Ecological studies of COVID-19 and air pollution: How useful are they?

Paul J. Villeneuve[a,*], Mark S. Goldberg[b–e]

Background: Results from ecological studies have suggested that air pollution increases the risk of developing and dying from COVID-19. Drawing causal inferences from the measures of association reported in ecological studies is fraught with challenges given biases arising from an outcome whose ascertainment is incomplete, varies by region, time, and across sociodemographic characteristics, and cannot account for clustering or within-area heterogeneity. Through a series of analyses, we illustrate the dangers of using ecological studies to assess whether ambient air pollution increases the risk of dying from, or transmitting, COVID-19.
Methods: We performed an ecological analysis in the continental United States using county-level ambient concentrations of fine particulate matter ($PM_{2.5}$) between 2000 and 2016 and cumulative COVID-19 mortality counts through June 2020, December 2020, and April 2021. To show that spurious associations can be obtained in ecological data, we modeled the association between $PM_{2.5}$ and the prevalence of human immunodeficiency virus (HIV). We fitted negative binomial models, with a logarithmic offset for county-specific population, to these data. Natural cubic splines were used to describe the shape of the exposure-response curves.
Results: Our analyses revealed that the shape of the exposure-response curve between $PM_{2.5}$ and COVID-19 changed substantially over time. Analyses of COVID-19 mortality through June 30, 2021, suggested a positive linear relationship. In contrast, an inverse pattern was observed using county-level concentrations of $PM_{2.5}$ and the prevalence of HIV.
Conclusions: Our analyses indicated that ecological analyses are prone to showing spurious relationships between ambient air pollution and mortality from COVID-19 as well as the prevalence of HIV. We discuss the many potential biases inherent in any ecological-based analysis of air pollution and COVID-19.

Keywords: COVID-19; HIV; Air pollution; Ecological studies; Cross-level bias

Introduction

The emergence of the 2019 global pandemic caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has profoundly impacted global health. As of December 19, 2021, official statistics show more than 274 million cases diagnosed worldwide and 5.4 million deaths,[1] but the true impact is substantially higher given that underascertainment of both incidence and mortality data have been well documented.[2–5] As well, the development of the postacute COVID-19 syndrome ("long-COVID"), which likely represents a wide spectrum of conditions, affects a substantial number of people for long periods of time,[6] with the prevalence of persistent symptoms (>60 days) estimated from a systematic review to be about 70%.[7] Justifiably so, there has been a rush to develop methods to screen and diagnose COVID-19, develop efficacious vaccines and treatment regimens, and implement a battery of public health measures to reduce spread. These public health measures vary by jurisdiction and have oscillated over time as multiple waves of the pandemic have come and gone.

It is abundantly clear that this highly infectious virus, especially the new variants, such as Delta and now Omicron, is transmitted from person to person, with many indoor spaces providing a highly efficient means of transmission especially due to contaminated aerosols.[8] Thus, public health measures have focused largely on limiting transmission, from wearing of masks, distancing protocols, to curfews and lockdowns, as well as vaccination campaigns. Despite the well-appreciated dynamics of transmission, there has been a search for other factors that may increase in incidence or affect sequelae such as mortality,

*a*School of Mathematics and Statistics, Carleton University, Ottawa, ON, Canada; *b*Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, QC, Canada; *c*Department of Medicine, McGill University, Montreal, Canada; *d*Gerald Bronfman Department of Oncology, McGill University, Montreal, Canada; and *e*Centre for Outcomes Research and Evaluation, Research Institute of the McGill University Hospital Centre, Montreal, Canada.

What this study adds

Findings from ecological studies suggest that exposure to ambient air pollution increases the incidence or mortality from COVID-19. We have previously described some of the biases of these studies. Herein, we modeled ecological county-level data of ambient concentrations of $PM_{2.5}$, COVID-19, and the prevalence of HIV. We observed that associations between ambient $PM_{2.5}$ and COVID-19 mortality vary substantially over time, and that associations with prevalent HIV are inversely related to ambient $PM_{2.5}$. Our analyses show that ecological data can provide misleading associations and should not be used for drawing causal inferences between ambient air pollution and COVID-19 outcomes.

including vitamin D supplementation,[9] urban noise,[10] warm weather,[11] proximity to urban greenness,[12] solar ultraviolet radiation,[11,13] and selenium deficiency.[14]

Given our past research in air pollution, we were drawn to a series of studies that reported associations between ambient air pollution and an increased risk of transmission of COVID-19, and more severe outcomes including mortality. We published a detailed commentary about many of the challenges of making causal inferences between exposure to air pollution and incidence and mortality of COVID-19[15] and recommended, in particular, that etiological studies of this topic use personal rather than area-wide information, but that because of limitations in identifying cases, these studies may also be unable to provide valid results. Since the publication of that commentary, a series of cohort studies have also been used to investigate associations between ambient air pollution and COVID-19.[16–20] The recent findings by Kogevinas et al,[20] which found no associations between air pollution and COVID-19 transmission, are of particular interest because they establish infection to the virus by measuring antibodies in blood samples. This approach may help avoid possible biases in case ascertainment.

