk associations from multiple studies. We attempted to construct the IER for long-term O\textsubscript{3} exposure associated mortalities in this review, with statistically reproduced exposure levels to enhance the curve fitting, as illustrated in Supplementary Texts S1-S3. The exposure imputing had revealed high reliability, but the high uncertainties of the IER curves could still not be addressed, which should be attributed to the limited effective epidemiological evidences. Empirically, this approach would require sufficient studies to cover a wide range of exposure levels, which had been frequently adopted for particulate matter exposure researches, but seldomly used for O\textsubscript{3}-health studies. The main probable reason might be that the population long-term O\textsubscript{3} exposure levels would not be as comparably distinguishable to the particulate matters. In addition, a reasonable prescribed TMREL would be necessary to establish the IER curves, and hence the indeterminacy of the threshold level could exacerbate the uncertainties in the estimated concentration-response relationships. Therefore, more relevant studies on long-term O\textsubscript{3} exposure associated risks are urgently appealed for, based on which discussions, optimisations, or corrections on our enhanced exposure-response curve-fitting methodologies, are encouraged.

### 4.5 Hierarchical classification of diseases

The causes of mortalities analysed in our study followed hierarchical subordinate relationships, as the all-cause mortality consisted of cardiovascular diseases, respiratory diseases, cancer and other causes; chronic obstructive pulmonary disease belonged to respiratory category; and ischaemic heart disease, stroke, congestive heart failure and other cerebrovascular diseases all subordinated to cardiovascular symptoms. On this occasion, estimating all O\textsubscript{3}-exposure induced mortalities could follow a bottom-up scheme by adding up subgroups of diseases. However, for the historical O\textsubscript{3}-associated mortalities, GBD attributed all O\textsubscript{3}-associated mortalities onto COPD-induced premature deaths, which we thought were of spaces for further explorations. Long-term O\textsubscript{3} exposure had shown significant association with excess cardiovascular mortalities, and thus we should update the mortality estimations in further studies by including CVD altogether into consideration.
4.6 Application in mortality estimations

The widest applications of the estimated risk association strengths were to project how many people would be affected by long-term ambient O₃ exposure. For example, Malley et al. (2017) estimated 1.23 (95% UI: 0.85–1.62) million respiratory deaths attributable to O₃ exposure in 2010, using the risk strength by Turner et al. (2016) as HR = 1.12 (95% UI: 1.08–1.16). This estimation was much higher than the 2019 GBD report: 0.31 (95% UI: 0.15–0.49) million, as had been highlighted in another recent study, which should be attributed to the use of high HR value among all included studies. We had also found some other studies using one singular HR value for population risk estimations, but we would still encourage further relevant studies to consider multi-study pooled RRs, which could effectively reduce the potential biases from a single study. The adaptability of the pooled RRs could be verified from the coverage of exposure levels, as the 25 studies identified in our review had embraced a wide range of exposure concentrations (Supplementary Text S1) to encompass the global surface O₃ variability. On the other hand, the leave-one-out sensitivity analyses (Table S6) had revealed the robustness of the meta-analysis results when including sufficient numbers of studies, which was a circumstantial reflection for the representativeness of the synthesised risk association strengths. The annual GBD reports were also based on the generalisability presumption of the synthesised epidemiological evidences, but cohort-based researches in the unstudied regions are always appealed for to provide more convincing discoveries.

4.7 Limitations

Although the total number of studied participants for risk pooling was adequately high to ensure the statistical power, the cohort-based O₃-health studies were factually rare according to our literature search, and thus long-term follow-up studies are urgently encouraged. Additionally, current literatures seldomly reported grouped RRs, which made meta-analyses by sub-categories (like gender, age, socio-economic status, smoking and alcohol history, etc.) unfeasible. Scarcity of credible evidences also restricted the effects of conventional approaches to construct exposure-response curves, and our methodological innovation would require further relevant studies for substantiation. The cross-metric linear conversion factors were estimated relying on observations from available sites, which however might not be sufficiently representative of the global residential areas, as observational sites in India, Africa, and Latin America were still sparse. With ever-increasing popularisation of the in situ monitoring networks, the cross-
metric conversion factors might need calibration with more comprehensive observations, so that the pooled RRs should also be updated accordingly.

