Environmental Pollution 268 (2021) 115859

Contents lists available at ScienceDirect

Environmental Pollution

journal homepage: www.elsevier.com/locate/envpol

Links between air pollution and COVID-19 in England☆

Marco Travaglio [1], Yizhou Yu [1], Rebeka Popovic, Liza Selley, Nuno Santos Leal, Luis Miguel Martins*

MRC Toxicology Unit, University of Cambridge, UK

ARTICLE INFO

Article history:
Received 29 June 2020
Received in revised form
17 September 2020
Accepted 13 October 2020
Available online 19 October 2020

Keywords:
SARS-CoV-2
COVID-19
Air pollution
Nitrogen oxides
Ozone
$PM_{2.5}$
PM_{10}
Mortality

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) presumed to have jumped species from another mammal to humans. This virus has caused a rapidly spreading global pandemic. To date, over 300,000 cases of COVID-19 have been reported in England and over 40,000 patients have died. While progress has been achieved in managing this disease, the factors in addition to age that affect the severity and mortality of COVID-19 have not been clearly identified. Recent studies of COVID-19 in several countries identified links between air pollution and death rates. Here, we explored potential links between major fossil fuel-related air pollutants and SARS-CoV-2 mortality in England. We compared current SARS-CoV-2 cases and deaths from public databases to both regional and subregional air pollution data monitored at multiple sites across England. After controlling for population density, age and median income, we show positive relationships between air pollutant concentrations, particularly nitrogen oxides, and COVID-19 mortality and infectivity. Using detailed UK Biobank data, we further show that $PM_{2.5}$ was a major contributor to COVID-19 cases in England, as an increase of 1 m^3 in the long-term average of $PM_{2.5}$ was associated with a 12% increase in COVID-19 cases. The relationship between air pollution and COVID-19 withstands variations in the temporal scale of assessments (single-year vs 5-year average) and remains significant after adjusting for socioeconomic, demographic and health-related variables. We conclude that a small increase in air pollution leads to a large increase in the COVID-19 infectivity and mortality rate in England. This study provides a framework to guide both health and emissions policies in countries affected by this pandemic.

1. Introduction

In December 2019, a high number of pneumonia cases with an unknown aetiology were detected in Wuhan, China. A molecular analysis of samples from affected patients revealed that their symptoms were caused by an infection with a novel coronavirus, later named severe acute respiratory syndrome (SARS) coronavirus (CoV) 2 (SARS-CoV-2), the pathogenic agent of coronavirus disease 19 (COVID-19) (Zhu et al., 2020a). Within five months, this disease had affected more than 210 countries and became a global pandemic, causing devastating consequences to public health (Wang et al., 2020a). Coronaviruses are a genus of enveloped, non-segmented, positive-sense RNA viruses belonging to the family Coronaviridae and classified within the Nidovirales order (Yi et al., 2020). Historically, illnesses caused by coronaviruses have ranged in severity, with some, including human coronaviruses-229E and -OC43, causing common cold symptoms, but SARS-CoV and Middle East respiratory syndrome coronavirus have initiated outbreaks of life-threatening pneumonia (Yi et al., 2020). While the initial symptoms of COVID-19 include fever with or without respiratory syndrome, a crescendo of pulmonary abnormalities may subsequently develop in patients (Huang et al., 2020). According to recent studies, most patients present with only a mild illness, but approximately 25% of hospital-admitted patients require intensive care because of viral pneumonia with respiratory complications (Wang et al., 2020a).

While extensive research into the pathogenesis of COVID-19 suggests that the severe disease likely stems from an excessive inflammatory response (Cao, 2020), the exact predisposing factors contributing to increased clinical severity and death in patients remain unclear. Individuals over the age of 60 years or with

☆ This paper has been recommended for acceptance by Da Chen.

* Corresponding author.
E-mail addresses: martins.lmiguel@gmail.com, lm795@cam.ac.uk (L.M. Martins).

[1] Equal contributions

https://doi.org/10.1016/j.envpol.2020.115859

for the year 2018 (single-year model) or temporally averaged levels for the years 2014–2018 (multi-year model) to capture the multi-year trajectory of historical change in air pollution in England. Air pollutant levels are reported in μg/m^3, except for ozone, whose metric is the number of days on which the daily max 8-hr concentration is greater than 120 μg/m^3. A detailed quality report regarding this data is available at the following website: https://uk-air.defra.gov.uk/assets/documents/reports/cat09/1903201606_AQ0650_2017_MAAQ_technical_report.pdf. We obtained the longitude and latitude of each local authority using OpenCage Geocoder (https://opencagedata.com/). The air pollutant levels for each authority was approximated by averaging 10 values nearest the centre of authority. This area covers approximately 12 km^2. Detailed descriptions of the methodology and analysis workflow are available in our GitHub repository. For the UK Biobank data, we matched the location coordinate each participant reported to their nearest modelled value. The level of each pollutant was measured less than 2 km away from the self-reported address.

