

REVIEW

Open Access



Methodological limitations in studies assessing the effects of environmental and socioeconomic variables on the spread of COVID-19: a systematic review

Maria A. Barceló^{1,2} and Marc Saez^{1,2*}

Abstract

Background: While numerous studies have assessed the effects of environmental (meteorological variables and air pollutants) and socioeconomic variables on the spread of the COVID-19 pandemic, many of them, however, have significant methodological limitations and errors that could call their results into question. Our main objective in this paper is to assess the methodological limitations in studies that evaluated the effects of environmental and socioeconomic variables on the spread of COVID-19.

Main body: We carried out a systematic review by conducting searches in the online databases PubMed, Web of Science and Scopus up to December 31, 2020. We first excluded those studies that did not deal with SAR-CoV-2 or COVID-19, preprints, comments, opinion or purely narrative papers, reviews and systematic literature reviews. Among the eligible full-text articles, we then excluded articles that were purely descriptive and those that did not include any type of regression model. We evaluated the risk of bias in six domains: confounding bias, control for population, control of spatial and/or temporal dependence, control of non-linearities, measurement errors and statistical model. Of the 5631 abstracts initially identified, we were left with 132 studies on which to carry out the qualitative synthesis. Of the 132 eligible studies, we evaluated 63.64% of the studies as high risk of bias, 19.70% as moderate risk of bias and 16.67% as low risk of bias.

Conclusions: All the studies we have reviewed, to a greater or lesser extent, have methodological limitations. These limitations prevent conclusions being drawn concerning the effects environmental (meteorological and air pollutants) and socioeconomic variables have had on COVID-19 outcomes. However, we dare to argue that the effects of these variables, if they exist, would be indirect, based on their relationship with social contact.

Keywords: COVID-19, Environmental (meteorological and air pollutants) variables, Socioeconomic variables, Social contacts

Background

Numerous studies have assessed the effects of environmental and socioeconomic variables on the spread of the COVID-19 pandemic. Most of them have addressed the influence meteorological variables have, although there are also quite a few that have considered the effects of air pollutants and socioeconomic variables. Those which assessed the effects of meteorological variables were

*Correspondence: marc.saez@udg.edu

¹ Research Group On Statistics, Econometrics and Health (GRECS), and CIBER of Epidemiology and Public Health (CIBERESP), University of Girona, Carrer de la Universitat de Girona 10, Campus de Montilivi, 17003 Girona, Spain

Full list of author information is available at the end of the article

the first to appear, specifically between the last week of March and the first week of April 2020. In other words, very close to COVID-19 being officially declared a global pandemic (11 March 2020) [1]. Later, there were those which evaluated the effects of air pollutants, the first of which appeared between the end of April and the first week of May 2020. Finally, the last ones to appear were those related to socioeconomic variables; the first of which was mid-May 2020.

The studies differ in their outcomes (new and cumulative cases, mortality, reproductive number, etc.), study populations (the world, countries, regions, cities), confounders as well as in the way of controlling for them, and in the modelling strategies adopted. However, with the exception of socioeconomic variables, several systematic reviews attempting to synthesize the evidence have already been published.

For instance, with regard to meteorological variables, Mecenas et al. carried out a bibliographic search until the end of March 2020 [2]. In reviewing 17 studies (most of them preprints), they found that warm wet climates seemed to reduce the spread of COVID-19. However, the role of temperature and humidity on the spread of the virus was very moderate, since these variables alone could not explain most of the variability in the disease's transmission. Smit et al., in a systematic review carried out in July 2020 (that is, of studies that used data from the first wave), critically evaluated 42 articles published in scientific journals and 80 preprints [3]. They concluded that the evidence suggested that either there was no modulating effect of the summer weather conditions (i.e., high temperature and low humidity reduce the transmission rate of the virus) or, along the same lines as Mecenas et al., if it did exist, it was weak. Smit et al. also found similar results for other meteorological variables, such as ultraviolet radiation and wind speed [3]. McClymont and Hu discussed 23 articles with moderate or high ratings (out of a total of 86 eligible peer-reviewed articles) published until October 1 (also contemplating only the first wave) [4], and found that temperature and humidity were associated with COVID-19 incidence. However, while the decrease in temperature was associated with increases in incidence, in the variations in humidity the results were mixed (positive and negative associations were found). They also found that wind speed and rainfall results were not consistent across studies [4].

In relation to air pollutants, Copat et al. carried out a systematic review of 15 studies (13 articles and 2 preprints) published between April 2020 and July 6th, 2020 [5]. They found a consistent association between some air pollutants (fine particles, $PM_{2.5}$ with a diameter of 2.5 microns (μm) or less, and nitrogen dioxide, NO_2 , and with a less extent coarse particles, PM_{10} , with

a diameter of 10 μm or less) and a higher incidence and mortality from COVID-19. They pointed out, however, that there were important limitations for any direct comparison of the results and that more studies were needed to strengthen scientific evidence. Malecki et al. carried out a systematic review of 19 studies, published through to October 31, 2020, that assessed the association of particulate matter (i.e., PM_{10} and $PM_{2.5}$) pollution and the spread of SARS-CoV-2 [6]. They pointed out that although there were suggestions that particulate matter (PM) played a role in the spread of SARS-CoV-2, PM concentration alone cannot be effective in spreading the COVID-19 disease, and that other meteorological and environmental variables were also involved.

Until today (June 2021), no peer-reviewed systematic reviews have been published concerning the influence socioeconomic variables have on the spread of the pandemic. However, let us advance some of our results here by noting that in ecological studies the results were not conclusive. In some, especially those carried out in the United States, the areas with greater economic deprivation had a higher incidence and also a higher mortality. That said, in others no association was found, or deprivation was even found to be a protective factor. What was consistently observed was the fact that the higher the population density was, the greater incidence and mortality were. In individual studies, however, individuals with lower incomes or from more disadvantaged groups were at greater risk of hospitalization and death.

Nevertheless, all the reviews state that many of the studies have significant methodological limitations and errors that could bring their results into question. Our main objective here is to assess the methodological limitations in the studies that evaluated the effects environmental and socioeconomic variables have had on the spread of COVID-19. Furthermore, we discuss the results of those studies that were, in fact, able to control those very limitations.

Methods

Systematic review

The protocol for this review is registered in the Prospective Register of Systematic Reviews (PROSPERO 2020 CRD42020201540). In the review process, we followed the preferred reporting items for systematic reviews and meta-analysis (PRISMA) protocols [7]. The literature search, study selection, data extraction, and quality assessment were performed by each of us independently. In case of any discrepancy between us, we all reached an agreement on the final decision.

By combining the keyword 'COVID-19' with the keywords 'temperature', '(meteorological variables)', '(air pollutants)', '(environmental variables)', and '(socioeconomic

variables), through the Boolean connector ‘AND’ we conducted a search in the online databases PubMed, Web of Science and Scopus, up to December 31, 2020. We did not impose any language restrictions, nor did we contact any author for additional information.

All the articles retrieved underwent an initial title and abstract screening, where any duplicates were discarded, followed by a full-text screening for eligible abstracts. We made a first exclusion of those studies that did not deal with SARS-CoV-2 or COVID-19, preprints (non-peer-reviewed articles), comments, opinion or purely narrative papers, reviews and systematic literature reviews (Fig. 1). Among the eligible full-text articles, we made a second exclusion of those articles that were purely descriptive (including only plots or maps, etc.) and those that did not include any type of regression model (those that only included the analysis of correlations, for example).

We extracted the following data from the articles included in the qualitative analysis: first author, study population, study period, outcome, explanatory variables, covariates, the statistical method (including the model specification and the methods to control the confounding), and the study findings.

Methodological limitations

The usual assessment tools for observational studies were not entirely suitable for assessing the risk of bias of the studies we reviewed. We preferred to adapt the tool proposed by Parmar et al. [8] who, in turn, adapted the Newcastle–Ottawa scale [9] and the RTI item bank [10]. Specifically, we used six domains: two from Parmar et al. [8]—confounding bias and measurement errors in the outcome and/or in the exposure variables; one based on the dimension ‘unobserved confounding’ in Saez et al.