4.8 Further study suggestions

We suggest that further environmental epidemiology studies, especially long-term O₃ exposure related researches, clearly report i) the methodologies to obtain ambient O₃ concentrations, the spatiotemporal resolution, and prediction accuracy of the database; ii) the exposure metrics used for risk estimation; and iii) the statistical distribution of the O₃ exposure concentrations. The data-oriented methodologies to accomplish full spatial coverage ambient air O₃ concentrations for individual-level exposure assignment should be transparent as the construction credibility of air pollution concentration database should also be rigorously assessed, which were the foundation of epidemiological follow-up studies. We would advocate the report of exposure metrics in future O₃-health studies so as to avoid confusions when comparing the risks with literature and conducting meta-regression; and according to the recent consensus, warm-season average (6mDMA8) shall be preferred as epidemiological study metrics.¹⁹ We recommend future studies estimate risks with multiple O₃ metrics for reference; and describing the statistical distribution of the O₃ exposure levels is another suggested element to assess the reliability of risk estimation models, which can also be useful in exposure-response tendency exploration. We also propose future cohort studies estimate subgroup-specific RRs which can be conducive to identify the vulnerable populations.

Our review highlights a deficiency existing in current environmental health research literatures, that studies on long-term O₃ exposure health effects are still rather rare compared to particulate matter-based studies.⁷⁸ Also, the meta-analysis results might be geographically-biased, since large-scale O₃ exposure health risk studies till 2022 did not cover Asia, Africa or Latin America regions. However, there are some thriving cohorts such as the Multi-Country Multi-City (MCC) Collaborative Research Network covering over 22 countries or regions,⁷⁹ and the China Kadoorie Biobank (CKB) enrolling over 0.5 million people,⁸⁰ working on environmental exposure projects. We are optimistic that more research will come out to fill in the literature gap of multi-region population-based studies.
5. CONCLUSION

Our state-of-the-science systematic review has summarised cohort studies exploring the associations between long-term ambient O₃ exposure and multi-cause mortality risks. Current studies support O₃-exposure attributable additional mortalities caused from all causes, respiratory diseases, chronic obstructive pulmonary disease, cardiovascular diseases, and congestive heart failure, but no significant mortality risks are found for ischaemic heart diseases, all-type cerebrovascular diseases, and lung cancer. Exposure metrics are crucial for health risk estimations of long-term O₃ exposure and meta-analysis to pool the multi-study risks, for which we develop a linear conversion approach to harmonise the different metrics. Further researches on long-term O₃ observations and exposure-induced mortalities are encouraged to corroborate or contradict our linear conversion factors and meta-analysis results by providing more solid evidences, so as to strengthen the O₃-health literatures.
Competing Interest Statement

All authors declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; and no other relationships or activities that could appear to have influenced the submitted work.

Author Contributions

ATA and YG conceived the idea for the review; HZS, CL and PY performed the literature search; HZS and PY conducted statistical analyses; HZS, ATA, YG, PY, SH, JM, HS and LY contributed to discussions; HZS wrote the article; and MW examined the languages. HZS is the guarantor who accepted the full responsibility for the finished article, owned full access to all relevant data, and controlled the decision to publish. The two joint corresponding authors, ATA and YG, attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted. The guarantor, HZS, affirms that the manuscript is an honest, accurate, and transparent account of the study being reported, and no important aspects of the study have been omitted. No ethical approval is needed for a systematic review and meta-analysis.

Acknowledgement

This study is funded by UK Natural Environment Research Council (NERC), UK National Centre for Atmospheric Science (NCAS), Australian Research Council (DP210102076) and Australian National Health and Medical Research Council (APP2000581). HZS, MW, and SH receive funding from Engineering and Physical Sciences Research Council (EPSRC) via the UK Research and Innovation (UKRI) Centre for Doctoral Training in Application of Artificial Intelligence to the study of Environmental Risks (AI4ER, EP/S022961/1). ATA acknowledges funding from NERC (NE/P016383/1) and through the Met Office UKRI Clean Air Programme. YG is supported by a Career Development Fellowship of the Australian National Health and Medical Research Council (APP1163693). All contents in this study are solely the responsibility of the grantees and do not represent the official views of the supporting agencies.
Special appreciations to Dr Xiao Lu (School of Atmospheric Sciences, Sun Yat-sen University) for his insightful discussion on the quality control of TOAR products, Dr Liuhua Shi (Rollins School of Public Health, Emory University) for her supplementary information on Medicare beneficiary cohort information, and 4 anonymous reviewers together with the editor for their meticulous efforts in improving the manuscript.

Data Availability

The surface O₃ observations are archived in the Tropospheric Ozone Assessment Report (TOAR, https://b2share.fz-juelich.de/communities/TOAR) repository and China National Environmental Monitoring Centre (CNEMC, http://www.cnemc.cn/en/) repository, which are accessible to the public. The cohort-based long-term O₃ exposure-associated cause-specific mortality risks are all available in the main text or supplementary materials of the selected studies.

