ISGlobal
Barcelona Institute for Global Health
info@isglobal.org
www.isglobal.org
Campus Clínic
Rosselló, 132
08036 Barcelona, Spain
T +34 93 227 1806
Campus Mar
Doctor Aiguader, 88
08003 Barcelona, Spain
T +34 93 214 7300

Application title: Air pollution in relation to COVID-19 morbidity and mortality: a large population-based cohort study in Catalonia, Spain (COVAIR-CAT)

RFA Number and Title: RFA 20-1B, Air pollution, COVID-19, and Human Health

Principal investigator: Cathryn Tonne, Associate Research Professor, ISGlobal, cathryn.tonne@isglobal.org

Key co-investigators:

ISGlobal: Xavier Basagaña, Associate Research Professor; Payam Dadvand, Associate Research Professor; Otavio Ranzani, Postdoctoral researcher; Maria Foraster, Assistant Research Professor; Joan Ballester, Associate Research Professor; Jordi Sunyer, Research Professor; Mark Nieuwenhuijsen, Research Professor; Carlos Chaccour, Assistant Research Professor.
IDIAP JGol: Talita Duarte-Salles, Senior Epidemiologist

Catalan Agency for Health Quality and Assessment (AQuAs): Rosa Vivanco Hidalgo, Director of Health Technology Assessment

Study duration: 2 year project, estimated total budget $500,000

Rationale and aims There are wide disparities between and within countries in the health burden from COVID-19. Although demographic factors explain part of these differences, much of the observed heterogeneity remains poorly understood. Exposure to air pollution may contribute to these differences as it is an important risk factor for several communicable and non-communicable diseases that appear to increase vulnerability to hospital admission or death due to COVID-19. However, to date, nearly all studies investigating air pollution exposure and COVID-19 hospital admissions or mortality have been ecological in design.[1,2] Individual-level studies are needed to provide robust evidence, allowing for more comprehensive control of confounding, and investigation of vulnerable population subgroups.

Spain has been particularly affected by the COVID-19 pandemic, with the highest number of deaths per population (598 deaths per million as of June 4) in the world. The autonomous region of Catalonia has the highest number of deaths in Spain after Madrid. Urban areas of Catalonia are burdened by high levels of traffic-related air pollution, and high levels of ozone are frequent downwind of metropolitan areas. With wide spatial variation in air pollution levels, a good spatial distribution of COVID-19 cases (20% occurred in rural areas), and powerful electronic health record databases from a universal health system covering nearly the entire population, Catalonia is an ideal setting to conduct well-powered studies of the role of air pollution in COVID-19 hospital admission or mortality and to identify vulnerable population subgroups.

There is a critical need for practical and credible knowledge regarding the potential for air quality control to contribute to the prevention of COVID-19 morbidity and mortality. While COVID-19 is an immediate public health priority, evidence is needed to put the role of air pollution in COVID-19 in the broader context. Understanding whether the role of air pollution in COVID-19 related health outcomes is different from its role in the wider suite of respiratory infections is essential for identifying public health priorities. Our **overarching objective** is to test whether long or short-term exposure to air pollution increases the risk of COVID-19 hospital admissions or mortality and to identify vulnerable subgroups. Specifically, we aim to:

1) test whether long-term (e.g. previous 1 to 5 years) exposure to air pollution is associated with COVID-19 hospital admission or mortality in the general population.
2) test whether short-term exposure (on the order of days, weeks) to air pollution is associated with COVID-19 hospital admission or mortality following COVID-19 diagnosis.

3) test whether the influence of long or short-term exposure to air pollution on COVID-19 outcomes differs according to individual-level sociodemographic factors, comorbidities, area-level socioeconomic factors, or other environmental exposures.
4) test whether the influence of long-term air pollution exposure on COVID-19 hospital admissions and mortality differs from that for respiratory infections not due to SARS-COV2 infection.