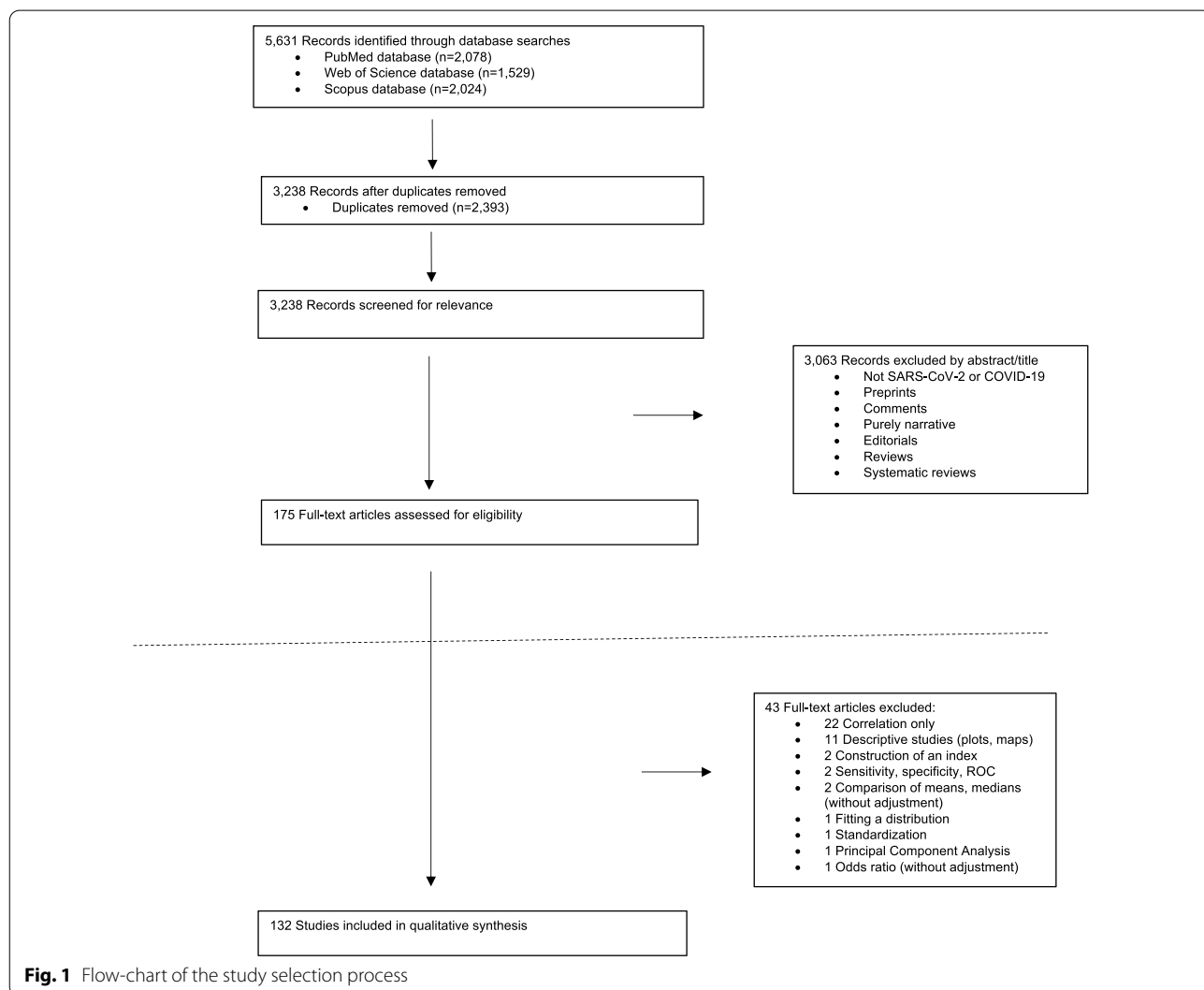


Fig. 1 Flow-chart of the study selection process

[11]—control of the spatial and/or the temporal dependence; and three that we added ex novo in this paper—control for the population, statistical model, and control of non-linearities.

In each study, each of the six domains were rated as: 1—low risk of bias, 2—moderate risk of bias, or 3—high risk of bias) (Table 1). For the overall rating of each study, we evaluated it as 'strong' (low risk of bias) if, at most, one of the six domains was rated as high risk of bias (i.e., a rating of 3), 'moderate' (moderate risk of bias) if up to two domains were rated as weak, or 'weak' (high risk of bias) if three or more domains were rated as high risk of bias. For the rating of both the domains and the studies, we rely on Parmar et al. [8].

Three of the six dimensions corresponded to the specification error known as omission of relevant variables: confounding bias, control of the population and control of the spatial and/or of the temporal dependence. This

specification error leads to biased and inconsistent estimators (that is, the estimators biased even asymptotically, i.e., when the number of observations is very high) and, in addition, the variances of the estimators are also misleading [12]. In any case, the inference of those studies that do not control for this error is highly compromised.

Confounding bias

None of the studies included all possible confounders, especially if the studies were ecological (as most of them were). However, as regards the spread of COVID-19, there is a confounder that, at a minimum, must be controlled for, namely, social contact.

The main route of transmission for COVID-19 is through the direct or indirect contact with an infected subject via the small droplets that occur when they cough or sneeze [13]. Thus, this contact must be controlled for in the models, even if indirectly. The control, although

Table 1 Bias assessment tool

Bias domain	Question to Consider	Indicator	Score (1: Low Risk of Bias; 2: Medium Risk; 3: High risk)
Confounding bias	Did the study analysis adjust potential confounders appropriately?	Confounders adjusted	1: Many confounders and unobserved confounding 2: Some confounders (in particular mobility or socioeconomic variables) or unobserved confounders 3: None or cannot tell
Control of the population	Did the study control for population?	Population, age structure of the population	1: Control and/or including population density 2: Control only by including population density 3: No control
Control of the spatial and/or temporal dependence	Did the study control the spatial and/or temporal extra variability?	Spatial and temporal dependence	1: Control 2: Partial control (control of only one dependence, for example) 3: No control
Control of non-linearities	Did the study control for non-linearities?	Non-linearities (parametric or non-parametric)	1: Control 3: No control
Measurement errors	What was the heterogeneity of indicators used in the study?	Measurement errors in the explanatory variables (exposure variables and covariates)	1: Control 2: Partial control (including lags, for example) 3: No control
Statistical model	Did the study use appropriate statistical model?	Statistical model	1: Models for count data response variables 2: Control of heteroscedasticity and rates as response variables 3: Models with normally distributed errors and count data response variable; or no control of heteroscedasticity and rates as response variables
Overall Study Rating	Strong (low risk of bias): one domain, at most, was rated as 3 Moderate: up to 2 domains were rated as 3 Weak (high risk of bias): ≥ 3 domains were rated as 3		

partial, can be carried out through mobility or, much more indirectly, through socioeconomic variables. In general, greater mobility implies greater levels of contact. Likewise, areas with high population densities are known to have greater social contact. Furthermore, some occupations present a greater risk, particularly those that were less able to switch to teleworking and, therefore, require greater mobility and the resulting higher level of social contact.

Unobserved confounding (i.e., residual confounding) including, for example, random effects that capture heterogeneity, should also be controlled for. In other words, unobserved variables specific to the unit of analysis (area or individual) that could influence the risk of, in this case, the spread of the COVID-19.

We scored this domain with a 3 if the confounding was not controlled for by any method, with a 2 if the observed confounding was controlled for with a moderate number of confounders (up to two maximum), in particular mobility or socioeconomic variables, or with a 1 if the observed confounding was controlled for with a large number of confounders (more than two) and/or unobserved confounding was also controlled for.

Control of the population

Perhaps the main relevant variable that should not be omitted by any study is that of population at risk, either in the study area (in ecological studies) or in the area in which the subject resides (in individual studies). It is evident that both incidence and mortality, as well as other outcomes (hospitalizations, ICU admissions, etc.), depend both on the population of the area under study and on the age structure of that population.

Population control can be carried out in various ways: using rates, including the population or the expected value of the outcome in each area under study in the model as an offset, or controlling, as covariates, the size of the population or its structure (for example, percentage of population aged 65 years or more).

A control of the population can also be achieved by including population density (i.e., the number of people per unit of area, usually per square kilometre) as a covariate. However, it is possible that, in this case, control would only be partial. On one hand, an area with a higher population density does not always have more population than another, but it depends, logically, on its surface. On the other hand, population density could be capturing other socioeconomic variables.

This domain was scored with a 3 if the population was not controlled for by any method, a 2 if the population was controlled for by only including population density as a covariate, or a 1 if the population was controlled for, in addition to including the population density by other additional method.

Control of the spatial and/or of the temporal dependence

Several studies analyze, as outcome, cumulative cases and cumulative deaths. Many others, however, use a temporal design. This is a design, where both the outcome and its possible explanatory variables, as well as the covariates, are measured in the form of time series. Time series are observed with a certain periodicity, usually regular (for example, daily) over a given period of time.

In this case there is temporal dependency. The outcome observations are not independent but are related, so their future behavior is predictable. In general, this dependence can be long or short term. A long-term dependency, or trend, could be defined as a movement or tendency in the data. As is known, in the case of COVID-19 there have been between two and four waves, depending on the country. That is, long-term swings have occurred. Periods in which the outcome values are persistently high, followed by others in which the values have been low. Short-term dependency, also called serial autocorrelation, refers to the relationship of the values of an outcome on, for example, a given day with the values of the previous days, especially with those of the preceding day.

Most studies use a spatial or spatio-temporal design. In other words, they observe the outcome in different geographical areas, and sometimes over time. When a spatial design is available, it is important to distinguish two sources of variation. In the first place, the most important source is usually the so-called 'spatial dependence' and is a consequence of the correlation of the spatial unit with neighboring spatial units, generally those that are geographically contiguous. In this way, the risks (for example, of transmission) of contiguous or nearby areas are more similar than the risks of spatially distant areas. Part of this dependency is not really a structural dependency but is mainly due to the existence of uncontrolled variables, that is, not included in the analysis. Meanwhile, the second source, the existence of spatially independent and unrelated variation called 'spatial heterogeneity,' must be assumed. This is a consequence of the existence of unobserved variables without spatial structure that could influence risk [14].

The temporal and the spatial dependence must be controlled for, because, otherwise, in the best of cases, the variances of the estimators will be misleading (when the outcome is a continuous variable, normally distributed, and least squares methods are used for the inference) and in most cases, not only will the variances be biased, but the estimators will also be biased (when the outcome is not a continuous variable, not normally distributed, and least squares methods cannot be used) [12].

In some studies, the control of temporal or spatial dependence is not applicable. Thus, in studies with a time series design but in which a very short period of time is

analyzed, it does not make sense to control for temporal dependence. Likewise, in those studies with a spatial (or spatio-temporal) design but that analyze very spatially distant territories (for example, several countries in the world) it does not make sense to control for the spatial dependence.

We scored this domain with a 3 when neither temporal nor spatial dependency was controlled and should have been; a 2 when the control was partial, controlling only one dependency and not controlling the other; and a 1 when they were controlled.