Supplementary Information

Further detailed information can be found in the Supplementary Materials (PDF) consisting of 25 pages with 5 sections of texts as “imputation procedures for exposure distribution” (S1), “enhanced integrated exposure-response curve-fitting” (S2), “demonstrative procedures of enhanced exposure-response trend curve-fitting” (S3), “undefined metric imputation” (S4), and “interpretation and procedure of cross-metric linear conversion” (S5), together with 6 tables and 7 figures to strengthen the results and discussions presented in the main text. A PRISMA Checklist (PDF) was provided to verify the integrity of this study.
## TABLES

### Table 1 Summary of cohort characteristics included for meta-analysis.

<table>
<thead>
<tr>
<th>Study</th>
<th>Cohort</th>
<th>Country</th>
<th>Follow-up Duration</th>
<th>Population Type</th>
<th>Sample Size</th>
<th>Sex</th>
<th>Age</th>
<th>Key Confounding Adjustment</th>
<th>Mortality Causes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abbey et al. 1999&lt;sup&gt;38&lt;/sup&gt;</td>
<td>AHS</td>
<td>USA</td>
<td>1977-1992</td>
<td>Occupational</td>
<td>6,182</td>
<td>FM</td>
<td>27-95</td>
<td>age, sex, BMI, smoking, individual demographic features&lt;sup&gt;3&lt;/sup&gt;, lifestyle features&lt;sup&gt;4&lt;/sup&gt;, medical history</td>
<td>AC, RESP, LC</td>
</tr>
<tr>
<td>Lipfert et al. 2006&lt;sup&gt;39&lt;/sup&gt;</td>
<td>WU-EPRI</td>
<td>USA</td>
<td>1976-1996</td>
<td>General</td>
<td>67,108</td>
<td>M</td>
<td>51 (12)&lt;sup&gt;3&lt;/sup&gt;</td>
<td>age, ethnicity, BMI, smoking, traffic density, NO&lt;sub&gt;2&lt;/sub&gt;, CO</td>
<td>AC</td>
</tr>
<tr>
<td>Jerrett et al. 2009&lt;sup&gt;40&lt;/sup&gt;</td>
<td>ACS CPS II</td>
<td>USA</td>
<td>1977-2000</td>
<td>General</td>
<td>448,850</td>
<td>FM</td>
<td></td>
<td>age, sex, ethnicity, BMI, smoking, individual demographic features, lifestyle features</td>
<td>AC, RESP, CVD</td>
</tr>
<tr>
<td>Krewski et al. 2009&lt;sup&gt;41&lt;/sup&gt;</td>
<td>ACS CPS II</td>
<td>USA</td>
<td>1982-2000</td>
<td>General</td>
<td>488,370</td>
<td>FM</td>
<td>≥ 30</td>
<td>age, sex, ethnicity, BMI, smoking, individual demographic features, lifestyle features, PM&lt;sub&gt;2.5&lt;/sub&gt;</td>
<td>AC, IHD, LC</td>
</tr>
<tr>
<td>Smith et al. 2009&lt;sup&gt;42&lt;/sup&gt;</td>
<td>ACS CPS II</td>
<td>USA</td>
<td>1982-2000</td>
<td>General</td>
<td>352,242</td>
<td>FM</td>
<td></td>
<td>age, ethnicity, BMI, smoking, lifestyle features, medical treatment</td>
<td>AC, RESP, CVD, IHD, CHF, CEVD, LC</td>
</tr>
<tr>
<td>Lipsett et al. 2011&lt;sup&gt;43&lt;/sup&gt;</td>
<td>CTS</td>
<td>USA</td>
<td>1998-2005</td>
<td>Occupational</td>
<td>124,614</td>
<td>F</td>
<td>≥ 20</td>
<td>age, sex, ethnicity, medical history</td>
<td>COPD, CHF</td>
</tr>
<tr>
<td>Zanobetti et al. 2011&lt;sup&gt;44&lt;/sup&gt;</td>
<td>Medicare</td>
<td>USA</td>
<td>1985-2006</td>
<td>General</td>
<td>8,894,473</td>
<td>FM</td>
<td>≥ 65</td>
<td>age, sex, BMI, smoking, individual demographic features</td>