Despite the publication of these individual-level longitudinal studies, ecological studies on this topic continue to be published.[21–32] We remain concerned especially about the use of ecological studies to inform on causality. As they rely on group-level data, this allows analyses to be carried out quickly and cheaply, and often by using data that can be extracted readily from administrative sources. It has been known for many decades that this study design has significant limitations for making inferences about causality, because it is uses grouped data, usually from administrative jurisdictions, rather than information on individuals. A key limitation of ecological studies is the inability to identify and account for heterogeneity within the geographical areas used as the unit of analysis.[33] The bias may be so severe as to completely reverse the direction of the association, such as shown in a parallel analyses of ecological and individual-level case-control data of radon and lung cancer.[34] The issue of cross-level bias,[35] as well as other serious problems associated with the ascertainment of incidence and mortality of COVID-19 and other methodological issues, have led us to repeatedly raise concerns about possible spurious findings from studies of air pollution and COVID-19.[36–38] Although many authors have acknowledged the limitation of ecological studies, their willingness to publish these studies suggest they are essentially accepting that these measures of association are valid and, thus, either implicitly or explicitly conveying that these measures represent causal associations.

As indicated earlier, ecological studies of COVID-19 are particularly problematic because of substantial underascertainment of health outcomes. The recent article by Kogevinas et al. found that 40% of individuals with COVID-19 were asymptomatic[20] suggesting that self-reported measures of COVID-19 or those based on hospitalization or death data introduce bias. The completeness of ascertainment is inextricably linked to social determinants of health, equity, and health care capacity, and these same factors have been shown to be associated with ambient concentrations of air pollution.[39] A recent American report, for example, found that COVID-19 infections and deaths had disproportionately affected Hispanic or Latino, and Black populations.[20] Complicating matters further is that ascertainment of outcomes of COVID-19 change over time and are dependent on public health policies that can vary both between and within jurisdictions, vaccination uptake, and effectiveness.[40] Moreover, newer variants differ from the original ones with respect to features of transmissibility and severity. For example, it has been suggested that the viral load for those with the Delta strain of COVID-19 is 1,000 times higher than previous variants,[41] and that SARS-CoV-2 vaccines are less effective.[42] The recent emergence of the Omicron variant is a stark reminder how quickly the transmissibility features of COVID-19 can change. For all of the above reasons and others, we hypothesized that it might would only be possible to identify environmental factors that may affect COVID-related incidence, mortality, and other comorbidities at a future date—but only if the data are of sufficiently high quality.[15]

Thus, the present commentary is motivated by the continuing publication of ecological studies of air pollution and outcomes of COVID-19 from many jurisdictions.[21–32] Using US county-level measures of mass concentrations of fine particulate matter (PM$_{2.5}$) and COVID-19 mortality, similar to that used by Wu et al[43] that spurred many of the studies mentioned above, the objective of the present article is to estimate the shape of the response function and investigate whether these functions vary by time. We emphasize that the goal of these analyses was not to replicate the results of Wu and colleagues[43] but rather illustrate key methodological aspects of these types of analyses and their interpretation. In addition, as our a priori assumption is that ecological associations with many area-wide variables are spurious, we also modeled the association between ambient air pollution and the prevalence of human immunodeficiency virus (HIV). We reasoned that there should not be a causal association for PM2.5 given that HIV is caused through sexual contact as well as contact with infected bodily fluids, such as blood.[44] Thus, a second objective was to determine, through a similar ecological design, the association between the prevalence of HIV and ambient concentrations of PM$_{2.5}$.

Methods

Data sources

COVID-19

Cumulative numbers of deaths attributed to COVID-19, at the county level, were downloaded from the Johns Hopkins University Coronavirus Resource Center (https://coronavirus.jhu.edu/map.html) through to June 2020, December 2020, and June 2021, and we conducted three separate analyses using cumulative COVID-19 mortality data for these three time periods.

Fine particulate data

Ambient concentrations of the mass of PM$_{2.5}$, in μg/m³, were estimated at a resolution of 0.01° × 0.01° for the period 2000–2016 using validated atmospheric chemistry and machine learning models.[45] We used publicly available county-level concentrations of PM$_{2.5}$ by Wu et al.,[46,47] averaged across the period 2000–2016.

HIV data

The prevalent number of HIV cases, by county, for 2018 was extracted from the US Centers for Disease Control website.[48]

Covariable data

We used the census data of each county that were obtained from the published datasets from the Wu et al. study.[43] These data included county-level population and socioeconomic and demographic variables from 2009 to 2016 US CDC Compressed Mortality Data (https://wonder.cdc.gov/cmf-ICD10.html). These data also included county-level risk factor data from the US CDC Behavioral Risk Factor Surveillance system (https://www.cdc.gov/brfss/). Finally, we appended to our dataset county-level data for vaccinations for the analyses covering the latter period (through June 30, 2021).