Control of non-linearities

Along with the omission of relevant variables, the error in the functional form constitutes the most important specification error. The relationships between environmental variables and COVID-19 outcomes are not usually linear. Thus, for example, in Fig. 2, we show the smoothed curves for the relationship between the daily temperature and the daily levels of nitrogen dioxide (NO₂) and the daily number of cases for Spain in the period between January 1, 2020 and April 14, 2021. Specifically, we draw the estimated curves in a generalized additive model in which we use smoothing splines with a quasi-likelihood Poisson link, i.e., taking into account over-dispersion.

As can be seen, in none of the cases was the relationship linear. These non-linearities must be controlled in the models, because, otherwise, as when relevant variables are omitted, the estimators will be inconsistent and their variances misleading.

We scored this dimension with a 3 if non-linearities were not controlled for (again, when applicable) or a 1 if they had been controlled.

Measurement errors

Measurement errors (also known as misclassification) can occur in both the response variable and in the exposure variables.

The definition of the response variable can vary in space and time, even within the same country, leading to differential misclassification. In Spain, for example, the Catalan government, on the one hand, defined a death from COVID-19 as being a positive result on some test (PCR or fast test) or symptoms presented at some point which a health professional subsequently classified as a possible case, but the individual did not have a diagnostic test with a positive result [15], whereas on the other hand, the Spanish government, defined a death from COVID-19 as being someone who presented a positive PCR result [16], thus providing significantly lower figures. This misclassification continued until May 21, 2020, when the Government of Spain adopted the same definition as the Government of Catalonia [17].

However, the measurement errors in the response variable are not attributable to the investigators, although they should certainly discuss them if appropriate. Furthermore, fortunately, when measurement errors occur in the dependent variable, the estimators remain consistent, although they are not efficient [12], that is, not very precise, thus leading to wider confidence intervals than if there had been no measurement errors.

There is, however, an important problem if measurement errors occur in the explanatory variables (exposure or covariates). If the explanatory variables are measured with error, the estimators will be inconsistent [12].

Even in studies at the individual level, the exposure variables and, obviously, the contextual variables (for

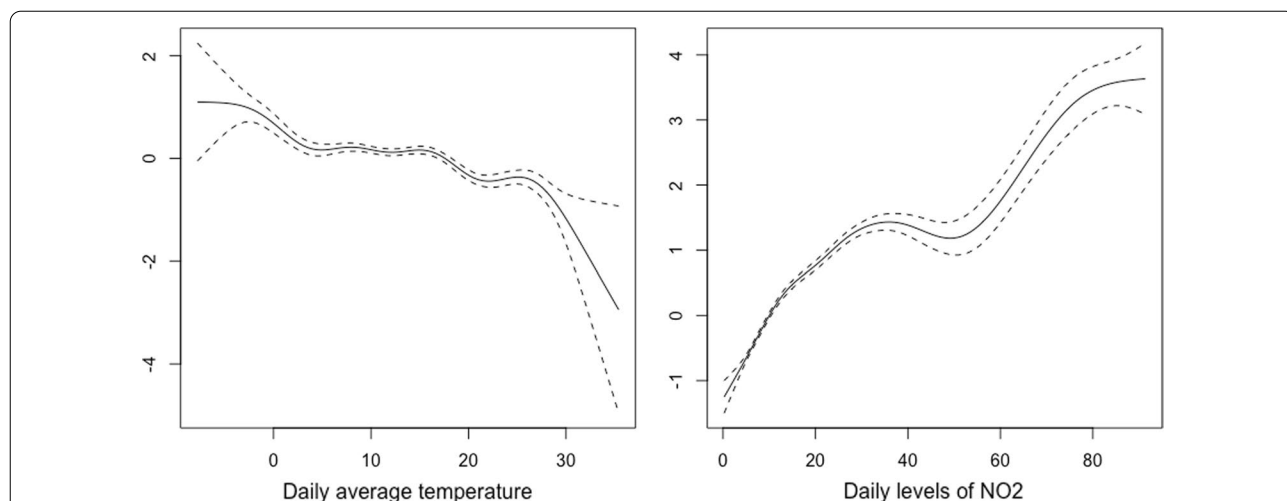


Fig. 2 Smoothed curves for the relationships between daily temperature and daily levels of nitrogen dioxide and the number of daily cases of COVID-19, Spain, January 1, 2020 to April 14, 2021. The data were obtained from: [16]. Environmental data [81, 82]

example, the socioeconomic ones) are not observed at the individual level, but are aggregated at the level of the area under study. Nevertheless, not all residents in the area under study are actually exposed to the same mean values of the explanatory variables, which leads to a measurement error. If the misclassification is non-differential (over time and over space within the area under study) and, furthermore, if the between-area variability of the variable measured with error is much greater than the within-area variability of such variable, that is, that the area under study is not very heterogeneous (for example, because it is a small area), then the effect of the measurement error on the estimator consistency may be negligible [18]. This is what happens in the case of contextual socioeconomic variables as long as the area under study is not very large.

In the exposure variables (both air pollutants and meteorological variables), however, there is differential misclassification, because the exposure exhibits spatial variation across the area under study. If the spatial structure (i.e., spatial dependence) of the data is ignored, the estimators will be biased and inconsistent [19]. Many studies use the measurements observed in the area under study to estimate, by means of point estimators, exposure levels for that entire area. The estimators most widely used are the arithmetic mean of the values of the exposure, observed in several monitoring or meteorological stations in the area, and sometimes the inverse-distance weighted average of these values.

This measurement error in the exposure variables must be controlled for, either explicitly incorporating the spatial dependence, in the ecological studies, or by correcting the misalignment between the locations of the observation points of the exposure variables and that locations of the individuals, in the studies at the individual level.

In studies with an ecological spatial design, the 'modifiable areal unit problem' (MAUP) occurs [20]. The MAUP is a consequence either because areas of different sizes are added (scale effect) and/or because of the way the area is divided (zoning effect) [21]. In either case, it is a potential source of bias. For example, Wang and Di found that the association between nitrogen dioxide (NO₂) and COVID-19 deaths varies when the data is aggregated at different levels: a risk factor when the area is smaller (aggregation of districts and cities) and a protective factor at the province level [22]. Similarly, we also found a positive association between NO₂ and deaths as a consequence of COVID-19 at the level of a county-like area [17] and no association at a lower level of aggregation [23].

When using a temporal design, the 'modifiable temporal unit problem' (MTUP) [24] also occurs, whereby the

results depend on the way data are temporally aggregated [21]. Furthermore, in this type of design, temporal misalignment can occur. In other words, the relationship between exposure and the occurrence of COVID-19 outcome is not contemporary, but rather is distributed over time as a consequence of the incubation period of COVID-19 and due to the diagnostic delays of the outcome. This temporal misalignment must be controlled by including lags, for example.

We scored this dimension with a 3 if measurement errors in the exposure variables are not controlled at all, a 2 if they are only partially controlled (not including lags, for example) or the areas under study are very large (countries, for example) and a 1 if they have been controlled for.

Statistical model

Many of the studies, even though the response variable is a count data, used regression models with normally distributed errors (linear regression models, generalized linear and additive models with Gaussian link, etc.). Using this type of models leads to biased results, unless the number of counts is very large. However, this was not the case in most studies.

Some studies did not model the counts but rather the rates, dividing the dependent variable by the size of the population. However, since the numerator, being a count data, is actually distributed following a Poisson distribution, the variance is proportional to the mean, so it is not constant, leading to heteroscedasticity (i.e., overdispersion). This must be controlled for, otherwise, the variances of the estimators are misleading.

To illustrate the effects on the results of erroneously using a regression model with normally distributed errors, we used the data in Filippini et al. [25]. Their objective was to investigate the link between the transmission of SARS-CoV-2 infection and long-term exposure to NO₂ in the provinces of three regions of Northern Italy (Lombardia, Veneto and Emilia Romagna), between March 8 and April 5, 2020 ($n = 84$). Using their data, we first estimated a linear regression model including, as a dependent variable, the number of new daily SARS-CoV-2 positive cases (count data variable). We found that long-term NO₂ levels to which the inhabitants of the provinces of the Italian regions studied had been exposed to be positively associated with the total number of cases that occurred in the period considered. Specifically, for every 1 µg/m³ increase in the NO₂ levels, the number of cases increased by 18.478 for the entire period (95% confidence interval, 95% CI 10.285–27.210). However, the residuals of the model were not normally distributed (Fig. 3). We then modelled the rates (cases per 100,000 inhabitants) using a linear regression model, although we

did not control for heteroscedasticity. For every 1 $\mu\text{g}/\text{m}^3$ increase in NO_2 , the number of cases increased by 1.207 cases per 100,000 inhabitants (95% CI 0.050–2.364). However, the residuals presented a clear heteroscedasticity behavior (the scatter plot of the residuals against the adjusted values did not present a constant dispersion, i.e., variance), and furthermore, they were not normally distributed (Fig. 3). When we estimated a generalized Poisson model, in which we took into account the overdispersion, and in which we included the population size as an offset, we could not reject the null hypothesis that the parameter associated with the long-term exposure of NO_2 was equal to zero (95% CI: $-0.004, 0.001$).