<td>AC, RESP, LC</td>
</tr>
<tr>
<td>Carey et al. 2013&lt;sup&gt;45&lt;/sup&gt;</td>
<td>CPRD</td>
<td>UK</td>
<td>2003-2007</td>
<td>General</td>
<td>824,654</td>
<td>FM</td>
<td>40-89</td>
<td>age, sex, smoking, individual demographic features</td>
<td>AC, RESP, LC</td>
</tr>
<tr>
<td>Jerrett et al. 2013&lt;sup&gt;46&lt;/sup&gt;</td>
<td>ACS CPS II</td>
<td>USA</td>
<td>1982-2000</td>
<td>General</td>
<td>73,711</td>
<td>FM</td>
<td>57 (11)</td>
<td>age, sex, smoking, individual demographic features, lifestyle features</td>
<td>AC, RESP, CVD, IHD, LC</td>
</tr>
<tr>
<td>Bentayeb et al. 2015&lt;sup&gt;47&lt;/sup&gt;</td>
<td>GAZEL</td>
<td>France</td>
<td>1989-2013</td>
<td>Occupational</td>
<td>20,327</td>
<td>FM</td>
<td>44 (4)</td>
<td>age, sex, BMI, smoking, individual demographic features, lifestyle features</td>
<td>AC, RESP, CVD</td>
</tr>
<tr>
<td>Crouse et al. 2015&lt;sup&gt;48&lt;/sup&gt;</td>
<td>CANCHEC</td>
<td>Canada</td>
<td>1991-2006</td>
<td>General</td>
<td>2,521,525</td>
<td>FM</td>
<td>≥ 25</td>
<td>age, sex, individual and area-level demographic features, PM&lt;sub&gt;2.5&lt;/sub&gt;, NO&lt;sub&gt;2&lt;/sub&gt;</td>
<td>AC, RESP, COPD, CVD, IHD, CEVD, LC</td>
</tr>
<tr>
<td>Tonne et al. 2016&lt;sup&gt;49&lt;/sup&gt;</td>
<td>MINAP</td>
<td>UK</td>
<td>2003-2010</td>
<td>MI Survivors&lt;sup&gt;6&lt;/sup&gt;</td>
<td>18,138</td>
<td>FM</td>
<td>68 (14)</td>
<td>age, sex, ethnicity, smoking, medical history, area-level demographic features</td>
<td>AC</td>
</tr>
<tr>
<td>Turner et al. 2016&lt;sup&gt;50&lt;/sup&gt;</td>
<td>ACS CPS II</td>
<td>USA</td>
<td>1982-2004</td>
<td>General</td>
<td>669,046</td>
<td>FM</td>
<td>≥ 30</td>
<td>age, sex, BMI, smoking, individual and area-level demographic features, PM&lt;sub&gt;2.5&lt;/sub&gt;, NO&lt;sub&gt;2&lt;/sub&gt;</td>
<td>AC, RESP, COPD, CVD, CHF, IHD, CEVD</td>
</tr>
<tr>
<td>Di et al. 2017&lt;sup&gt;51&lt;/sup&gt;</td>
<td>Medicare</td>
<td>USA</td>
<td>2000-2012</td>
<td>General</td>
<td>60,925,443</td>
<td>FM</td>
<td>≥ 65</td>
<td>age, sex, ethnicity, BMI, smoking, individual and area-level demographic features, meteorological features, PM&lt;sub&gt;2.5&lt;/sub&gt;</td>
<td>AC</td>
</tr>
<tr>
<td>Weichenthal et al. 2017&lt;sup&gt;52&lt;/sup&gt;</td>
<td>CANCHEC</td>
<td>Canada</td>
<td>2001-2011</td>
<td>General</td>
<td>2,448,500</td>
<td>FM</td>
<td>25-89</td>
<td>age, sex, ethnicity, individual and area-level demographic features</td>
<td>AC, RESP, CVD</td>
</tr>
<tr>
<td>Cakmak et al. 2018&lt;sup&gt;53&lt;/sup&gt;</td>
<td>CANCHEC</td>
<td>Canada</td>
<td>1991-2011</td>
<td>General</td>
<td>2,291,250</td>
<td>FM</td>
<td>≥ 25</td>
<td>age, sex, individual demographic features, PM&lt;sub&gt;2.5&lt;/sub&gt;</td>
<td>AC, COPD, IHD, LC</td>
</tr>
<tr>
<td>Hvidtfeldt et al. 2019&lt;sup&gt;54&lt;/sup&gt;</td>
<td>DDCH</td>
<td>Denmark</td>
<td>1993-1997</td>
<td>General</td>
<td>49,596</td>
<td>FM</td>
<td>50-64</td>
<td>age, sex, BMI, smoking, individual and area-level demographic features, noise</td>
