Links between air pollution and COVID-19 in England

Marco Travaglio ¹, Yizhou Yu ¹, Rebeka Popovic, Nuno Santos Leal, Luis Miguel Martins

MRC Toxicology Unit, University of Cambridge
Correspondence to: L. Miguel Martins, e-mail: martins.lmiguel@gmail.com
¹Equal contributions

Revision version v03
Main revisions: We built on the negative binomial regression models from our last version. Our new model predicts COVID-19 cases and deaths from air pollutants, while taking into account population densities in the different English regions.

Running title: Air Pollution and COVID-19 in England
ABSTRACT

In December 2019, a novel disease, coronavirus disease 19 (COVID-19), emerged in Wuhan, People's Republic of China. COVID-19 is caused by a novel coronavirus (SARS-CoV-2) thought to have jumped species from another mammal to humans. This virus has caused a rapidly spreading global pandemic. To date, thousands of cases of COVID-19 have been reported in England, and over 18,000 patients have died. While there has been progress in managing this disease, it is not clear which factors, in addition to age, affect the severity and mortality of COVID-19. A recent analysis of COVID-19 in Italy identified links between air pollution and death rates. Here, we explored potential links between three major air pollutants related to fossil fuels and SARS-CoV-2 mortality in England. We compared current, SARS-CoV-2 cases and deaths recorded in public databases to region-level air pollution data monitored at over 120 sites across England. We found that the levels of some markers of poor air quality, nitrogen oxides and ozone, were associated with COVID-19 mortality in different English regions, after adjusting for population density. We conclude that the levels of some air pollutants are linked to COVID-19 cases and morbidity. We consider that our study provides a useful framework to guide health policies in countries affected by this pandemic.

Keywords: SARS-CoV-2; COVID-19; Air pollution; Nitrogen oxides, Ozone, mortality.
Abbreviations: SARS, severe acute respiratory syndrome
INTRODUCTION

In December 2019, a high number of pneumonia cases of unknown aetiology were detected in Wuhan, China. A molecular analysis of samples from affected patients revealed that their symptoms were caused by an infection by a novel coronavirus, later named severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the pathogenic agent of coronavirus disease 19 (COVID-19)\(^1\). Coronaviruses are enveloped, non-segmented positive-sense RNA viruses belonging to the family Coronaviridae and classified within the Nidovirales order\(^2\).

In a previous coronavirus outbreak, another coronavirus was found to induce SARS, leading to a wide spectrum of clinical manifestations, including respiratory failure and death\(^2\). While the initial symptoms of COVID-19 include fever with or without respiratory syndrome, a crescendo of pulmonary abnormalities can develop later in patients\(^3\). Recent studies have shown that most patients present with only mild illness, but approximately 25% of hospital-admitted cases require intensive care because of viral pneumonia with respiratory complications\(^4\).

While abundant research into the pathogenesis of COVID-19 suggests that this disease likely stems from an excessive inflammatory response\(^5\), the exact predisposing factors contributing to increased clinical severity and death in patients remain unclear. Moreover, the risks associated with COVID-19 are unknown. Previous studies suggested that individuals over the age of 60 or with underlying health conditions, including cardiovascular disease, diabetes, chronic respiratory disease and cancer, are at highest risk of severe disease and death\(^6\).

Curiously, long-term exposure to air pollutants has been shown to be a risk factor mediating the pathogenesis of these health conditions\(^7\). In fact, prolonged exposure to common road transport pollutants, including nitrogen oxides and ground-level ozone, a by-product of chemical reactions between nitrogen oxides and volatile organic compounds, can significantly exacerbate cardiovascular morbidity, airway oxidative stress and asthma\(^8\). Such pollutants can also cause a persistent inflammatory response and increase the risk of infection by viruses that target the respiratory tract\(^9,10\). We therefore addressed whether infectivity and mortality rates of COVID-19 in England are linked to air pollution.

This hypothesis is supported by observations that the geographical patterns of COVID-19 transmission and mortality among countries, and even among regions of the same country, closely align with local levels of air pollutants\(^10\). For example, increased contagiousness and COVID-19-related mortality in northern Italian regions, including Lombardia, Veneto and Emilia Romagna, have been correlated with high levels of air pollutants in these regions\(^10\).
These findings suggest that individuals exposed to chronic high levels of air pollutants may have increased susceptibility to SARS-CoV-2 infection because of compromised immune system defence responses.