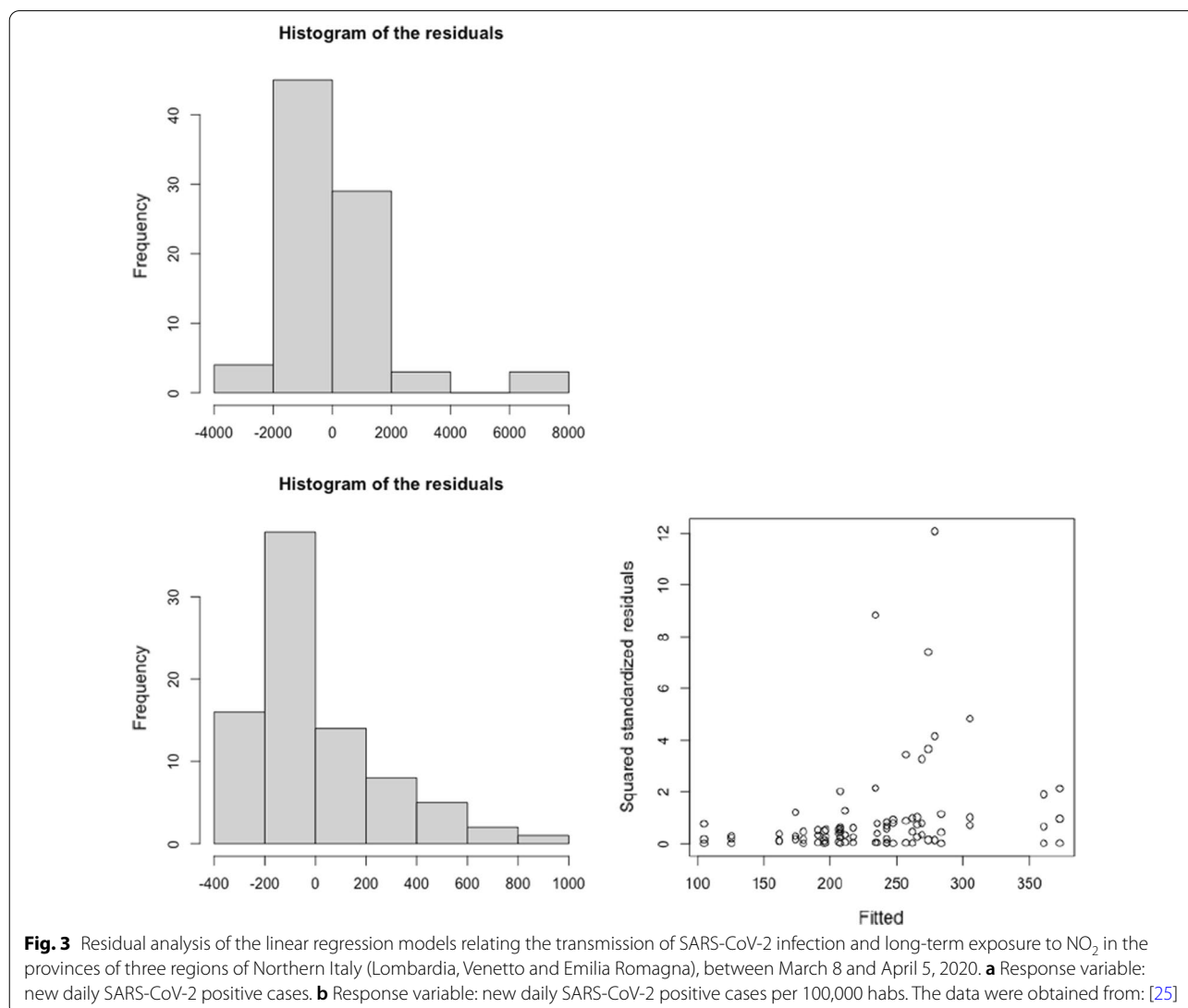
We scored this dimension with a 3 when the outcome was a count data and regression models with normally distributed error were used. We also scored a 3 when rates were modelled but heteroscedasticity was not

controlled for. Meanwhile, we scored a 2 if rates were modelled and heteroscedasticity was controlled, and a 1 if models for count data response variables were used (Poisson regression, negative binomial regression, etc.).

Results

Systematic review

Figure 1 shows a flowchart of the review process. Of the 5631 abstracts initially identified, and after excluding duplicates, we were left with 3238 studies. From these we excluded 3063 studies that did not refer to SARS-CoV-2 or COVID-19, preprints, comments, those purely narrative studies, editorials and reviews and systematic reviews, thus leaving us with 175 eligible studies. As we said, we excluded 43 studies that were purely descriptive and those that did not include any type of regression model (Additional file 1: Table S1). In the end we were



left with 132 studies with which to carry out the qualitative synthesis (Additional file 1: Tables S2 and S3).

Of the 132 studies, 92 referred to meteorological variables, 40 to socioeconomic variables and 34 to air pollutants. Seventy-one of the studies referred only to meteorological variables, 21 only to socioeconomic variables and 16 only to air pollutants. Of the 92 studies that referred to meteorological variables, 16 also considered air pollutants, 14 meteorological variables and socioeconomic variables. Four of the studies referring to air pollutants also referred to socioeconomic variables but not to meteorological variables. Nine referred to meteorological variables and socioeconomic variables but not to air pollutants. Finally, five studies considered meteorological variables, air pollutants and socioeconomic variables (Additional file 1: Figure S1).

Of the 132 studies finally selected, 124 used an ecological design and nine an individual design. Most ecological studies considered different regions (states, regions, provinces, counties, cities, etc.) within the same country as study populations (71 studies). This is followed by those that considered countries or cities in the world (34 studies) and, finally, those that considered individual cities or smaller areas (19 studies). Seven of the eight studies with an individual design, analyzed the influence of socioeconomic variables, while only two considered socioeconomic variables and air pollutants.

Most of the studies (129 out of 132) analyzed data referring up to August 1, 2020 (i.e., only considering the first wave). In fact, only three consider the first two waves of the pandemic.

Methodological limitations

Table 2 shows the evaluation of the studies included in the qualitative synthesis. Of the 132 eligible studies, we evaluated 63.64% (84 of 132) as weak (high risk of bias), 19.70% (26 of 132) as moderate (moderate risk of bias) and 16.67% (22 of 132) as strong (low risk of bias). Only four studies did not have any dimension scored with a 3 (high risk of bias) [17, 26–28].

In decreasing order of the studies that considered socioeconomic variables, 62.50% (25 of 40 studies) were evaluated as moderate (15 studies, 37.50%) or strong (10, 25.00%). Of the 34 which considered air pollutants, 41.18% (14 studies) were evaluated as moderate (9 studies, 26.47%) or strong (5 studies, 14.71%). Finally, of the 92 studies that considered meteorological variables, 25.00% (23 studies) were evaluated as moderate (11 studies, 11.96%) or strong (12 studies, 13.04%).

However, in the case of studies that consider socioeconomic variables, it should be noted that the high risk of bias could be underestimated. As is known, socioeconomic variables are contextual variables measured

at an ecological level in a geographic area and invariant over time. Their influence on COVID-19, if any, is highly unlikely to be non-linear. Consequently, in many cases this dimension was not evaluated.

The dimension in which we evaluated more studies with a high risk of bias was that of measurement errors (90 of 132 studies, 68.19%), followed by the control of the spatial and temporal dependence dimension (80 studies, 60.61%) and of the statistical model (77 studies, 58.33%) and control of non-linearities (73 studies, 55.30%) dimensions. The dimensions with fewer studies with a high risk of bias were confounding bias (47 studies, 35.61%) and control of the population (53 studies, 40.15%).

Findings from studies assessed as moderate or strong

In relation to the studies that considered meteorological variables, the ones that we evaluated as moderate or strong [28–48] have not consistently found an attenuating effect of meteorological variables. That is, they have not found that high temperature and low humidity were associated with lower incidence or mortality from COVID-19. In seven of 22 studies, temperature was either positively associated or not statistically associated with incidence [27, 30, 34, 35, 44], transmission (reproductive number) [46] and mortality [28] (four out of 11 studies were assessed as strong and another three out of 11 studies assessed as moderate). Among the studies that found an attenuating effect, five (three evaluated as strong [31, 38, 40] and one as moderate [41]) did not include lags and, therefore, assumed that the effect of the meteorological variables was contemporaneous. The studies that did include lags were evaluated with high risk of bias in some dimension. In particular, control of non-linearities [32, 39, 42, 45, 48], confounding bias [37, 43, 48], and measurement errors [37, 42, 47], followed by control of population [29, 36] and control of spatial and/or temporal dependence [33, 39]. Interestingly, Xie et al. [27], whose units of analysis were 122 Chinese cities, (a study that we evaluated as strong and did not have any dimension evaluated as high risk), points out that there is no evidence supporting that case counts of COVID-19 could decline when the weather becomes warmer.

There was very little evidence in relation to other meteorological variables such as wind speed (only two strong [33, 49] and one moderate [34] study analyzed it and found a negative association between wind speed and incidence); cloud percentage [29] or solar radiation [42] (both evaluated as moderate and with contradictory results: higher percentage of cloud was associated with higher incidence, while no association was found with solar radiation); or precipitation (considered in only one strong study that found a significant negative association with incidence [31]).