Study design and methods We will address these specific aims using a large population-based cohort derived from electronic health records covering over 80% of the population of Catalonia. We will include all individuals registered in the public health system in 2015 and follow them prospectively through the end of 2020. We are already assembling this cohort as part of a separate study focused on air pollution and stroke occurrence, and can efficiently expand the cohort through data linkage to include first diagnosis of COVID-19 through the primary care system, hospital admission, and mortality.

Data sources and access to air quality and health data *Health outcomes:* We will use data from primary care records integrated with testing, hospital, and mortality registries allowing for a complete characterization of COVID-19 cases starting from symptom onset. This unique dataset has been described previously.[3]

We will investigate the association between long-term exposure to air pollution and COVID-19 hospital admission (to date, around 30,000 admissions) and all-cause and cause-specific mortality (to date, around 7,000 COVID-19 deaths). COVID-19 hospital admissions and deaths will be identified through specific ICD10 codes (currently being standardized by national and international agencies). We will focus on COVID-19 outcomes occurring during the period 1 Mar –31 Dec 2020. To address Aim 4, we will focus on hospital admissions and mortality due to 1) acute respiratory infections and 2) respiratory conditions more generally during 1 Jan 2017- 31 Jan 2020, prior to the COVID-19 outbreak.

Air pollution exposure: We leverage air pollution exposure models developed through several previous projects. Residential addresses for all individuals included in the public health system have already been geocoded as part of a previous project. Geocoded addresses will be linked to the following exposure data allowing for comparisons across models and pollutants:

- Daily PM_{10} and $PM_{2.5}$ at 1km resolution from models developed through the HEI-funded APACHE project. We will update models as part of this project through the end of 2020.
- Annual average $PM_{2.5}$, NO_2, O_3, and BC from the HEI-funded ELAPSE project, with temporal adjustment using monitoring data to account for temporal trends over the study period.
- In nested analyses based on the Barcelona population (the majority of COVID-19 cases in Catalonia), finer spatially resolved exposure to traffic-related air pollution ($PM_{2.5}$, $PM_{2.5}$ light absorption, and NO_2) based on a hybrid land use regression-dispersion model developed as part of the HEI-funded FRONTIER project.

Individual and area-level covariates: Individual-level data are available on age, sex, nationality, tobacco and alcohol use, socioeconomic position, and residence in nursing home. Data on a range of comorbidities are available through the primary care system including for example: obesity, previous diagnosis of hypertension, ischemic heart disease, dementia, diabetes mellitus, cerebrovascular disease and chronic respiratory disease, as well as long-term medication use.[3] A range of area-level indicators can be linked to the individuals in the cohort regarding socioeconomic position and socio-demographics (e.g. % minority ethnic group, % unemployment, urban vulnerability index, and average income tax).

Environmental co-exposures: A wide range of additional environmental exposures are being linked to the cohort through ongoing work including indicators of noise exposure and attributes of the built environment such as surrounding green, blue, and gray spaces, streetscape, and walkability.

COVID-19 dynamics: The majority of COVID-19 diagnoses occurred in large cities, which also have the highest levels of air pollution. Population density, seroprevalence, and other aspects of infection dynamics are likely to confound the relationship between air pollution and COVID-19 morbidity and mortality. To account for these, we will leverage seroprevalence data at the province level (4 within Catalonia),[4] spatially resolved data (300m x 300m) on COVID-19 symptoms reported to the Catalan Health System via mobile application,[5] as well as number of COVID-19 diagnosis from the primary health care system at the census tract level.

Plan for statistical analyses We will use semi-parametric hazard models with time-varying covariates to quantify associations between air pollution and COVID-19 hospital admission and mortality. Analyses will account for individual and area-level covariates and for competing-risks (e.g. Fine and Gray's semi-parametric proportional subhazards model) where relevant. For all analyses, we will use directed acyclic graphs to identify potential confounders and mediators and guide analyses plans and will account for spatial autocorrelation.