Here, we explored the relationship between air pollution and SARS-CoV-2 mortality in England and showed an association of airborne toxins released by fossil fuels with susceptibility to viral infection by SARS-CoV-2 and COVID-19 mortality in England.

RESULTS

We analysed cumulative counts of COVID-19 cases and deaths against the annual average concentrations of three major air pollutants recorded between 2018 and 2019, when no COVID-19 cases were reported. Our study used publicly available data (Table 1) from seven regions in England, where a minimum of 2,000 SARS-CoV-2 infections and 200 deaths were reported from February 1 to April 8, 2020, approximately two weeks after the UK was placed into lockdown.

The spatial pattern of COVID-19 deaths appeared to mimic the geographical distribution of COVID-19-related cases, with the largest numbers of COVID-19 deaths occurring in London and in the Midlands (Figure 2). Previous studies suggest that those two areas present the highest annual average concentration (μg/m³) of nitrogen oxides. In addition, ground-level ozone concentrations have been previously shown to vary significantly with latitude and altitude, depending on concentration of ozone in the free troposphere, long-range transport and its precursor emissions. Therefore, we sought to determine if spatial variations in nitrogen dioxide, nitrogen oxide as well as ground-level ozone concentrations in England are associated with increased COVID-19 infection and mortality. To this end, we constructed three generalised linear models to estimate the association between each air pollutant with the cumulative number of both COVID-19 cases and deaths in England.

To assess the association between the number of COVID-19 related cases and the spatial variation of the three air pollutants, we initially built an ordinary least square (OLS) regression model (Table 2, column 1). We added population density, a confounding factor, as an independent variable to this model to account for differences in the number of inhabitants across regions. After the OLS model was fitted to the data, the analysis demonstrated a large residual standard error, indicating poor accuracy of the model. Therefore, we next fitted a Poisson (Table 2, column 2) and a negative binomial regression model (Table 2, column 3), which are known to be more appropriate for count data. The negative binomial model arose
as a more parsimonious model, supported by its higher log likelihood and a lower Akaike Information Criterion (AIC)\textsuperscript{13}. We, therefore, used the negative binomial model for further analysis. This model showed that population density is a significant predictor of COVID-19 cases (\(p\)-value < 0.01) as well as levels of nitrogen oxide dioxide (\(p\)-value < 0.01). This shows that levels of nitrogen oxides increase the accuracy of the model, indicating that they act as significant predictors of COVID-19 cases. We next applied a similar workflow for the number of COVID-19 deaths (Table 3). The negative binomial regression remains the optimal model, and shows that ozone, nitrogen oxide and nitrogen dioxide are significantly associated with COVID-19 deaths, together with population density.

Taken together, the negative binomial models of both COVID-19 cases and deaths (third column of Tables 2 and 3) show that nitrogen dioxide, nitrogen oxide and ozone are significant predictors of COVID-19 related death, after accounting for population density. This provides the first evidence that SARS-CoV-2 cases and deaths are associated with the levels of pollutants in England.

**METHODS**

**Data sources**

The number of SARS-CoV-2 infection cases in England were obtained from Public Health England (PHE) and analysed according to the following statistical regions: London, Midlands, north-west, north-east and Yorkshire, south-east, east and south-west of England. Regional-level data on the cumulative number of SARS-CoV-2-related deaths in England was retrieved from the National Health System (NHS). This source includes the most comprehensive region-specific records of COVID-19 deaths in England. The daily deaths summary included deaths of patients who died in hospitals in England and had tested positive for SARS-CoV-2 at the time of death. While this online repository is updated on a daily basis, it is important to note that figures are subject to change due to \textit{post-mortem} diagnosis confirmation. Furthermore, data do not include deaths outside the hospital and therefore serve as an approximation of the total number of deaths in England. All deaths are recorded as the date of death rather than the day in which the death was announced. The total numbers of COVID-19 deaths for each region were collected up to the 8th of April 2020 (Table 1).
Air pollution data between 2018 and 2019 were obtained from the European Environmental Agency (EEA). This agency maintains one of the largest online repositories of current, quality-controlled environmental information in Europe. For the present study, we collected annual aggregated air quality (AQ) values determined by the EEA based on direct observations obtained from over 120 monitoring stations across England. Stations located near airports (i.e., Luton) or near/within national reserves were excluded from the analysis, as measured observations were sparse and inconsistent with their respective surroundings. Because of limited data availability, we also restricted our analysis to individual pollution indices for three major air pollutants, namely, nitrogen dioxide, nitrogen oxide and ozone, across the prespecified regions (Figure 2). Nitrogen dioxide, nitrogen oxide and ozone AQ values are expressed in µg/m³ and represent the annual average of daily measurements for each air polluting substance from 2018 to 2019 in each specified region. The identification of each monitoring station was matched to each available city by accessing the Department for Environment, Food and Rural Affairs (DEFRA) website (Figure 1). This website contains a resource called DEFRA’s Air Quality Spatial Object Register, which allows users to view and retrieve information on the air quality-related spatial and non-spatial data objects from the UK’s Air Quality e-Reporting data holdings. The annual average of daily measurements for each pollutant in each monitoring area was analysed to determine the influence of toxic exposure on the number of SARS-CoV-2 cases and deaths across England (Figure 1).