<td>AC, RESP, CVD</td>
</tr>
<tr>
<td>Kazemiparkouhi et al. 2019&lt;sup&gt;55&lt;/sup&gt;</td>
<td>Medicare</td>
<td>USA</td>
<td>2000-2008</td>
<td>General</td>
<td>22,159,190</td>
<td>FM</td>
<td>≥ 65</td>
<td>age, sex, ethnicity, area-level demographic features, PM&lt;sub&gt;2.5&lt;/sub&gt;</td>
<td>AC, RESP, COPD, CVD, IHD, CHF, CEVD, LC</td>
</tr>
<tr>
<td>Study</td>
<td>Cohort</td>
<td>Country</td>
<td>Follow-up Duration</td>
<td>Population Type</td>
<td>Sample Size</td>
<td>Sex</td>
<td>Age</td>
<td>Key Confounding Adjustment</td>
<td>Mortality Causes</td>
</tr>
<tr>
<td>------------------------</td>
<td>--------------</td>
<td>------------------</td>
<td>--------------------</td>
<td>-----------------</td>
<td>-------------</td>
<td>-----</td>
<td>-----</td>
<td>-------------------------------------------------------------------------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Lim et al. 201916</td>
<td>NIH-AARP</td>
<td>USA</td>
<td>1995-2011</td>
<td>General</td>
<td>548,780</td>
<td>FM</td>
<td>50-71</td>
<td>age, sex, ethnicity, BMI, smoking, individual demographic features, PM$_{2.5}$, NO$_2$, daily maximum temperature</td>
<td>AC, RESP, COPD, CVD, IHD, CHF, CEVD</td>
</tr>
<tr>
<td>Paul et al. 202047</td>
<td>ONPHEC</td>
<td>Canada</td>
<td>1996-2015</td>
<td>Diabetes</td>
<td>452,590</td>
<td>FM</td>
<td>35-85</td>
<td>age, sex, area-level demographic features, smoking, individual and area-level demographic features</td>
<td>CVD</td>
</tr>
<tr>
<td>Shi et al. 202118</td>
<td>Medicare</td>
<td>USA</td>
<td>2001-2017</td>
<td>General</td>
<td>44,684,756</td>
<td>FM</td>
<td>≥ 65</td>
<td>age, sex, ethnicity, BMI, smoking, individual and area-level demographic features, lifestyle features, PM$_{2.5}$, NO$_2$, medical history</td>
<td>AC</td>
</tr>
<tr>
<td>Strak et al. 202119</td>
<td>ELAPSE</td>
<td>6 countries¹</td>
<td>1985-2015</td>
<td>General</td>
<td>325,367</td>
<td>FM</td>
<td>49 (13)</td>
<td>age, sex, ethnicity, BMI, smoking, individual and area-level demographic features, PM$_{2.5}$, NO$_2$, BC</td>
<td>AC, RESP, COPD, CVD, IHD, CEVD</td>
</tr>
<tr>
<td>Yazdi et al. 202160</td>
<td>Medicare</td>
<td>USA</td>
<td>2000-2016</td>
<td>General</td>
<td>44,430,747</td>
<td>FM</td>
<td>≥ 65</td>
<td>age, sex, ethnicity, BMI, smoking, individual and area-level demographic features, lifestyle features, PM$_{2.5}$, NO$_2$, medical history</td>
<td>AC</td>
</tr>
<tr>
<td>Bauwelinc et al. 202261</td>
<td>BC2001</td>
<td>Belgium</td>
<td>2001-2011</td>
<td>General</td>
<td>5,474,470</td>
<td>FM</td>
<td>≥ 30</td>
<td>age, sex, individual and area-level demographic features, PM$_{2.5}$, NO$_2$, BC</td>
<td>AC, RESP, CVD, LC</td>
</tr>
<tr>
<td>Stafoggia et al. 202252</td>
<td>ELAPSE</td>
<td>7 countries²</td>
<td>2000-2017</td>
<td>General</td>
<td>28,153,138</td>
<td>FM</td>
<td>≥ 30</td>
<td>age, sex, ethnicity, BMI, smoking, individual and area-level demographic features, PM$_{2.5}$, NO$_2$, BC</td>
<td>AC, RESP, CVD, LC</td>
</tr>
</tbody>
</table>