Statistical analyses

We conducted all analyses in R[49] using negative binomial regression models (function "glm.nb," MASS library[50]). The outcome variables were the number of deaths from COVID-19 or the number prevalent cases of HIV, and we included an offset using the natural logarithm of the total population of each county.

For COVID-19 mortality, we carried out three separate analyses using the cumulative number of deaths until June 30, 2020, December 31, 2020, and June 30, 2021. For each analysis, we estimated the functional forms of each variable separately using natural cubic spline functions and determined from visual inspection that three degrees of freedom (df) adequately described the patterns for each variable. We did not use goodness of fit statistics to determine the "optimal" shape of the response curve because likelihood ratio tests cannot be used for smoothers having different degrees or freedom and because they are not nested. In addition, statistics such as the Akaike information criterion require the same number of observations between models which was not the case because of missing data in covariables.

We constructed an adjusted model by entering $PM_{2.5}$ as the sole variable, and then added each covariable consecutively in the following order: smoking rate; population density; percent older population; hospital beds; median household income; percent of the population under the poverty line; percent of the population who were Black; mean summer temperature, mean body mass index; percent of the population who owned their own home; and percentages of the population who were Hispanic, Asian, and White. For the analyses of COVID-19 mortality until June 30, 2021, we also included a smoothing term for fully vaccinated individuals. For the analyses of HIV, we used the same covariables as above, but the outcome was county-specific prevalence of HIV for 2018. We show marginal predicted effects (in numbers of deaths) for $PM_{2.5}$ and other covariables using Lüdecke's R package sjPlot.[51]

Results

COVID-19 and $PM_{2.5}$

We obtained data for 3,224 counties that included all American states as well as Puerto Rico. Because of missing values on various covariables, the fully adjusted models comprised about 1,990 counties. Included in these missing data were 128 counties with missing concentrations of $PM_{2.5}$ in some counties in Alaska, Hawaii, South Dakota, Virginia, and Puerto Rico. The range of $PM_{2.5}$ was between 2.1 and 15.8 μg/m³.

The total number of counties with no recorded deaths were as follows: end of June 2020, 1,103; end of December 2020, 159; and end of April 2021, 120. The supplement shows the response functions for each of the covariables for cumulative mortality from COVID-19 until the end of April 2021.

As well, we show at the end of the supplement diagnostic plots for the full model. The pattern of residuals by predicted values is not uniformly distributed suggesting that the model assumptions do not hold well. We interpret this as due to difference in the structure of counties and the inherent heterogeneity within counties.

Figure 1 shows the fitted marginal response functions for $PM_{2.5}$ for the three periods of time adjusted for all covariables, and that accounts for the total population of each county using an offset on a logarithmic scale.

For the first two periods of time, nonlinear associations were found, exhibiting a sinusoidal pattern (Wu et al. previously published risk estimates based on a linear assumption), but the latter period showed a response consistent with a linear trend. Had we assumed linear functions for $PM_{2.5}$, the estimated rate ratios for a unit increase of $PM_{2.5}$ were: June 2020, 1.09; end of December 2020, 1.02; and end of April 2021, 1.04. The results

from Wu et al[43] for the period until middle of June showed an 11% increase per unit increase of $PM_{2.5}$.

HIV prevalence, voting practices, and $PM_{2.5}$

The exposure-response function between HIV prevalence in 2018 and $PM_{2.5}$ was modeled using data from 2,001 counties. This pattern was nonlinear and reduced rates of HIV were observed in counties with higher ambient concentrations of $PM_{2.5}$ (Figure 2).

Additional findings from sensitivity analysis are provided in the accompanying efile; http://links.lww.com/EE/A176.

Discussion

Our ecological analyses of COVID-19 mortality data at the US county-level showed that the functional form of the exposure-response function between ambient concentrations of $PM_{2.5}$ and COVID-19 mortality was variable over time, and was only consistent with a linear association with outcomes through the end of June 2021. Regardless of the shape of the exposure-response curves (i.e., linear or nonlinear) observed in our models, the limitations of ecological data preclude drawing any causal inferences. Indeed, we found an inverse association between the prevalence of HIV and $PM_{2.5}$. This example illustrates that using ecological data can generate spurious results, and in our view, this danger is exacerbated with an emerging disease like COVID-19 that is poorly ascertained.

In all the other previously published ecological studies of COVID-19 and air pollution, it was assumed that the association was linear. Our analyses using smoothers that showed highly nonlinear associations indicate that this assumption needs to be verified. Despite finding an increasing and almost monotonic relationship for the latter period this should not be interpreted as a causal association.[15] Indeed, adding in another 6 months of data as Delta, and now Omicron, would profoundly alter the observed number of deaths and likely the shape of the ecological exposure-reponse curve. This is particularly relevant given the changes in the key variants and emerging ones, vaccinations of populations using different vaccines with different levels of effectiveness that change with time,[52] breakthrough infections, and the like.