- To address **Aim 1**, we will model associations between long-term exposure to several pollutants and COVID-19 hospital admission and mortality in the full cohort (n=6 million individuals).
- To investigate whether short term exposure to air pollution increases the risk of hospital admission or mortality in **Aim 2,** we will construct a nested cohort based on individuals diagnosed with COVID-19 in the primary care system in 2020 (to date, approximately 150,000). Individuals will be followed for 30 days post-diagnosis for hospital admission or mortality. We will use daily PM_{10} and $PM_{2.5}$ from the APACHE models to investigate whether particulate air pollution in the days preceding hospital admission or death was associated with event occurrence.[6,7] We will consider a range of lagged exposures to investigate the relevant timing of exposure. In addition to individual and area-level covariates, models to address Aim 2 will include long-term air pollution exposure, as well as daily temperature and humidity. In sensitivity analyses, we will include only individuals with a positive PCR test for SARS-COV2 (to date, around 60,000), rather than doctor diagnosis.
- We will conduct extensive analyses of predefined subgroups to test whether the influence of long and short-term air pollution on COVID-19 related health outcomes differs according to individual and area-level characteristics and co-exposure to other environmental factors (**Aim 3**).
- To address **Aim 4**, we will fit similar models to those used to address Aim 1 but instead focus on hospital admissions for respiratory infections in general, and pneumonia in particular, prior to the onset of COVID-19. We will compare effect estimates for COVID-19 hospital admission and mortality with similar outcomes not due to SARS-COV2 infection to identify whether the influence of air pollution on COVID-19 related outcomes differs from respiratory infections more generally.

Study limitations The strengths of our study are its large size and representativeness (covering 6 out of 7.5 million residents of Catalonia), and data on COVID-19 at first diagnosis through the primary care system. The analysis should be well powered to identify main effects as well as for subgroup analyses. The study team is highly experienced in air pollution epidemiology, use of electronic health records (including investigators based within the organizations responsible for managing these data), and infectious disease epidemiology (including investigators involved in the local COVID-19 response). The main limitation will be missing values among covariates in the electronic health records. We will explore in detail patterns of missingness and appropriate methods for accounting for missing data (e.g. multiple imputation). Data on asymptomatic cases of SARS-COV2 infection are not available; however, we will use multiple indirect surrogates of infections to minimize residual confounding from spatio-temporal patterns in infections.

References

1. Wu X, Nethery RC, Sabath BM, Braun D, Dominici F. Exposure to air pollution and COVID-19 mortality in the United States: A nationwide cross-sectional study. *medRxiv* 2020. DOI:10.1101/2020.04.05.20054502.
2. Domingo JL, Rovira J. Effects of air pollutants on the transmission and severity of respiratory viral infections. *Environ Res* 2020; **187**: 109650.
3. PRIETO-ALHAMBRA D, Ballo E, Coma-Redon E, *et al.* Hospitalization and 30-day fatality in 121,263 COVID-19 outpatient cases. *medRxiv Prepr* 2020. DOI:10.1101/2020.05.04.20090050.
4. Institue of Health Carlos III. National Study of Sero-epidemiology of infections with SARS-COV2 in Spain (Round 2). 2020 https://www.mscbs.gob.es/ciudadanos/ene-covid/docs/ESTUDIO_ENE-COVID19_SEGUNDA_RONDA_INFORME_PRELIMINAR.pdf.
5. Generalitat de Catalunya A. STOP COVID CAT. 2020.
6. Szpiro AA, Sheppard L, Adar SD, Kaufman JD. Estimating acute air pollution health effects from cohort study data. *Biometrics* 2014; **70**: 164–74.
7. Lepeule J, Rondeau V, Filleul L, Dartigues J-F. Survival Analysis to Estimate Association between Short-Term Mortality and Air Pollution. *Environ Health Perspect* 2006; **114**: 242–7.