2.3. UK biobank data sources

We used data from the UK Biobank under application #60124. Details regarding the geographical regions, recruitment processes, and other characteristics have been previously described (Sudlow et al., 2015),[14], and are found on ukbiobank.co.uk. The UK Biobank has received ethical approval from the North West – Haydock Research Ethics Committee, 11/NW/0382 to gather data from participants. A detailed list of the variables analysed in the present study is presented in Supplementary Table 1 (https://m1gus.github.io/AirPollutionCOVID19/). Notably, we defined hypertension using the criteria of a diastolic blood pressure ≥90 mmHg OR systolic blood pressure ≥140 mmHg. We assigned an average of annual pollutant concentrations from the PCM data to each study subject, on the basis of a six-digit postcode. Individual-level data were collected from the UK Biobank on April 26, 2020. This dataset contained information on individuals that tested positive for COVID-19. No COVID-19 test data were available for UKB assessment centers in Scotland and Wales, thus data from these centers were not included.

2.4. Regional heatmaps

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

2.5. Statistical analysis

In our regional exploratory analysis, we fitted generalised linear models to our data using COVID-19 deaths and cases as the outcomes 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/km^2) data correspond to subnational mid-year population estimates of the resident population in England and excludes visitors or short-term immigrants (<12 months). We modelled both the number of cases and deaths using negative binomial regression analyses since the response variables are overdispersed count data. We used the same model type for our subregional analysis, adding mean annual earnings and median age as confounding factors.

For the UK Biobank models, we fitted a binomial regression model because the response variable, COVID-positive or -negative, is defined as either 0 or 1.

Methods for assessing the fit of the model included residual analyses, diagnostic tests, and information criterion fit statistics. The goodness of fit of each regression model was determined using the log-likelihood and Akaike Information Criterion (AIC) statistics.

For all models, we calculated the odds or risk ratios and their 95% confidence intervals to quantify the effects of the independent variables on the response variables. The models were built using the MASS package (www.stats.ox.ac.uk/pub/MASS4/) in R. The comparison tables were generated using the Stargazer package (Hlavac, 2018). The analysis source code, detailed quality checks and all Supplementary material are available in GitHub (https://github.com/M1gus/AirPollutionCOVID19). The analysis notebook is available at the following link: https://m1gus.github.io/AirPollutionCOVID19/. Statistical significance was defined as $p \leq 0.05$.

3. Results

3.1. Links between regional nitrogen oxide and ozone levels and COVID-19 in England

We analysed the associations between cumulative numbers of COVID-19 cases and deaths with the concentrations of three major air pollutants recorded between 2018 and 2019, when no COVID-19 cases were reported. Due to differences in data availability for each air pollutant, we only included annual mean values of daily measurements, which was the most consistent aggregation type reported for all air pollutants described in this analysis. We started by analysing publicly available data from seven regions in England (Table 1). For each region, a minimum of 2000 SARS-CoV-2 infections and 200 deaths were reported by PHE from February 1 to April 8, 2020, which was approximately two weeks after the UK was placed into lockdown (Fig. 1).

The spatial pattern of COVID-19 deaths matched the geographical distribution of COVID-19-related cases, with the largest numbers of COVID-19 deaths occurring in London and in the Midlands (Fig. 2). According to previous studies, those two areas present the highest annual average concentration (μg/m^3) of nitrogen oxides (Pannullo et al., 2017). In addition, ground-level ozone concentrations have been previously shown to vary significantly with latitude and altitude, depending on the concentration of ozone in the free troposphere, long-range transport and emission of its precursor (Hagenbjork et al., 2017). Therefore, we sought to determine if spatial variations in the levels of nitrogen oxides, in particular nitrogen dioxide (NO_2) and nitrogen oxide (NO), as well as ground-level ozone concentrations in England are associated with increased numbers of COVID-19 infections and mortality. We applied a negative binomial regression model to estimate the association between each air pollutant with the cumulative number of both COVID-19 cases and deaths at the regional level (Supplementary Tables 2 and 3). The model was chosen based on the data type (count data) and log likelihood and AIC scores (Akaike et al., 1998). Population density, a confounding factor, was added to this model as an independent variable to account for differences in the number of inhabitants across regions. The levels of nitrogen oxide and nitrogen dioxide are significant predictors of COVID-19 cases ($p < 0.05$), independent of the population density (Supplementary Table 2). We next applied a similar method to assess the association with the number of COVID-19 deaths (Supplementary Table 3). Ozone, nitrogen oxide and nitrogen dioxide levels are significantly associated with COVID-19 deaths, together with the population density.

Taken together, the negative binomial regression models of both COVID-19 cases and deaths (Supplementary Tables 2 and 3) show that nitrogen dioxide, nitrogen oxide and ozone levels are significant predictors of COVID-19-related death, after accounting for the population density. This study provides the first evidence that