Bradford Hill, in his seminal 1965 article,[53] highlighted the value of biological plausibility when assessing causal associations. A key component in this assessment is the possibility of exposure, and for COVID-19, it is now recognized that the virus is spread through aerosol transmission.[8,54] As well, past studies reported associations between ambient air pollution and increased incidence of respiratory viral infections, transmissibility of viruses, and hospital visits.[55] It has also been suggested that air pollution facilitates the transport of viral particles[56–58] and other plausible pathways whereby air pollution can worsen case severity, including effects on the immune system[59] and inflammation. We have also highlighted in our previous editorial that exposure to air pollution increases the incidence of a number of health conditions, and individuals with these underlying comorbidities may have worsened prognosis following diagnosis of COVID-19.

Despite these hypothetical mechanisms, it would appear transmission of SARS-COV-2 rarely occurs in outdoor settings. Specifically, a study in Wuhan, China, showed that detailed contact tracing found that among the 7,324 COVID-19 cases only one outbreak, consisting of two cases, could be linked to transmission in an outdoor environment.[60] Elsewhere, the Health Protection Surveillance Centre in the Republic of Ireland reported that only 262 of the 232,164 cases of COVID through March 24, 2021, were linked to outdoor transmission.[61] Further support of indoor transmission comes from the now famous case of a Church choir in Skagit County, Washington,[62] as well as studies showing that concentrations of viral particles in ambient air was very low.[63] A systematic review found that the
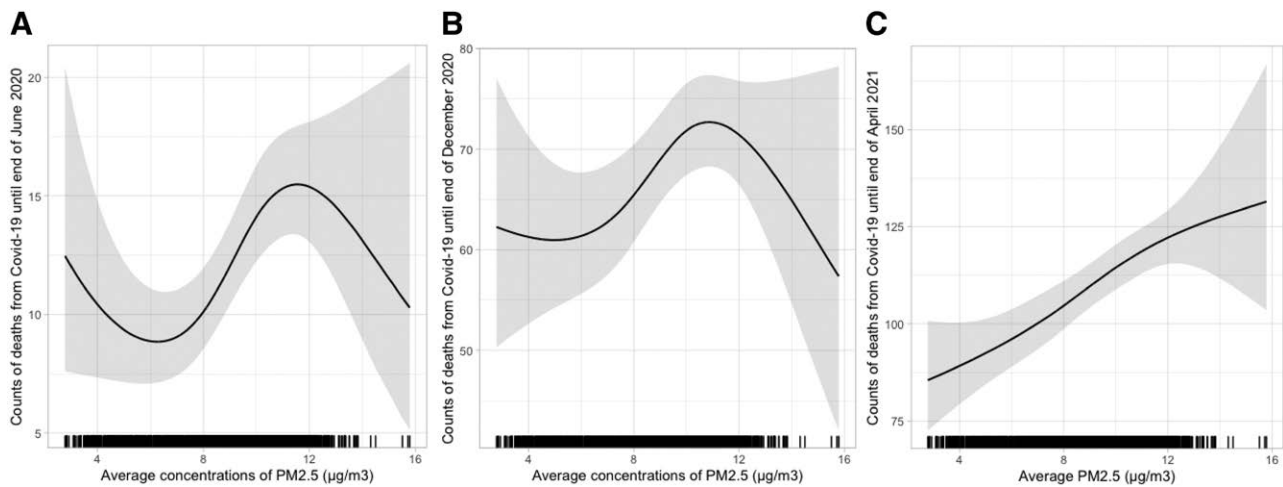
Figure 1. Adjusted exposure-response curves for county-level measures of PM$_{2.5}$ and cumulative COVID-19 mortality (solid lines) and 95% CIs (shaded areas) using natural cubic spline models with 3 df, for three different time periods. Models were adjusted, using natural cubic spline functions, for county-level measures of smoking, population density, older population, hospital beds, median household income, poverty, summer temperature, mean body mass index, and percentage of population that is White, Hispanic, Asian, White, and Black. For the June 2021 analysis, the model included the county-level proportion of fully vaccinated persons. Note the different scales for the y axes in the different plots. A, Cumulative deaths until end of June 2020; B, Cumulative deaths until end of December 2020; and C, Cumulative deaths until end of June 2021.

odds of contracting COVID-19 was close to 19 times higher for indoor versus outdoor transmission.[64] Previous studies that have assessed links between outdoor air pollution and COVID have not collected information on the source or location of infection of study participants.

Limitations of the data and analysis

Other limitations of the ecological data we modeled include missing data, where we found that many counties had considerable missing data in some covariables, and we dealt with this problem by simply excluding those counties with missing data.