Table 2 Evaluation of bias for the studies in the systematic review

Manuscript	Confounding bias	Control of the population	Control of the spatial and/or temporal dependence	Control of non-linearities	Measurement errors	Statistical model	Overall rating
Aabed	3	2	3	1	3	1	3
Adekunle	3	3	1	1	2	3	3
Adhikari	2	3	2	3	2	1	2
Ahmadi	1	2	3	3	3	3	3
Azar	1	2	3	NA	2	1	1
Azuma	1	1	2	3	3	3	3
Behnood	3	2	3	1	3	3	3
Briz-Redón	1	1	1	1	3	3	2
Byass	3	1	3	3	3	1	3
Carleton	3	1	2	1	3	3	3
Chadeau-Hyam	1	3	2	3	2	1	2
Chakrabarty	3	3	3	1	3	1	3
Chakraborty	1	1	3	2	3	1	2
Chaudhry	1	2	3	NA	3	1	2
Chien	1	1	1	1	3	1	1
Coccia a	3	2	3	3	3	3	3
Coccia b	1	2	3	3	2	3	3
Coker	1	1	2	3	1	1	1
Das a	3	1	3	NA	3	1	3
Das b	1	3	3	NA	3	1	3
Demongeot	3	3	3	3	3	3	3
DiMaggio	1	1	1	NA	2	1	1
Dogan	3	3	3	3	2	3	3
Drefahl	1	3	3	NA	2	1	2
Falcão Sobral	2	2	3	3	3	3	3
Fattorini	3	3	3	3	3	3	3
Fazzini	3	3	3	3	3	3	3
Fiasca	3	2	3	3	3	3	3
Filippini	1	1	3	1	3	3	3
Fu	2	2	2	3	2	1	2
Guasp	2	2	3	3	3	3	3
Guo C	1	1	3	1	2	1	1
Guo XJ	3	3	3	3	3	3	3
Gupta	3	3	NA	3	3	3	3
Han	1	3	1	3	2	3	3
He	3	3	1	1	3	1	3
Hoang a	2	3	3	1	3	3	3
Hoang b	2	3	3	1	1	3	3
Hutter	2	2	3	3	3	1	3
Iqbal MM	3	3	NA	3	3	3	3
Iqbal N	2	3	3	1	3	3	3
Isaia	1	1	3	3	3	3	3
Islam ART	3	3	3	1	2	1	3
Islam N	2	1	3	2	3	1	2
Jamshidi	1	2	3	3	3	3	3
Jiang	3	3	3	3	3	1	3
Jüni	1	2	NA	3	2	2	1

Table 2 (continued)

Manuscript	Confounding bias	Control of the population	Control of the spatial and/or temporal dependence	Control of non-linearities	Measurement errors	Statistical model	Overall rating
Rehman	1	1	3	3	3	1	3
Richmond	1	1	3	NA	2	3	2
Rodríguez-Villamizar	1	1	2	3	3	1	2
Rozenfeld	1	3	3	NA	2	1	2
Rubin	1	3	3	1	2	1	2
Runkle	1	2	2	3	3	1	2
Saez	1	1	1	1	2	1	1
Sajadi	1	3	NA	3	3	3	3
Sánchez-Lorenzo	3	3	NA	3	3	3	3
Sannigrahi	1	2	3	NA	2	3	2
Sarkodie	1	3	1	3	2	3	3
Scarpone	1	1	1	1	2	3	1
Sehra	1	1	3	3	3	1	3
Shahzad F	3	3	3	3	3	3	3
Shahzad K	3	3	3	3	3	3	3
Shao	2	3	NA	3	3	1	3
Shi	3	1	2	1	2	1	1
Stieb	1	1	2	3	3	1	2
Su	1	3	NA	3	3	3	3
Sun	1	1	1	3	3	3	3
Tagaki a	2	1	3	3	3	3	3
Tagaki b	3	3	3	3	3	3	3
To	3	1	2	3	3	3	3
Tobías	2	1	2	3	2	1	1
Tzampoglou	1	1	3	3	3	3	3
Ujiie	1	1	3	3	3	1	3
Wang Q	1	1	2	1	3	1	1
Wang Y	3	3	3	1	3	3	3
Ward a	3	1	3	1	3	1	3
Ward b	3	1	3	1	3	1	3
Wu X	1	1	2	1	2	1	1
Wu Y	2	2	NA	1	3	3	2
Xie J	2	1	2	1	2	1	1
Xie Z	1	3	1	3	3	3	3
Xu	3	1	2	3	2	1	2
Yao a	1	3	3	3	3	2	3
Yao b	1	3	3	3	2	3	3
You	1	2	1	NA	3	3	2
Zakeri	1	2	3	NA	2	1	1
Zhang	2	3	3	1	3	3	3
Zhu L	3	3	2	1	2	3	3
Zhu Y	3	3	NA	3	3	2	3
Dimensions				Overall			
Number of 3 s	47	53	80	73	90	77	85
Number of 2 s	26	32	23	2	40	4	26
Number of 1 s	60	48	30	58	3	52	22

Bold: Strong (rate 1); *Italic:* Moderate (rate 2); **Bolditalic:** Weak (rate 3)

Greater consistency was found in the association between greater exposure to levels of air pollution, especially long-term exposure, and an increase in COVID-19 outcomes, both in ecological [17, 28, 29, 44, 48–51, 54–56] and individual studies [52, 53]. The areas that were most exposed to air pollution were those with the highest incidence (new daily cases, new positive tests, and cumulative cases) [17, 29, 44, 48, 49, 53, 54] and the highest mortality [17, 28, 29, 49–52, 55, 56] from COVID-19. This result occurs, above all, for fine particles, PM_{2.5} [28, 44, 50–56], but also for ozone, O₃ [29, 49, 50], coarse particles, PM₁₀ [17, 50], nitrogen dioxide, NO₂ [17, 50], benzene [55] and for an air quality index [48]. In Saez et al. [17] (which we evaluated as strong) as in Adhikari et al. [29] and Rodríguez-Villamizar et al. [56] (these last two evaluated as moderate), some of the pollutants were not found to be associated with mortality (PM₁₀ in Saez et al., O₃ in Adhikari et al., PM_{2.5} in Rodríguez-Villamizar et al.).

In relation to studies that considered socioeconomic variables, as we said, we must distinguish between the findings of ecological [17, 28, 54–66] and individual studies [53, 67–72]. In the ecological studies, there was no consistent association between socioeconomic contextual variables and COVID-19 outcomes. In just over half of the studies, the socioeconomic variables were risk factors and in the rest they were either protective factors or no statistically significant association was found. Even in some studies, such as Saez et al. [17] or Wu X et al. [28] (both of which we evaluated as strong and did not have any dimension evaluated with high risk of bias), apparently contradictory results were found. Thus, in Saez et al. [17], whose unit of analysis were small areas (counties and health zones, some made up of census tracts, others by municipalities) in Catalonia, Spain, the higher the percentage of poor housing in the small area and the more economically deprived the area was, the greater the risk of a positive result and death. Conversely, the higher the unemployment rate and the percentage of foreigners in the small area, the lower the risk of a positive result and death. In Wu et al. [28], whose units of analysis were US counties, while percent of the adult population with less than high school education and percent of Black residents, both in the county, were found to be positively associated with the number of deaths in the county, the median household income, the percentage of owner-occupied housing and, marginally, the median house value were also found positively associated. Meanwhile, others, such as the percentage of people in the county in poverty, were not found to be statistically significant associated.

More consistency has been found in relation to population density. In the areas with a higher population

density, there was a higher incidence, a higher number of positives, a higher transmission (measured by the reproductive number) and a higher number of deaths than in others less densely populated areas. In Wu et al. [28], however, the higher the population density, the lower the risk of mortality (although statistical significance only occurs in the fourth quintile).

Of the seven individual studies that we evaluated as moderate or strong, five found an association between both individual socioeconomic status (income, non-white ethnicity—especially Blacks-, lower educational attainment, being an immigrant from a low- or middle-income country) and contextual (income of the area, where the subject resided, residing in a neighborhood with financial insecurity) and various COVID-19 outcomes (positive tests, hospital admissions and deaths). We did, however, find one exception. In Price-Haywood et al. (a study that we evaluated as strong), whose study population was the Ochsner Health facility in New Orleans, Louisiana, USA, Black race was not associated with higher in-hospital mortality than white race, after adjustment for differences in sociodemographic and clinical characteristics on admission [70].

Discussion

Our results, both with regard to the methodological limitations that we found in the review and the results of the studies that control them, were similar to those of other reviews. Regarding the methodological limitations, we will refer, in order of publication, to two reviews (not systematic): one that considered air pollutants [73] and the other meteorological variables [74]. Villeneuve and Goldberg review six studies on COVID-19 (only two were peer-reviewed) and two on SARS, published up to May 2020 [74]. Hunter Kerr et al. review 43 studies (23 of them peer-reviewed), published in 2020 [74]. Both reviews found, as we did, that all studies have methodological limitations in one way or another. Almost all the methodological limitations that we have pointed out here were also considered in these two reviews. There are, however, some differences. Hunter Kerr et al. did not consider choosing a statistical model with normally distributed errors [74] as a limitation. Villeneuve and Goldberg, for their part, did not consider the error of the functional form (i.e., control of non-linearities), at least directly, inasmuch as they do so indirectly by pointing out, as a limitation, the inadequate evaluation of effect modification [73]. In contrast, Villeneuve and Goldberg point out, as the most important error, possible cross-level bias in ecological studies.

Regarding the influence of environmental variables (meteorological and air pollutants) in COVID-19 outcomes, the findings of the studies evaluated as moderate or strong in our review, coincided with the findings of the other reviews (both systematic and non-systematic).

We cannot conclude that there was an attenuating effect of weather conditions on the spread of the COVID-19 pandemic. In addition to the fact that, as mentioned, we did not find a systematic behaviour in the reviewed studies, so the attenuation shown by some of them could actually be a consequence of an inadequate adjustment. Thus, on the one hand, the study period of all the studies reviewed by the systematic reviews of Mecenas et al. [2], Smit et al. [3] and McClymont and Hu [4] as well as by the Hunter Kerr et al.'s review [74], corresponded to the first wave. The same occurs with most of the studies in our review (all except three). However, with a single exception [45], none of the studies controlled for non-pharmaceutical interventions either as containment or suppression strategies undertaken in that period. Thus, in this case, the reduction in the spread of the pandemic as temperature increased and humidity decreased, could have been confounded by the effects of lockdowns and other restrictions. Although Tobías and Molina [45] controlled for the effects of lockdown (and also those of seasonality as a consequence of weekends), they did not adjust for other confounders. Consequently, and perhaps for this reason, they found a significant effect only in the contemporary association (the same day) between an increase in temperature and a reduction in the incidence rate. We believe that, if they exist, the effect of meteorological variables on the spread of COVID-19 would be indirect. In the spring–summer of 2020, better weather conditions (higher temperature, lower relative humidity, lower wind speed, etc.) and a relaxation of restrictions, led to greater mobility and, therefore, greater social contact that, in turn, led to an increase in transmission and, consequently, in incidence. This was what happened, for example, in Spain during the second wave (which began in August 2020) [23].