Regional heatmaps

Heatmap legends were generated using GraphPad Prism 8 (www.graphpad.com), and regions are labelled with the mapped colour values.

Statistical analysis

We fitted and compared three types of generalised linear models to our data using COVID-19 deaths and cases as the outcome and nitrogen oxide, nitrogen dioxide and ozone as the exposures of interest, adding the corresponding population density values as a confounding variable. Population density (person/km²) data corresponds to subnational mid-year population estimates of the resident population in England for the year 2018 and excludes visitors or short-term immigrants (< 12 months). First, we fitted a simple OLS model, which makes basic assumptions and guides further analyses. Poisson regression models are generally preferred for modelling count data, but the Poisson distribution is restricted as it does not account for over dispersed data. Negative binomial models have higher flexibility,
offering narrower confidence intervals if the conditional distribution of the outcome variable is over-dispersed. The fitting goodness of each regression model was determined by the log-likelihood and AIC statistics. The models were built using the MASS package (www.stats.ox.ac.uk/pub/MASS4/) in R. The model comparison tables were generated using the Stargazer package. The analysis workflow is available in GitHub (https://m1gus.github.io/AirPollutionCOVID19/).

DISCUSSION

In this study, we identified an association between air pollution and SARS-CoV-2 lethality in England, expanding previous evidence linking high mortality rates in Europe with increased toxic exposure to air pollutants. In addition, a new study showed an association between fine particulate matter levels and COVID-19-related deaths in the US by demonstrating that 1 µg/m³ increase in this particulate matter led to a 15% increase in COVID-19 death rates. Notably, these findings are in line with studies conducted during the previous SARS outbreak, where long-term exposure to air pollutants had a detrimental effect on the prognosis of SARS patients in China.

Previous studies have shown that the estimated model used to predict the probability of an event greatly affects analysis outcomes in epidemiological studies. Therefore, three distinct analysis methods were employed in this study to select the most appropriate, to examine the links between air pollution and COVID-19 in England.

Among the models employed, the negative binomial model emerged as the most appropriate based on likelihood test statistics. This model shows that nitrogen oxides and ozone have a significant effect on the cumulative number of COVID-19 deaths, and nitrogen oxides have a significant effect on the cumulative number of COVID-19 cases, independently of population density.

These models do not account for other confounders such as age, which other studies have shown to be important. Furthermore, we built models by aggregating all variables into seven regions, losing granularity. Based on the above, we did not consider appropriate to calculate mortality rate ratios or suggest that an increase of 1 µg/m³ in air pollutants will increase the number of cases or death. Therefore, we will use additional data to improve these models and their parameter estimates. Nonetheless, our results show that the levels of
nitrogen oxides have an effect on COVID-19 related cases and death while ozone only has an effect on COVID-19 deaths.

Nitrogen oxides gases result from a chemical reaction between nitrogen and oxygen during the combustion of fossil fuels and therefore represent a significant source of air pollution in areas with high traffic. Previous studies have shown that exposure to nitrogen dioxide is associated with a significant decrease in pulmonary function and promotes an inflammatory response in the airway. In fact, a recent study showed that intensive care unit (ICU) ventilation duration was associated with preadmission exposure to nitrogen dioxide. Furthermore, Faustini and colleagues reported that a 10 µg/m³ increase in the annual concentration of nitrogen dioxide was associated with 13% and 2% increases in cardiovascular and respiratory mortality, respectively. As respiratory and cardiovascular diseases represent potential risk factors for increased COVID-19 mortality, our results are in line with the hypothesis that long-term exposure to nitrogen dioxide is linked with an increased risk of COVID-19-related mortality.