We did not find that the negative binomial models fit the data well even after adjusting for many covariables. In fact, we adjusted for several possible "risk factors," particularly social determinants of health, that are likely related to both air pollution and COVID-19, including income, race, access to health care, and adherence to public health recommendations, such as physical distancing and wearing face masks. More importantly, however, unlike individual studies in which causal diagrams, such as directed acyclic graphs, can be used to create a theoretical model, we know of no way in an ecological analysis to determine a plausible causal model. This is due to the nature of causality as it is defined in epidemiology: "… a cause of a specific disease event as an antecedent event, condition, or characteristic that was necessary for the occurrence of the disease at the moment it occurred, given that other conditions are fixed."[65] A causal diagram could be defined from first principles but its application in an ecological study becomes problematic if one considers the presence of cross-level bias that cannot be accounted for as well as that counter-factual arguments cannot be applied as they are based on risk in individuals and not in heterogeneous groups.

Regardless of this lack of a theoretical model and the issue of missing data, Wakefield[33] showed that drawing causal inferences in ecological data are impossible unless there is no within-area heterogeneity. Briefly, he described that the within-area distributions of exposure, covariables, and outcome can be described by three frequencies, and he showed that the marginal prevalence of these risk variables, as available in ecological analyses, cannot be used to characterize their joint distribution unless the confounder and exposure of interest are independent. Wakefield noted that without individual-level data, this assumption cannot

be assessed. Thus, accurately defining sets of covariables to estimate causal associations to block backdoor paths cannot be made from ecological data, and thus causal inference methods are inappropriate.

Our concerns about the use of ecological studies have been expressed by other authors.[66] Ecological studies are easy to carry out, but the issue of severe bias should be taken seriously before
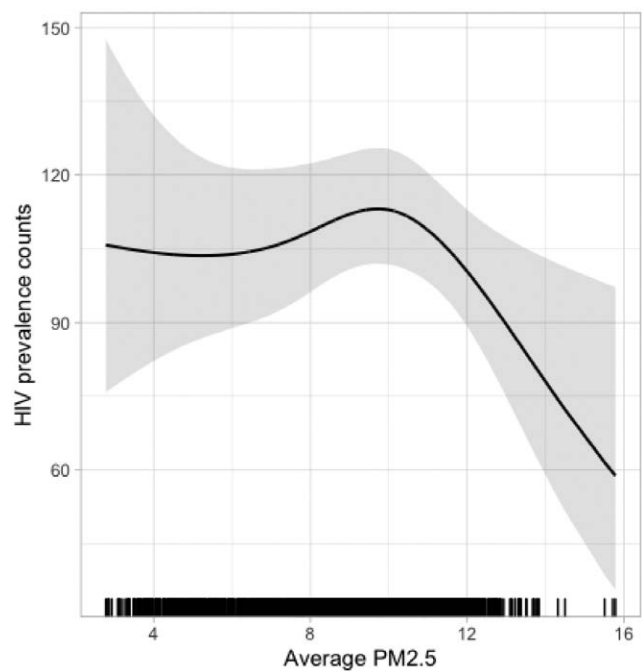

Figure 2. Adjusted exposure-response curves for PM$_{2.5}$ and the prevalence of HIV in 2018 (solid lines) and 95% CIs (shaded areas) using natural cubic spline models with 3 df, for three different time periods Adjusted, using natural cubic spline functions, for county-level measures of smoking, population density, older population, hospital beds, median household income, poverty, summer temperature, mean body mass index, percentage of population that is White, Hispanic, Asian, White, Black, and proportion of fully vaccinated persons added to model.

an investigation is initiated. Given the issues that we outlined in our review of COVID-19 and air pollution,[15] we reiterate, especially considering the possibility of nonsensical associations being generated, that studies of COVID-19 and any risk factor may be premature because of the serious biases inherent in the ecological data. The issues with data quality include under-ascertainment of incidence and mortality that are dependent on jurisdiction, not accounting for nonindependence of outcomes, inability to account for clustering in the data (after all, human-to-human transmission is a clustered process), using grouped data that cannot account for heterogeneity or confounding leading to cross-level bias, lack of highly spatially resolved air pollution measures, and not accounting for regional variations in the timing of outbreaks' temporal changes in at-risk populations.

In addition, clustering of cases occurring in a neighborhood can bias both point estimates of risk as well as the associated standard errors. Indeed, many social determinants of health that lead to strong spatial and temporal correlations are linked to the spread of COVID-19, including crowded living especially in lower income areas, race, community access to health care, employment (e.g., working multiple jobs), the occupational environment (e.g., meat packing plants), and in nursing homes where large clusters of cases have occurred. It is our view that these factors all contribute to spreading infections, and this is the point of view taken by all sensible public health measures that include physical distancing, isolation, contact tracing, wearing masks, improving indoor ventilation, and the like. As shown in the studies quoted above on transmission in outdoor air, and implicit in public health guidelines, outdoor air pollution is not a contributing factor and, thus, the positive associations from ecological studies are spurious.