The results of all reviews, including ours, suggest that there is an association between exposure to air pollutants (particularly in the long term but also in the short term) and COVID-19 outcomes. In fact, two hypotheses have been suggested that would explain this association. First, some studies have proposed that air particulate matter can operate as a virus carrier, promoting the spread of the SARS-CoV-2 [74–76]. It should be noted, however, that these studies were either not eligible as they used only correlation analysis to test their hypothesis [75] or they were eligible but were assessed as a high risk of bias [76].

A second hypothesis has been proposed which suggests there could be potential biological mechanisms that may explain the association between air pollutants and respiratory viral infections. According to this, the effects of exposure to air pollutants would occur not so much on transmission or incidence but on the worsening of the disease (hospitalization, ICU admissions, mortality). Exposure exacerbates the severity of COVID-19 infection symptoms and worsens the prognosis of COVID-19 patients [73]. In this sense, Wu X et al. [28] argue that long-term exposure to PM_{2.5} could cause alveolar angiotensin-converting enzyme 2 (ACE-2) receptor overexpression and impairs host defences [77]. This could cause a more severe form of COVID-19 in ACE-2–depleted lungs, increasing the likelihood of poor outcomes, including death [78]. We, however, believe that air pollutants have actually been surrogates of other variables, such as the mobility of residents and several socioeconomic conditions (high population density, poor housing, use of public transport, occupations in which it is not possible to telecommute, etc.) that facilitate social contact [17]. In fact, Dey and Dominici, in a very recent editorial commenting on the study by Wu et al. [28], and of which Dominici is a co-author, point out that the health risks of some racial subgroups are spiraling as they have higher levels of exposure to air pollutants, hence being more susceptible to mortality from COVID-19 [79]. We do not deny that exposure to air pollutants had an independent effect on, above all, the worsening of the disease among those diagnosed with COVID-19. However, we are convinced that this effect cannot be observed using an ecological design.

As we noted, we have found a consistency in the effects of socioeconomic variables on COVID-19 outcomes only in individual studies and in indicators also at the individual level (ethnicity—particularly being Black—education, etc.). We believe that the effect, if it exists, would be indirect. Poorer socioeconomic conditions would be associated, on the one hand, with greater social contact, which would affect the transmission of the virus and the incidence of COVID-19 and, on the other, with a greater number of comorbidities and greater difficulties in accessing health care which would affect a poorer prognosis of the disease. Furthermore, poorer socioeconomic conditions could be related both to a differential exposure to air pollution and to a differential susceptibility to its effects (i.e., modification of the effect) [80].

In short, a large part of the methodological problems that we have encountered and, therefore, of the uncertainty in the findings, are the consequence of using an ecological design. In this sense, we could not agree more with Hunter Kerr et al. [74], who recommend, as an

epidemiological design, a longitudinal study with individual-level data, in which those diagnosed with COVID-19 would be followed through time.

Our study may have three limitations. First, some studies published during 2020 may have escaped us. That said, this is unlikely, since, as of January 2021, we have been regularly reviewing PubMed and periodically reviewing the other databases. Nevertheless, it is not impossible that a study may have eluded us. Second, both the information extraction and the quality control we carried out could have some subjectivity. We have tried to minimize this as much as possible.

Finally, as we noted, the rating of both the domains and the studies are based on Parmar et al. [8], with the only difference being that in Parmar et al., an overall rating of strong was given if none of its domains was rated as weak. In our case, this assignment seemed too restrictive. In fact, applying this criterion would imply that only one of the studies could be rated as strong. In our case, we observed some biases that were not contemplated in Parmar et al., such as the lack of control of the population and of the spatial and/or temporal dependence, the non-control of non-linearity and the inappropriate use of statistical models. In our case, the probability that at least one of these biases occurred was very high. In any case, we admit that there could be some degree of arbitrariness in the assignment of the overall rating to one category or another.

Conclusions

All the studies we reviewed have methodological limitations to a greater or lesser extent. Even those that we have evaluated as strong (16.67% of the studies reviewed) and, among them, those in which we did not evaluate any dimension as having a high risk of bias (4 studies), have the limitation of using an ecological epidemiological design or, in any case, either of measuring the exposure in an ecological way (exposure misclassification). These limitations prevent conclusions about the effects of environmental (meteorological and air pollutants) and socioeconomic variables on COVID-19 outcomes being drawn. However, we dare to argue that the effects of these variables, if they exist, would be indirect, based on their relationship with social contact. In any case, the estimation of these independent effects requires the use of an individual design and the control of the methodological limitations explained in this work. Among them, an estimate of individual exposure free of biases (non-differential misclassification, non-existence of spatial-temporal misalignment, etc.).

Abbreviations

COVID-19: Novel coronavirus disease; PM_{2.5}: Fine particles with a diameter of 2.5 microns (µm) or less; NO₂: Nitrogen dioxide; PM₁₀: Coarse particles with a diameter of 10 µm or less; SARS-CoV-2: Severe acute respiratory syndrome coronavirus 2; PM: Particulate matter; PROSPERO: Prospective Register of Systematic Reviews; PRISMA: Preferred reporting items for systematic reviews and meta-analysis; RTI: Item Bank for Assessment of Risk of Bias and Precision for Observational Studies of Interventions or Exposures; PCR: Polymerase chain reaction; MAUP: Modifiable areal unit problem; MTUP: Modifiable temporal unit problem.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12302-021-00550-7>.

Additional file 1: Table S1. List of studies excluded. **Table S2.** Studies included in the qualitative synthesis **Table S3.** List of studies included in the qualitative synthesis. **Figure S1.** Number of studies by type of explanatory variable analyzed

Acknowledgements

This study was carried out within the 'Cohort-Real World Data' subprogram of CIBER of Epidemiology and Public Health (CIBERESP).

Authors' contributions

MS had the original idea for the paper. MS designed the study. The bibliographic search and the writing of the introduction were carried out by MS and MAB. The methods were chosen and performed by all authors. MAB created the tables and figures. All authors wrote the results and the discussion. The writing and final editing was done by all authors. All authors reviewed and approved the manuscript.

Funding

This work was partially financed by the SUPERA COVID19 Fund from SAUN: Santander Universidades, CRUE and CSIC, and by the COVID-19 Competitive Grant Program from Pfizer Global Medical Grants. The funding sources did not participate in the design or conduct of the study, the collection, management, analysis, or interpretation of the data, or the preparation, review, or approval of the manuscript.

Availability of data and materials

All the studies, as well as the code to make the figures, can be requested from the corresponding author (marc.saez@udg.edu).

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interest

The authors declare that they have no competing interests.

Author details

¹Research Group On Statistics, Econometrics and Health (GRECS), and CIBER of Epidemiology and Public Health (CIBERESP), University of Girona, Carrer de la Universitat de Girona 10, Campus de Montilivi, 17003 Girona, Spain. ²CIBER of Epidemiology and Public Health (CIBERESP), Madrid, Spain.

Received: 26 June 2021 Accepted: 3 September 2021
Published online: 10 September 2021