Cohort abbreviations: AHSMOG, Adventist Health Study of Smog; WU-EPRI, Washington University–Electric Power Research Institute; ACS CPS, American Cancer Society Cancer Prevention Study; CTS, California Teacher Study; CPRD, Clinical Practice Research Datalink; GAZEL, GAZ de France and Électricité; CANCHEC, Canadian Census Health and Environment Cohort; MINAP, National Audit of Myocardial Infarction Project; DDCH, Danish Diet, Cancer and Health; NIH-AARP, National Institute of Health, American Association of Retired Persons; ONPHEC, Ontario Population Health and Environment Cohort; BC2001, Belgian 2001 Census.

1. Demographic features included marital status, education attainment, employment status and occupational class, aboriginal ancestry, visible minority ethnicities, immigrant status and residence location (urban or rural), income level and socioeconomic status (SES), and regional population density. Different studies adjusted various combinations of demographic features.

2. Lifestyle features included consumptions of alcohol, dietary fat, vegetables (dietary fibre), and vitamins, together with physical activity frequency. Different studies adjusted various combinations of lifestyle features.

3. Population ages were reported by mean with standard deviation (in bracket).

4. MI, Myocardial Infarction.

5. Sweden, Denmark, France, Netherlands, Germany and Austria.

6. Belgium, Denmark, England, Netherlands, Norway, Switzerland and Italy.
### Table 2 Data sources and statistical methods of O₃ exposure assignment.
Methodological ratings were based on spatial interpolation and multi-data assimilation approaches. Spatial resolutions, exposure metrics, and levels of incremental risk ratio were also listed.

<table>
<thead>
<tr>
<th>Study</th>
<th>Data Sources</th>
<th>Methods</th>
<th>Resolution</th>
<th>Rating</th>
<th>Metrics</th>
<th>Level of incremental risk ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abbey et al. 1999¹³</td>
<td>monitoring station observations</td>
<td>IDW interpolation</td>
<td>NR¹</td>
<td>Low</td>
<td>ADMA8</td>
<td>12.03 ppbV</td>
</tr>
<tr>
<td>Lipfert et al. 2006³⁹</td>
<td>monitoring station observations</td>
<td>nearest matching (assumed)²</td>
<td>NR</td>
<td>Low</td>
<td>ADMA1</td>
<td>40 ppbV</td>
</tr>
<tr>
<td>Jerrett et al. 2009⁴⁰</td>
<td>monitoring station observations</td>
<td>nearest matching (assumed)</td>
<td>NR</td>
<td>Low</td>
<td>6mDMA1</td>
<td>10 ppbV</td>
</tr>
<tr>
<td>Krewski et al. 2009⁴¹</td>
<td>monitoring station observations</td>
<td>ordinary kriging interpolation</td>
<td>NR</td>
<td>Low</td>
<td>6mDMA1</td>
<td>10 ppbV</td>
</tr>
<tr>
<td>Smith et al. 2009⁴²</td>
<td>monitoring station observations</td>
<td>nearest matching (assumed)</td>
<td>NR</td>
<td>Low</td>
<td>6mDMA1</td>
<td>1 µg/m³</td>
</tr>
<tr>
<td>Lipsatt et al. 2011⁴³</td>
<td>monitoring station observations</td>
<td>IDW interpolation</td>
<td>250 m</td>
<td>Low</td>
<td>ADA24</td>
<td>22.96 ppbV</td>
</tr>
<tr>
<td>Zanobetti et al. 2011⁴⁴</td>
<td>monitoring station observations</td>
<td>nearest matching (assumed)</td>
<td>NR</td>
<td>Low</td>
<td>6mDMA8</td>
<td>5 ppbV</td>
</tr>
<tr>
<td>Carey et al. 2013⁴⁵</td>
<td>monitoring station observations</td>
<td>interpolation (IDW assumed)</td>
<td>1 km</td>
<td>Low</td>
<td>ADA24</td>
<td>3.0 µg/m³</td>
</tr>
<tr>
<td>Jerrett et al. 2013⁴⁶</td>
<td>monitoring station observations</td>
<td>IDW interpolation</td>
<td>NA</td>
<td>Low</td>
<td>ADA24</td>
<td>24.1782 ppbV</td>
</tr>
<tr>
<td>Bentayeb et al. 2015⁴⁷</td>