Although additional ecological studies of air pollution continue to be published, new investigations are making use of other study designs, as well as using individual-level data. It is not our intent to provide a detailed critique of these studies within this article. Well-designed cohort studies can provide insights about whether air pollution impacts the transmission and severity of COVID-19. Indeed, the recent study by Kogevinas et al[20] has made important advances through the use of antibody measures, which avoid some of the possible biases in case ascertainment. Their study also had several other strengths including analyses of outcomes before vaccines became available, and a small geographical area with good surveillance data.[67] The lack of information on sources of infection, as in all previous studies, remains a glaring limitation. Along with these challenges the virus continues to mutate, and there are substantial complexities in modeling the impacts of vaccination for a risk factor (air pollution) that may be weakly associated with COVID-19 outcomes. This ultimately leads us to conclude that we are unlikely to definitively determine whether air pollution, or other environmental exposures, are causally related to COVID-19 outcomes.

Conflicts of interest statement

The authors declare that they have no conflicts of interest with regard to the content of this report.

References

1. Johns Hopkins University. COVID-19 Dashboard by the Center for Systems Science. 2021. Available at: https://coronavirus.jhu.edu/map.html Accessed July 16, 2021.
2. The Economist. Leaders section. Vaccinating the world. Ten million reasons to vaccinate the world. *The Economist*. 2021, Available at: https://www.economist.com/leaders/2021/05/15/ten-million-reasons-to-vaccinate-the-worl.
3. Angulo FJ, Finelli L, Swerdlow DL. Estimation of US SARS-CoV-2 infections, symptomatic infections, hospitalizations, and deaths using seroprevalence surveys. *JAMA Netw Open*. 2021;4:e2033706.
4. Lau H, Khosrawipour T, Kocbach P, Ichii H, Bania J, Khosrawipour V. Evaluating the massive underreporting and undertesting of COVID-19 cases in multiple global epicenters. *Pulmonology*. 2021;27:110–115.
5. Institute of Health Metrics and Evaluation. Estimation of total mortality due to COVID-19. 2021 Available at: http://www.healthdata.org/special-analysis/estimation-excess-mortality-due-covid-19-and-scalars-reported-covid-19-deaths. [Accessed 20 December 2021].
6. Nalbandian A, Sehgal K, Gupta A, et al. Post-acute COVID-19 syndrome. *Nat Med*. 2021;27:601–615.
7. Nasserie T, Hittle M, Goodman SN. Assessment of the frequency and variety of persistent symptoms among patients with COVID-19: a systematic review. *JAMA Netw Open*. 2021;4:e2111417.
8. Greenhalgh T, Jimenez JL, Prather KA, Tufekci Z, Fisman D, Schooley R. Ten scientific reasons in support of airborne transmission of SARS-CoV-2. *Lancet*. 2021;397:1603–1605.
9. Diaz-Curiel M, Cabello A, Arboiro-Pinel R, et al. The relationship between 25(OH) vitamin D levels and COVID-19 onset and disease course in Spanish patients. *J Steroid Biochem Mol Biol*. 2021;212:105928.
10. Díaz J, Antonio-López-Bueno J, Culqui D, Asensio C, Sánchez-Martínez G, Linares C. Does exposure to noise pollution influence the incidence and severity of COVID-19? *Environ Res*. 2021;195:110766.
11. Tobías A, Molina T, Rodrigo M, Saez M. Meteorological factors and incidence of COVID-19 during the first wave of the pandemic in Catalonia (Spain): a multi-county study. *One Health*. 2021;12:100239.
12. Klompmaker JO, Hart JE, Holland I, et al. County-level exposures to greenness and associations with COVID-19 incidence and mortality in the United States. *medRxiv*. Preprint posted online September 1, 2020. doi: 10.1101/2020.08.26.20181644.
13. Isaia G, Diémoz H, Maluta F, et al. Does solar ultraviolet radiation play a role in COVID-19 infection and deaths? An environmental ecological study in Italy. *Sci Total Environ*. 2021;757:143757.
14. Zhang HY, Zhang AR, Lu QB, et al. Association between fatality rate of COVID-19 and selenium deficiency in China. *BMC Infect Dis*. 2021;21:452.
15. Villeneuve PJ, Goldberg MS. Methodological considerations for epidemiological studies of air pollution and the SARS and COVID-19 coronavirus outbreaks. *Environ Health Perspect*. 2020;128:95001.
16. Bowe B, Xie Y, Gibson AK, et al. Ambient fine particulate matter air pollution and the risk of hospitalization among COVID-19 positive individuals: Cohort study. *Environ Int*. 2021;154:106564.
17. Mendy A, Wu X, Keller JL, et al. Long-term exposure to fine particulate matter and hospitalization in COVID-19 patients. *Respir Med*. 2021;178:106313.
18. López-Feldman A, Heres D, Marquez-Padilla F. Air pollution exposure and COVID-19: a look at mortality in Mexico City using individual-level data. *Sci Total Environ*. 2021;756:143929.
19. Elliott J, Bodinier B, Whitaker M, et al. COVID-19 mortality in the UK Biobank cohort: revisiting and evaluating risk factors. *Eur J Epidemiol*. 2021;36:299–309.
20. Kogevinas M, Castaño-Vinyals G, Karachaliou M, et al. Ambient air pollution in relation to SARS-CoV-2 infection, antibody response, and COVID-19 disease: a Cohort Study in Catalonia, Spain (COVICAT Study). *Environ Health Perspect*. 2021;129:117003.
21. De Angelis E, Renzetti S, Volta M, et al. COVID-19 incidence and mortality in Lombardy, Italy: An ecological study on the role of air pollution, meteorological factors, demographic and socioeconomic variables. *Environ Res*. 2021;195:110777.
22. Gujral H, Sinha A. Association between exposure to airborne pollutants and COVID-19 in Los Angeles, United States with ensemble-based dynamic emission model. *Environ Res*. 2021;194:110704.
23. Hassan MS, Bhuiyan MAH, Tareq F, Bodrud-Doza M, Tanu SM, Rabbani KA. Relationship between COVID-19 infection rates and air pollution, geo-meteorological, and social parameters. *Environ Monit Assess*. 2021;193:29.
24. Lu B, Wu N, Jiang J, Li X. Associations of acute exposure to airborne pollutants with COVID-19 infection: evidence from China. *Environ Sci Pollut Res Int*. 2021;28:50554–50564.
25. Norouzi N, Asadi Z. Air pollution impact on the Covid-19 mortality in Iran considering the comorbidity (obesity, diabetes, and hypertension) correlations. *Environ Res*. 2022;204(Pt A):112020.
26. Pana TA, Bhattacharya S, Gamble DT, et al. Country-level determinants of the severity of the first global wave of the COVID-19 pandemic: an ecological study. *BMJ Open*. 2021;11:e042034.
27. Pansini R, Fornacca D. COVID-19 higher mortality in Chinese regions with chronic exposure to lower air quality. *Front Public Health*. 2020;8:597753.