References

1. World Health Organization (WHO) Director-General's opening remarks at the media briefing on COVID-19. <https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020>. Accessed 2 Apr 2021.
2. Mecenas P, Bastos RTDRM, Vallinoto ACR, Normando D (2020) Effects of temperatura and humidity on the spread of COVID-19: a systematic review. *PLoS ONE* 15(9):e0238339. <https://doi.org/10.1371/journal.pone0238339>
3. Smit AJ, Fitchett JM, Engelbrecht FA, Scholes RJ, Dzhivhuho G, Sweijd NA (2020) Winter is coming: a southern hemisphere perspective of the environmental drivers of SARS-CoV-2 and the potential seasonality of COVID-19. *Int J Environ Res Public Health* 17:5634. <https://doi.org/10.3390/ijerph17165634>
4. McClymont H, Hu W (2021) Weather variability and COVID-19 transmission: a review of recent research. *Int J Environ Res Public Health* 18(2):396. <https://doi.org/10.3390/ijerph18020396>
5. Copat C, Cristaldi A, Fiore M, Grasso A, Zuccarello P, Signorelli SS, Conti GO, Ferrante M (2020) The role of air pollution (PM and NO₂) in COVID-19 spread and lethality: a systematic review. *Environ Res* 191:110129. <https://doi.org/10.1016/j.envres.2020.110129>
6. Maleki M, Anvari E, Hopke PK, Noorimotlagh Z, Mirzaee SA (2021) An updated systematic review on the association between atmospheric particulate matter pollution and prevalence of SARS-CoV-2. *Environ Res* 195:110898. <https://doi.org/10.1016/j.envres.2021.110898>
7. Page MJ, McKenzie JR, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, Chou R, Glanville J, Grimshaw JM, Hróbjartsson A, Lalu MM, Li T, Lode EW, Mayo-Wilson E, McDonald S, McGuinness LA, Thomas J, Tricco AC, Welch VA, Whiting P, Moher D (2021) The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 372:n71. <https://doi.org/10.1136/bmj.n71>
8. Parmar D, Stavropoulou C, Ioannidis JP (2016) Health outcomes during the 2008 financial crisis in Europe: systematic literature review. *BMJ* 354:i4588. <https://doi.org/10.1136/bmj.i4588>
9. Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Losos M, Tugwell P. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp. Accessed 3 Apr 2021
10. Viswanathan M, Berkman ND (2012) Development of the RTI item bank on risk of bias and precision of observational studies. *J Clin Epidemiol* 65:163–178. <https://doi.org/10.1016/j.jclinepi.2011.05.008>
11. Saez M, Barceló MA, Saurina C, Cabrera A, Daponte A (2019) Evaluation of the biases in the studies that assess the effects of the Great Recession on health. A systematic review. *Int J Environ Res Public Health* 16(14):2479. <https://doi.org/10.3390/ijerph16142479>
12. Greene WH (2018) *Econometric analysis*, chapter 5, 8th edn. Pearson, Boston, London
13. Domingo JL, Marquès M, Rovira J (2020) Influence of airborne transmission of SARS-CoV-2 on COVID-19 pandemic. a review. *Environ Res* 188:109861. <https://doi.org/10.1016/j.envres.2020.109861>
14. Barceló MA, Saez M, Cano-Serral G, Martínez-Beneito MA, Martínez JM, Borrell C, Ocaña-Riola R, Montoya I, Calvo M, López-Abente G, Rodríguez-Sanz M, Toro S, Alcalá JT, Saurina C, Sánchez-Villegas P, Figueiras A (2008) Methods to smooth mortality indicators: application to analysis of inequalities in mortality in Spanish cities (the MEDEA Project) [in Spanish]. *Gac Sanit* 22(6):596–608. [https://doi.org/10.1016/s0213-9111\(08\)75362-7](https://doi.org/10.1016/s0213-9111(08)75362-7)
15. Open Government. Generalitat de Catalunya. Open data and COVID-19. http://governobert.gencat.cat/en/dades_obertes/dades-obertes-covid-19/index.html. Accessed 1 May 2021.
16. Secretaría General de Sanidad. Dirección General de Salud Pública, Calidad e Innovación. Ministerio de Sanidad. Gobierno de España. [in Spanish]. <https://www.msbs.gob.es/en/profesionales/saludPublica/ccayes/alertasActual/nCov-China/situacionActual.htm>. Accessed 1 May 2021.
17. Saez M, Tobias A, Barceló MA (2020) Effects of long-term exposure to air pollutants on the spatial spread of COVID-19 in Catalonia. *Spain Environ Res* 191:110177. <https://doi.org/10.1016/j.envres.2020.110177>
18. Elliott P, Savitz DA (2008) Design issues in small-area studies of environment and health. *Environ Health Perspect* 116(8):1098–1104. <https://doi.org/10.1289/ehp.10817>
19. Wannemuehler K, Lyles R, Waller L, Hoekstra R, Klein M, Tolbert P (2009) A conditional expectation approach for associating ambient air pollutant exposures with health outcomes. *Environmetrics* 20(7):877–894. <https://doi.org/10.1002/env.978>
20. Openshaw S (1984) *Concepts and techniques in modern geography* no. 38: the modifiable areal unit problem. Geo Books, Norwich
21. Helbich M, Mute Browning MHE, Kwan MP (2021) Time to address the spatiotemporal uncertainties in COVID-19 research: concerns and challenges. *Sci Total Environ* 764:142866. <https://doi.org/10.1016/j.scitotenv.2020.142866>
22. Wang Y, Di Q (2020) Modifiable areal unit problem and environmental factors of COVID-19 outbreak. *Sci Total Environ* 740:139984. <https://doi.org/10.1016/j.scitotenv.2020.139984>
23. Ribas V, Miralles F, Rey O, Rafael X, Subías P, Torrent M, Vicens JA, Saez M, Barceló MA, Ponce-de-León M, Valencia A, Arenas A, Saura P (2021) *Big Data i Intel·ligència Artificial per a la prevenció d'epidèmies. Monitoratge i predicció per a la detecció primerenca de brots epidèmics* [in Catalan]. Barcelona: Generalitat de Catalunya.
24. Cheng T, Adepeju M (2014) Modifiable temporal unit problem (MTUP) and its effect on space-time cluster detection. *PLoS ONE* 9:e100465. <https://doi.org/10.1371/journal.pone.0100465>
25. Filippini T, Rothman KJ, Goffi A, Ferrari F, Maffei G, Orsini N, Vinceti M (2020) Satellite-detected tropospheric nitrogen dioxide and spread of SARS-CoV-2 infection in Northern Italy. *Sci Total Environ* 739:140278. <https://doi.org/10.1016/j.scitotenv.2020.140278>
26. DiMaggio C, Klein M, Berry C, Frangos S (2020) Black/African American Communities are at highest risk of COVID-19: spatial modeling of New York City ZIP Code-level testing results. *Ann Epidemiol* 51:7–13. <https://doi.org/10.1016/j.annepidem.2020.08.012>
27. Xie J, Zhu Y (2020) Association between ambient temperature and COVID-19 infection in 122 cities from China. *Sci Total Environ* 724:138201. <https://doi.org/10.1016/j.scitotenv.2020.138201>
28. Wu X, Nethery RC, Sabath MB, Braun D, Dominici F (2020) Air pollution and COVID-19 mortality in the United States: Strengths and limitations of an ecological regression analysis. *Sci Adv* 6(45):eabd4049. <https://doi.org/10.1126/sciadv.abd4049>
29. Adhikari A, Yin J (2020) Short-term effects of ambient ozone, PM_{2.5}, and meteorological factors on COVID-19 confirmed cases and deaths in Queens, New York. *Int J Environ Res Public Health* 17(11):4047. <https://doi.org/10.3390/ijerph17114047>
30. Briz-Redón A, Serrano-Aroca A (2020) A spatio-temporal analysis for exploring the effect of temperature on COVID-19 early evolution in Spain. *Sci Total Environ* 728:138811. <https://doi.org/10.1016/j.scitotenv.2020.138811>
31. Chien LC, Chen LW (2020) Meteorological impacts on the incidence of COVID-19 in the U.S. *Stoch Environ Res Risk Assess.* 4:1–6. <https://doi.org/10.1007/s00477-020-01835-8>
32. Fu S, Wang B, Zhou J, Xu X, Liu J, Ma U, Li L, He X, Li S, Niu J, Luo B, Zhang K (2021) Meteorological factors, governmental responses and COVID-19: evidence from four European countries. *Environ Res* 194:110596. <https://doi.org/10.1016/j.envres.2020.110596> (Epub 2020 Dec 9)
33. Guo C, Bo Y, Changqing L, Li HB, Zeng Y, Zhang Y, Hossain S, Chan JWM, Yeung DW, Kwok KO, Wong SYS, Lau AKH, Lao XQ (2021) Meteorological factors and COVID-19 incidence in 190 countries: An observational study. *Sci Total Environ* 757:143783. <https://doi.org/10.1016/j.scitotenv.2020.143783> (Epub 2020 Nov 23)
34. Islam N, Bukhari Q, Jameel Y, Shabnam S, Erzurumluoglu AM, Siddique MA, Massaro JM, D'Agostino RB (2021) COVID-19 and climatic factors: a global analysis. *Environ Res* 193:110355. <https://doi.org/10.1016/j.envres.2020.110355> (Epub 2020 Oct 28)
35. Jüni P, Rothenbühler M, Bobos P, Thorpe KE, da Costa BR, Fisman DN, Slutsky AS, Gesink D (2020) impact of climate and public health interventions on the COVID-19 pandemic: a prospective cohort study. *CMAJ* 192(21):E566–E573. <https://doi.org/10.1503/cmaj.200920>
36. Liu J, Zhou J, Yao J, Zhang X, Li L, Xiaocheng X, He W, Wang B, Fu S, Niu T, Yan J, Shi Y, Ren X, Niu J, Zhu W, Li S, Luo B, Zhang L (2020) Impact of meteorological factors on the COVID-19 transmission: a multi-city study in China. *Sci Total Environ* 726:138513. <https://doi.org/10.1016/j.scitotenv.2020.138513>