<td>monitoring station observations, model simulation, other auxiliary predictors</td>
<td>universal kriging embedded land use regression</td>
<td>2 km</td>
<td>Good</td>
<td>6mDMA8</td>
<td>12.3 µg/m³</td>
</tr>
<tr>
<td>Crouse et al. 2015⁴⁸</td>
<td>monitoring station observations, model simulation</td>
<td>linear data assimilation</td>
<td>21 km</td>
<td>Good</td>
<td>6mDMA8</td>
<td>9.5 ppbV</td>
</tr>
<tr>
<td>Tonne et al. 2016⁴⁹</td>
<td>KCLurban air dispersion model simulation</td>
<td>NA</td>
<td>20 m</td>
<td>Moderate</td>
<td>ADA24</td>
<td>5.3 µg/m³</td>
</tr>
<tr>
<td>Turner et al. 2016⁵⁰</td>
<td>monitoring station observations, CMAQ model simulation</td>
<td>hierarchical Bayesian space-time data assimilation</td>
<td>12 km</td>
<td>High</td>
<td>ADMA8, 6mDMA8</td>
<td>10 ppbV</td>
</tr>
<tr>
<td>Di et al. 2017⁵¹</td>
<td>monitoring station observations, model simulation, satellite remote sensing observations, other auxiliary predictors</td>
<td>ensemble machine learning</td>
<td>1 km</td>
<td>High</td>
<td>6mDMA8</td>
<td>10 ppbV</td>
</tr>
<tr>
<td>Weichenthal et al. 2017⁵²</td>
<td>monitoring station observations, model simulation</td>
<td>linear data assimilation</td>
<td>21 km</td>
<td>Good</td>
<td>6mDMA8</td>
<td>10.503 ppbV</td>
</tr>
<tr>
<td>Cakmak et al. 2018⁵³</td>
<td>monitoring station observations, model simulation</td>
<td>linear data assimilation</td>
<td>21 km</td>
<td>Good</td>
<td>6mDMA8</td>
<td>10 ppbV</td>
</tr>
<tr>
<td>Hvidtfeldt et al. 2019⁵⁴</td>
<td>AirGIS dispersion model simulation</td>
<td>NA</td>
<td>1 km</td>
<td>Moderate</td>
<td>ADA24</td>
<td>10 µg/m³</td>
</tr>
<tr>
<td>Kazemiparkouhi et al. 2019⁵⁵</td>
<td>monitoring station observations</td>
<td>nearest matching (assumed)</td>
<td>6 km</td>
<td>Low</td>
<td>6mDMA1, 6mDMA8, 6mDA24</td>
<td>10 ppbV</td>
</tr>
<tr>
<td>Lim et al. 2019⁵⁶</td>
<td>monitoring station observations, CMAQ model simulation</td>
<td>Bayesian space-time downscaling</td>
<td>12 km</td>
<td>High</td>
<td>6mDMA8</td>
<td>10 ppbV</td>
</tr>
<tr>
<td>Paul et al. 2020⁵⁷</td>
<td>monitoring station observations, model simulation</td>
<td>linear data assimilation</td>
<td>21 km</td>
<td>Good</td>
<td>6mDMA8</td>
<td>6.4 ppbV</td>
</tr>
<tr>
<td>Shi et al. 2021⁵⁸</td>
<td>monitoring station observations, model simulation, satellite remote sensing observations, other auxiliary predictors</td>
<td>ensemble machine learning</td>
<td>1 km</td>
<td>High</td>
<td>6mDMA8</td>
<td>10 ppbV</td>
</tr>
<tr>
<td>Strak et al. 2021⁵⁹</td>
<td>monitoring station observations, model simulation, satellite remote sensing observations, other auxiliary predictors</td>
<td>universal kriging embedded land use regression</td>
<td>100 m</td>
<td>High</td>
<td>6mDMA8</td>
<td>10 µg/m³</td>
</tr>
<tr>
<td>Study</td>
<td>Methodology</td>
<td>Spatial Resolution</td>
<td>Air Quality</td>
<td>Concentration Unit</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>--------------------</td>
<td>-------------</td>
<td>--------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yazdi et al. 2021</td>
<td>Ensemble machine learning</td>
<td>1 km</td>
<td>High</td>
<td>6mDMA8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bauwelinc et al. 2022</td>
<td>Land use regression</td>
<td>100 m</td>
<td>High</td>
<td>6mDMA8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stafoggia et al. 2022</td>
<td>Universal kriging embedded land use regression</td>
<td>100 m</td>
<td>High</td>
<td>6mDMA8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