Five-year aggregated nitrogen oxide level data from both industrialised and rural areas in England indicated that the progression of respiratory syncytial virus (RSV) was directly linked to seasonal variations in nitrogen oxide levels. While the mechanisms by which environmental nitrogen oxide aggravates the clinical severity of RSV remain unclear, it has been noted that when inhaled, this gas depresses endogenous production of nitrogen oxide. Since endogenous nitrogen oxide is known to inhibit RSV replication in the lungs, it has been proposed that exogenous, inhaled nitrogen oxide may trigger RSV replication by repressing the production of endogenous nitrogen oxide. Although a similar mechanism of replication has not been shown for SARS-CoV-2, it is tempting to speculate a possible association between nitrogen oxide levels and SARS-CoV-2 lethality. Moreover, nitrogen oxide is a free radical and can be oxidized to nitrogen dioxide, which acts as a pulmonary irritant.

Ozone is a secondary by-product of traffic-related air pollution and is generated through sunlight-driven reactions between motor-vehicle emissions and volatile organic compounds. Generally, high ozone levels have been associated with reduced lung function and increased incidence of respiratory symptoms. We found the lowest levels of ozone in highly urbanised regions, such London and the Midlands. Given the highly reactive nature of ozone, decreased levels in these regions may indicate increased conversion of ozone to secondary gaseous species, a phenomenon previously reported for areas with heavy traffic. Moreover, ozone can readily react with other gaseous species and particulates in the
environment, resulting in the formation of respiratory irritants, such as terpene derivatives. Further research is necessary to determine the exact identity of these pollutants and their effect on COVID-19 severity and progression.

Our results, combined with those from the recent reports from northern Italy, Europe, and the USA, suggest that poor AQ increases the lethality of COVID-19. Future detailed studies may further elucidate these observations by addressing potential confounders, including socioeconomic status, comorbidities, age, race, and differences between regional health regulations and their ICU capacities. Nonetheless, our study highlights the importance of continuous implementation of existing air pollution regulations for the protection of human health, both in relation to the COVID-19 pandemic and beyond.

FIGURES & FIGURE LEGENDS

Figure 1. Analysis workflow.

This flowchart summarizes how raw data was extrapolated, processed and analysed. Population density data (person/km²) was derived from ONS and used to account for region-specific differences in population size across England; COVID-19 cases and deaths data were obtained from PHE and NHS, respectively; Air pollution data from each monitoring station was manually curated using DEFRA’s Air Quality Spatial Object Register and aggregated into statistical regions. ONS, Office for National Statistics; PHE, Public Health England; NHS, National Health Service; EEA, European Environmental Agency.
Figure 2. Regional heatmaps of COVID-19 and pollutants.

Regional English heatmaps of reported deaths and diagnosed COVID-19 cases through April 8, 2020 (top row), as well as air quality (AQ) values for the indicated pollutants (bottom row).
TABLES & TABLE LEGENDS

Table 1. Summary of data sources.

<table>
<thead>
<tr>
<th>Data type</th>
<th>Source</th>
<th>Download date</th>
<th>Measuring units</th>
</tr>
</thead>
<tbody>
<tr>
<td>COVID-19 Cases</td>
<td>Public Health England (<a href="https://coronavirus.data.gov.uk/#region">https://coronavirus.data.gov.uk/#region</a>)</td>
<td>April 9, 2020</td>
<td>Lab-confirmed cases per region up to and including April 8, 2020</td>
</tr>
<tr>
<td>Population data</td>
<td>Office for National Statistics (<a href="https://www.ons.gov.uk">https://www.ons.gov.uk</a>)</td>
<td>April 17, 2020</td>
<td>Region-level population density in England (person/km²)</td>
</tr>
</tbody>
</table>
Table 2. Effect of air pollutants on COVID-19 cases in England.