28. Prinz AL, Richter DJ. Long-term exposure to fine particulate matter air pollution: An ecological study of its effect on COVID-19 cases and fatality in Germany. *Environ Res*. 2022;204(Pt A):111948.

29. Qeadan F, Mensah NA, Tingey B, et al. The association between opioids, environmental, demographic, and socioeconomic indicators and COVID-19 mortality rates in the United States: an ecological study at the county level. *Arch Public Health*. 2021;79:101.

30. Sanchez-Piedra C, Cruz-Cruz C, Gamino-Arroyo AE, Prado-Galbarro FJ. Effects of air pollution and climatology on COVID-19 mortality in Spain. *Air Qual Atmos Health*. 2021; 1–7.

31. Stieb DM, Evans GJ, To TM, et al. Within-city variation in reactive oxygen species from fine particle air pollution and COVID-19. *Am J Respir Crit Care Med*. 2021;204:168–177.

32. Zheng P, Chen Z, Liu Y, et al. Association between coronavirus disease 2019 (COVID-19) and long-term exposure to air pollution: evidence from the first epidemic wave in China. *Environ Pollut*. 2021;276:116682.

33. Wakefield J. Ecologic studies revisited. *Annu Rev Public Health*. 2008;29:75–90.

34. Lagarde F, Pershagen G. Parallel analyses of individual and ecologic data on residential radon, cofactors, and lung cancer in Sweden. *Am J Epidemiol*. 1999;149:268–274.

35. Greenland S, Robins J. Invited commentary: ecologic studies–biases, misconceptions, and counterexamples. *Am J Epidemiol*. 1994;139:747–760.

36. Goldberg MS, Villeneuve PJ. Re: Long-term exposure to air-pollution and COVID-19 mortality in England: a hierarchical spatial analysis. *Environ Int*. 2021;150:106422.

37. Goldberg MS, Villeneuve PJ. Re: An ecological analysis of long-term exposure to PM2.5 and incidence of COVID-19 in Canadian health regions. *Environ Res*. 2021;194:110610.

38. Villeneuve PJ, Goldberg MS. Re: Links between air pollution and COVID-19 in England. *Environ Pollut*. 2021;274:116576.

39. Tessum CW, Paolella DA, Chambliss SE, Apte JS, Hill JD, Marshall JD. PM2.5 polluters disproportionately and systemically affect people of color in the United States. *Sci Adv*. 2021;7:eabf4491.

40. Rossen LM, Ahmad FB, Anderson RN, et al. Disparities in excess mortality associated with COVID-19 - United States, 2020. *MMWR Morb Mortal Wkly Rep*. 2021;70:1114–1119.

41. Reardon S. How the Delta variant achieves its ultrafast spread [published online ahead of print July 21, 2021]. *Nature*. doi: 10.1038/d41586-021-01986-w

42. Pouwels KB, Pritchard E, Matthews PC, et al. Impact of Delta on viral burden and vaccine effectiveness against new SARS-CoV-2 infections in the UK. *Nat Med*. 2021; 27:2127–2135.