† NR, not reported.
‡ The statistical methods were not clearly stated in literatures, and thus the most basic method was assumed. The nearest neighbourhood matching shall be the simplest way to assign spatially sparse observations onto cohort participants, and the inverse distance weighting (IDW) is the simplest spatial interpolation approach.

NA, not applicable. The chemical transport model simulations were directly used for individual exposure assignment without further statistical processing.
Table 3 Pooled RRs for long-term 10-ppbV incremental O₃-exposure attributable multi-cause mortalities by 4 most widely used metrics and crude risks without metric harmonisation.

<table>
<thead>
<tr>
<th>Mortality causes</th>
<th>6mDMA8</th>
<th>6mDA24</th>
<th>ADMA8</th>
<th>ADA24</th>
<th>Crude</th>
</tr>
</thead>
<tbody>
<tr>
<td>All causes (n = 23)</td>
<td>1.014 (1.009, 1.019)</td>
<td>1.023 (1.014, 1.032)</td>
<td>1.016 (1.010, 1.022)</td>
<td>1.027 (1.017, 1.037)</td>
<td>1.017 (1.011, 1.023)</td>
</tr>
<tr>
<td>Respiratory diseases (n = 16)</td>
<td>1.025 (1.010, 1.040)</td>
<td>1.042 (1.016, 1.069)</td>
<td>1.029 (1.011, 1.047)</td>
<td>1.049 (1.019, 1.081)</td>
<td>1.031 (1.017, 1.046)</td>
</tr>
<tr>
<td>Chronic obstructive pulmonary disease (n = 7)</td>
<td>1.056 (1.029, 1.084)</td>
<td>1.098 (1.050, 1.149)</td>
<td>1.066 (1.034, 1.098)</td>
<td>1.116 (1.058, 1.176)</td>
<td>1.055 (1.032, 1.078)</td>
</tr>
<tr>
<td>Cardiovascular diseases (n = 15)</td>
<td>1.019 (1.004, 1.035)</td>
<td>1.033 (1.006, 1.061)</td>
<td>1.022 (1.004, 1.041)</td>
<td>1.038 (1.007, 1.071)</td>
<td>1.024 (1.009, 1.038)</td>
</tr>
<tr>
<td>Ischaemic heart disease (n = 10)</td>
<td>1.012 (0.987, 1.039)</td>
<td>1.021 (0.977, 1.067)</td>
<td>1.014 (0.984, 1.045)</td>
<td>1.024 (0.973, 1.078)</td>
<td>1.017 (0.994, 1.041)</td>
</tr>
<tr>
<td>Congestive heart failure (n = 4)</td>
<td>1.074 (1.054, 1.093)</td>
<td>1.130 (1.094, 1.168)</td>
<td>1.086 (1.063, 1.110)</td>
<td>1.155 (1.110, 1.198)</td>
<td>1.083 (1.059, 1.107)</td>
</tr>
<tr>
<td>Cerebrovascular diseases (n = 6)</td>
<td>0.993 (0.979, 1.008)</td>
<td>0.988 (0.964, 1.013)</td>
<td>0.992 (0.976, 1.009)</td>
<td>0.986 (0.958, 1.015)</td>
<td>0.992 (0.979, 1.006)</td>
</tr>
<tr>
<td>Lung cancer (n = 12)</td>
<td>0.966 (0.926, 1.007)</td>
<td>0.943 (0.878, 1.012)</td>
<td>0.960 (0.915, 1.008)</td>
<td>0.933 (0.859, 1.014)</td>
<td>0.960 (0.909, 1.013)</td>
</tr>
</tbody>
</table>
FIGURES

Figure 1 Schematic flowchart of study assessment and selection processes for literature review and meta-analysis.
Figure 2 Cross-metric linear relationships and conversion accuracies. The cross-metric linear relationships were scaled by Pearson’s correlation coefficients. The cross-metric conversion factors with 95% confidence intervals (95% CI) were estimated by non-intercept linear regression models, accompanied with fitting accuracies quantified by coefficient of determination ($R^2$) and root mean square error (RMSE) in ppbV. The conversion factors were defined as multiples from the original metric by column into the target harmonised metric by row, e.g. ADMA8 = 1.671 ADA24, $R^2 = 0.9736$, RMSE = 7.78 ppbV. Note that by non-intercept linear regression, the values of $R^2$ should no longer be equal to the squared Pearson’s linear correlation coefficients. As the cross-metric conversion coefficients were estimated statistically, indirect conversions were not recommended, since regression noises restricted the validity of equation $k_{A\rightarrow B} = k_{A\rightarrow C} \cdot k_{C\rightarrow B}$.
**Figure 3** Pooled estimates of all-cause mortality risk associated with every 10-ppbV incremental O\(_3\) exposure by 6mDMA8 metric. Size of the shaded squares in the forest plot represents the weight of each study estimated by random-effect model.
Figure 4 Pooled estimates of respiratory diseases and COPD mortality risks associated with every 10-ppbV incremental O₃ exposure by 6mDMA8 metric.
Figure 5 Pooled estimates of cardiovascular diseases and congestive heart failure mortality risk associated with every 10-ppbV incremental O₃ exposure by 6mDMA8 metric.
Figure 6 Pooled estimates of ischaemic heart disease, cerebrovascular diseases, lung cancer, ischaemic stroke, and pneumonia mortality risks associated with every 10-ppbV incremental O₃ exposure by 6mDMA8 metric.
REFERENCES


68. EPA Integrated Science Assessment (ISA) for Particulate Matter; Washington, D.C., 2019.


national public health burden associated with exposure to ambient PM$_{2.5}$ and ozone. Risk Anal 2012, 32, (1), 81-95.