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>Number of cases until 8 April 2020</th>
<th>95% confidence intervals</th>
<th>p-values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS</td>
<td>Poisson</td>
<td>binomial</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>Average_Pop_density_personkm2</td>
<td>-3.805 (-8.339, 0.729)</td>
<td>-0.001*** (-0.001, -0.001)</td>
<td>-0.001*** (-0.001, -0.0004)</td>
</tr>
<tr>
<td>NO.levels</td>
<td>-2.778 (-4.781, -0.774)</td>
<td>-0.0005*** (-0.0005, -0.0004)</td>
<td>-0.0005*** (-0.001, -0.0003)</td>
</tr>
<tr>
<td>NO2.levels</td>
<td>4.118 (0.942, 7.294)</td>
<td>0.001*** (0.001, 0.001)</td>
<td>0.001*** (0.0005, 0.001)</td>
</tr>
<tr>
<td>O3.levels</td>
<td>0.238 (-0.726, 1.202)</td>
<td>0.0001*** (0.0001, 0.0001)</td>
<td>0.0001 (-0.00002, 0.0002)</td>
</tr>
</tbody>
</table>

Observations: 7
R²: 0.936
Adjusted R²: 0.809
Log Likelihood: -713.041
theta: 33.532* (17.929)
Akaike Inf. Crit.: 1,436.081
Residual Std. Error: 1,923.060 (df = 2)
F Statistic: 7.373 (df = 4; 2)

Note: *p<0.1; **p<0.05; ***p<0.01

Each column of the table corresponds to a different type of regression model as indicated at the top. The raw estimate values of each model are listed with their 95% confidence intervals in parentheses. The p-values are indicated using the number of asterisks beside the estimates. OLS, ordinary least square; Average_Pop_densitykm2, average population density per square kilometer; NO.levels, nitrogen oxide levels; NO2.levels, nitrogen dioxide levels; O3.levels, ozone levels; Akaike Inf. Crit., Akaike’s Information Criteria; Residual Std. Error, Residual standard error.
Table 3. Effect of air pollutants on COVID-19 deaths in England.

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>OLS</th>
<th>Poisson</th>
<th>negative binomial</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>Average_Pop_density_personkm2</td>
<td>-0.631 (-1.129, -0.133)</td>
<td>-0.001*** (-0.001, -0.001)</td>
<td>-0.001*** (-0.001, -0.001)</td>
</tr>
<tr>
<td>NO levels</td>
<td>-0.332* (-0.552, -0.112)</td>
<td>-0.0005*** (-0.0001, -0.0004)</td>
<td>-0.0005*** (-0.0001, -0.0004)</td>
</tr>
<tr>
<td>NO2.levels</td>
<td>0.601* (0.253, 0.950)</td>
<td>0.001*** (0.001, 0.001)</td>
<td>0.001*** (0.001, 0.001)</td>
</tr>
<tr>
<td>O3.levels</td>
<td>0.088 (-0.018, 0.194)</td>
<td>0.0001*** (0.0001, 0.0001)</td>
<td>0.0001*** (0.0001, 0.0001)</td>
</tr>
<tr>
<td>Constant</td>
<td>-12,339,490 (-21,816,650, -2,862,319)</td>
<td>-11,590*** (-13,372, -9,809)</td>
<td>-10,900*** (-14,604, -7,195)</td>
</tr>
<tr>
<td>Observations</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>R²</td>
<td>0.956</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>0.868</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log Likelihood</td>
<td>-49.547</td>
<td>-40.873</td>
<td></td>
</tr>
<tr>
<td>theta</td>
<td>196.421 (128.791)</td>
<td>196.421 (128.791)</td>
<td>196.421 (128.791)</td>
</tr>
<tr>
<td>Akaiake Inf. Crit.</td>
<td>109.094</td>
<td>91.745</td>
<td></td>
</tr>
<tr>
<td>Residual Std. Error</td>
<td>211.114 (df = 2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F Statistic</td>
<td>10.901* (df = 4; 2)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* p<0.1; ** p<0.05; *** p<0.01

This table can be interpreted using the same guidance given for Table 2.
CONFLICTS OF INTEREST
The authors declare no conflicts of interest.

AUTHOR CONTRIBUTIONS
MT, RP, YY and LMM planned and designed the study; MT, RP, YY, NSL collected the data; MT, YY and NSL treated and analysed the data; YY developed the models; MT, RP, YY wrote the manuscript with the support of NSL and LMM. MT, YY and RP conducted this work while in self-isolation due to the current pandemic.

ACKNOWLEDGEMENTS
We are grateful to all the staff members with critical functions in administration, operations and logistics at the MRC Toxicology Unit during the present crisis.

FUNDING
This work is funded by the UK Medical Research Council, intramural project MC_UU_00025/3 (RG94521).
REFERENCES