43. Wu X, Nethery RC, Sabath MB, Braun D, Dominici F. Air pollution and COVID-19 mortality in the United States: strengths and limitations of an ecological regression analysis. *Sci Adv*. 2020;6:eabd4049.

44. World Health Organization. Fact Sheet: HIV\AIDS. 2021. Available at: https://www.who.int/news-room/fact-sheets/detail/hiv-aids. Accessed August 8, 2021.

45. van Donkelaar A, Martin RV, Li C, Burnett RT. Regional estimates of chemical composition of fine particulate matter using a combined geoscience-statistical method with information from satellites, models, and monitors. *Environ Sci Technol*. 2019;53:2595–2611.

46. Wu X, Nethery RC, Sabath MB, Braun D, Dominici F. Data repository: Air Pollution and COVID mortality in the U.S. Available at: https://github.com/wxwx1993/PM_COVID. 2021. [Accessed 22 December 2021]

47. Wu X, Nethery RC, Sabath MB, Braun D, Dominici F. Air pollution and COVID-19 mortality in the United States. 2021. Available at: https://github.com/wxwx1993/PM_COVID. Accessed July 16, 2021.

48. Centers for Disease Control. *AlasPlus*. U.S. Centers for Disease Control, 2021.

49. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, 2020.

50. Venables WN, Ripley BD. *Modern Applied Statistics with S*. 4th ed. Springer, 2002.

51. D. L. *sjPlot: Data Visualization for Statistics in Social Science*. R package version 2.8.9., 2021.

52. Britton A, Jacobs Slifka KM, Edens C, et al. Effectiveness of the Pfizer-BioNTech COVID-19 vaccine among residents of two skilled nursing facilities experiencing COVID-19 outbreaks - Connecticut, December 2020-February 2021. *MMWR Morb Mortal Wkly Rep*. 2021;70:396–401.

53. Hill AB. The environment and disease: association or causation? *Proc R Soc Med*. 1965;58:295–300.

54. Zhang R, Li Y, Zhang AL, Wang Y, Molina MJ. Identifying airborne transmission as the dominant route for the spread of COVID-19. *Proc Natl Acad Sci U S A*. 2020;117:14857–14863.

55. Woodby B, Arnold MM, Valacchi G. SARS-CoV-2 infection, COVID-19 pathogenesis, and exposure to air pollution: what is the connection? *Ann N Y Acad Sci*. 2021;1486:15–38.

56. Frontera A, Martin C, Vlachos K, Sgubin G. Regional air pollution persistence links to COVID-19 infection zoning. *J Infect*. 2020;81:318–356.

57. Martelletti L, Martelletti P. Air Pollution and the novel Covid-19 disease: a putative disease risk factor. *SN Compr Clin Med*. 2020; 1–5.

58. Andersen ZJ, Hoffmann B, Morawska L, et al. Air pollution and COVID-19: clearing the air and charting a post-pandemic course: a joint workshop report of ERS, ISEE, HEI and WHO. *Eur Respir J* 2021;58:2101063.

59. Glencross DA, Ho TR, Camiña N, Hawrylowicz CM, Pfeffer PE. Air pollution and its effects on the immune system. *Free Radic Biol Med*. 2020;151:56–68.

60. Qian H, Miao T, Liu L, Zheng X, Luo D, Li Y. Indoor transmission of SARS-CoV-2. *Indoor Air*. 2021;31:639–645.

61. McGreevy R. Outdoor transmission accounts for 0.1% of State's Covid-19 cases. *Irish Times*. 2021. Available at: www.irishtimes.com/news/ireland/irish-news/outdoor-transmission-accounts-for-0-1-of-state-s-covid-19-cases-1.4529036

62. Hamner L, Dubbel P, Capron I, et al. High SARS-CoV-2 attack rate following exposure at a choir practice—Skagit County, Washington, March 2020. *MMWR Morb Mortal Wkly Rep*. 2020;69:606–610.

63. Chirizzi D, Conte M, Feltracco M, et al. SARS-CoV-2 concentrations and virus-laden aerosol size distributions in outdoor air in north and south of Italy. *Environ Int*. 2021;146:106255.

64. Bulfone TC, Malekinejad M, Rutherford GW, Razani N. Outdoor transmission of SARS-CoV-2 and other respiratory viruses: a systematic review. *J Infect Dis*. 2021;223:550–561.

65. Rothman KJ, Greenland S. Causation and causal inference in epidemiology. *Am J Public Health*. 2005;95(Suppl 1):S144–S150.

66. Heederik DJJ, Smit LAM, Vermeulen RCH. Go slow to go fast: a plea for sustained scientific rigour in air pollution research during the COVID-19 pandemic. *Eur Respir J*. 2020;56:2001361.

67. Hansell AL, Villeneuve PJ. Invited perspective: ambient air pollution and SARS-CoV-2: research challenges and public health implications. *Environ Health Perspect*. 2021;129:111303.