37. Ma Y, Zhao Y, Liu J, He X, Wang B, Fu S, Yan J, Niu J, Zhou J, Luo B (2020) Effects of temperature variation and humidity on the death of COVID-19 in Wuhan, China. *Sci Total Environ* 724:138226. <https://doi.org/10.1016/j.scitotenv.2020.138226>
38. Meyer A, Sadler R, Faverjon C, Cameron AR, Bannister-Tyrrell M (2020) Evidence that higher temperature are associated with a marginally lower incidence of COVID-19 cases. *Front Public Health* 8:367. <https://doi.org/10.3389/fpubh.2020.00367>
39. Pequeno P, Mendel B, Rosa C, Bosholn M, Souza JL, Baccaro F, Barbosa R, Magnusson W (2020) Air transportation, population density and temperature predict the spread of COVID-19 in Brazil. *PeerJ* 8:e9322. <https://doi.org/10.7717/peerj.9322>
40. Prata DN, Rodrigues W, Bermejo PH (2020) Temperature significantly changes COVID-19 transmission in (sub)tropical cities of Brazil. *Sci Total Environ* 729:138862. <https://doi.org/10.1016/j.scitotenv.2020.138862>
41. Qi H, Xiao S, Shi R, Ward MP, Chen Y, Tu W, Su Q, Wang W, Wang X, Zhang Z (2020) COVID-19 transmission in Mainland China is associated with temperature and humidity: a time-series analysis. *Sci Total Environ* 728:138778. <https://doi.org/10.1016/j.scitotenv.2020.138778>
42. Runkle JD, Sugg MM, Leeper RD, Rao Y, Matthews JL, Rennie JJ (2020) Short-term effects of specific humidity and temperature on COVID-19 morbidity in select US cities. *Sci Total Environ* 740:140093. <https://doi.org/10.1016/j.scitotenv.2020.140093>
43. Shi P, Dong Y, Yan H, Zhao C, Li X, Liu W, He M, Tang S, Xi S (2020) Impact of temperature on the dynamics of the COVID-19 outbreak in China. *Sci Total Environ* 728:138890. <https://doi.org/10.1016/j.scitotenv.2020.138890>
44. Stieb DM, Evans GJ, To TM, Brook JR, Burnett RT (2020) An ecological analysis of long-term exposure to PM_{2.5} and incidence of COVID-19 in Canadian health regions. *Environ Res* 191:110052. <https://doi.org/10.1016/j.envres.2020.110052>
45. Tobias A, Molina T (2020) Is temperature reducing the transmission of COVID-19? *Environ Res* 186:109553. <https://doi.org/10.1016/j.envres.2020.109553>
46. Wang Q, Zhao Y, Zhang Y, Qiu J, Li J, Yan N, Li N, Zhang J, Tian D, Sha X, Jing J, Yang C, Wang K, Xu R, Zhang Y, Yang H, Zhao S, Zhao Y (2021) Could the ambient higher temperature decrease the transmissibility of COVID-19 in China? *Environ Res* 193:110576. <https://doi.org/10.1016/j.envres.2020.110576> (**Epub 2020 Dec 3**)
47. Wu Y, Jing W, Liu J, Ma Q, Yuan J, Wang Y, Du M, Liu M (2020) Effects of temperature and humidity on the daily new cases and new deaths of COVID-19 in 166 countries. *Sci Total Environ* 729:139051. <https://doi.org/10.1016/j.scitotenv.2020.139051>
48. Xu H, Yan C, Fu Q, Xiao K, Yu Y, Han D, Wang W, Cheng J (2020) Possible environmental effects on the spread of COVID-19 in China. *Sci Total Environ* 731:139211. <https://doi.org/10.1016/j.scitotenv.2020.139211>
49. Coccia M (2021) How do low wind speeds and high levels of air pollution support the spread of COVID-19? *Atmos Pollut Res* 12(1):437–445. <https://doi.org/10.1016/j.apr.2020.10.002> (**Epub 2020 Oct 7**)
50. Liang D, Shi L, Zhao J, Liu P, Sarnat JA, Gao S, Schwartz J, Liu Y, Ebelt ST, Scovronick N, Chang HH (2020) Urban air pollution may enhance COVID-19 case-fatality and mortality rates in the United States. *Innovation (NY)* 1(3):100047. <https://doi.org/10.1016/j.xinn.2020.100047>
51. Pozzer A, Dominici F, Haines A, Witt C, Münzel T, Lelieveld J (2020) Regional and global contributions of air pollution to risk of death from COVID-19. *Cardiovasc Res* 116(14):2247–2253. <https://doi.org/10.1093/cvr/cvaa288>
52. López-Feldman A, Heres D, Márquez-Padilla F (2021) Air pollution exposure and COVID-19: a look at mortality in Mexico City using individual-level data. *Sci Total Environ* 756:143929. <https://doi.org/10.1016/j.scitotenv.2020.143929> (**Epub 2020 Nov 26**)
53. Chadeau-Hyam M, Bodinier B, Elliott J, Whitaker MD, Tzoulaki I, Vermeulen R, Kelly-Irving M, Delpierre C, Elliott P (2020) Risk factors for positive and negative COVID-19 tests: a cautious and in-depth analysis of UK biobank data. *Int J Epidemiol* 49(5):1454–1467. <https://doi.org/10.1093/ije/dyaa134>
54. Chakrabarty RK, Beeler P, Liu P, Gooswami S, Harvey RD, Pervez S, van Donkelaar A, Martin RV (2021) Ambient PM_{2.5} exposure and rapid spread of COVID-19 in the United States. *Sci Total Environ* 760:143391. <https://doi.org/10.1016/j.scitotenv.2020.143391> (**Epub 2020 Nov 9**)
55. Luo Y, Yan J, McClure S, (2021) Distribution of the environmental and socioeconomic risk factors on COVID-19 death rate across continental USA: a spatial nonlinear analysis. *Environ Sci Pollut Res Int* 28(6):6587–6599. <https://doi.org/10.1007/s11356-020-10962-2> (**Epub 2020 Oct 1**)
56. Rodríguez-Villamizar L, Belalcázar-Cerón LC, Fernández-Niño JA, Marín-Pineda DM, Rojas-Sánchez O, Acuña-Merchán LA, Ramírez-García N, Mangones-Matos SC, Vargas-González JN, Herrera-Torres J, Agudelo-Castañeda DM, Piñeros-Jiménez JG, Rojas-Roa NY, Herrera-Galindo M (2021) Air pollution, sociodemographic and health conditions effects on COVID-19 mortality in Colombia: an ecological study. *Sci Total Environ* 756:144020. <https://doi.org/10.1016/j.scitotenv.2020.144020> (**Epub 2020 Nov 26**)
57. Chaudhry R, Dranitsaris G, Mubashir T, Bartoszko J, Riaz S (2020) A country level analysis measuring the impact of government actions, country preparedness and socioeconomic factors on COVID-19 mortality and related health outcomes. *EClinicalMedicine* 25:100464. <https://doi.org/10.1016/j.eclinm.2020.100464>
58. Kaiser JC, Stathopoulos GT (2020) Socioeconomic correlates of SARS-CoV-2 and influenza H1N1 outbreaks. *Eur Respir J* 56(3):2001400. <https://doi.org/10.1183/13993003.01400-2020>
59. Lamb MR, Kandula S, Shaman J (2021) Differential COVID-19 case positivity in New York City neighborhoods: socioeconomic factors and mobility. *Influenza Other Respir Viruses* 15(2):209–217. <https://doi.org/10.1111/irv.12816> (**Epub 2020 Oct 14**)
60. Madhav KC, Oral E, Straif-Bourgeois S, Rung AL, Peters ES (2020) The effect of area deprivation on COVID-19 risk in Louisiana. *PLoS ONE* 15(12):e0243028. <https://doi.org/10.1371/journal.pone.0243028>
61. Plümper T, Neumayer E (2020) The pandemic predominantly hits poor neighbourhoods? SARS-CoV-2 infections and COVID-19 fatalities in German districts. *Eur J Public Health* 30(6):1176–1180. <https://doi.org/10.1093/eurpub/ckaa168>
62. Richmond HL, Tome J, Rochani H, Fung CH, Shah GH, Schwind JS (2020) The use of penalized regression analysis to identify county-level demographic and socioeconomic variables predictive of increased COVID-19 cumulative case rates in the state of Georgia. *Int J Environ Res Public Health* 17(21):8036. <https://doi.org/10.3390/ijerph17218036>
63. Rubin D, Huang J, Fisher BT, Gasparirini A, Tam V, Song L, Wang X, Kaufman J, Fitzpatrick K, Jain A, Griffis H, Cramer K, Morris J, Tasian G (2020) Association of social distancing, population density, and temperature with the instantaneous reproduction number of SARS-CoV-2 in counties across the United States. *JAMA Netw Open* 3(7):e2016099. <https://doi.org/10.1001/jamanetworkopen.2020.16099>
64. Sannigrahi S, Pilla F, Basu B, Basu AS, Molter A (2020) Examining the association between socio-demographic composition and COVID-29 fatalities in the European region using spatial regression approach. *Sustain Cities Soc* 62:102418. <https://doi.org/10.1016/j.scs.2020.102418>
65. Scarpone C, Brinkmann ST, Große T, Sonnenwald D, Fuchs M, Byron WB (2020) A multimethod approach for county-scale geospatial analysis of emerging infectious diseases: a cross-sectional case study of COVID-19 incidence in Germany. *Int J Health Geogr* 19(1):32. <https://doi.org/10.1186/s12942-020-00225-1>
66. You H, Wu X, Guo X (2020) Distribution of COVID-19 morbidity rate in association with social and economic factors in Wuhan, China: Implications for urban development. *Int J Environ Res Public Health* 17(10):3417. <https://doi.org/10.3390/ijerph17103417>
67. Azar KMJ, Shen Z, Romanelli RJ, Lockhart SH, Smits K, Robinson S, Brown S, Pressman AR (2020) Disparities in outcomes among COVID-19 patients in a large health care system in California. *Health Aff (Millwood)* 39(7):1253–1262. <https://doi.org/10.1377/hlthaff.2020.00598>
68. Drefahl S, Wallace M, Mussino E, Aradhya S, Kolk M, Brandén M, Malmberg B, Andersson G (2020) A population-based cohort study of socio-demographic risk factors for COVID-19 deaths in Sweden. *Nat Commun* 11(1):5097. <https://doi.org/10.1038/s41467-020-18926-3>
69. Marcielde Souza W, Fletcher Buss L, da Silva Candido D, Carrera JP, Li S, Zarebski AE, Moraes Pereira RH, Prete CA, de Souza-Santos AA, Parag KV, Belotti MC, Vincenti-González MF, Messina J, da Silva Sales FC, Dos Santos Andrade P, Heloiz Nascimento V, Ghilardi F, Abade L, Gutiérrez B, Kraemer MUG, Braga CKV, Santana Aguiar R, Alexander N, Mayaud P, Brady OJ, Marcilio I, Gouveia N, Li G, Tami A, Barbosa de Oliveira S, Gomes Porto VB, Ganem F, Ferreirade Almeida WA, Sutile Tardetti Fantinato FF, Marques Macário E, Kleberde Oliveira W, Nogueira ML, Pybus OG, Wu CH, Croda J, Sabino EC, Rodrigues Faria N (2020) Epidemiological and

- clinical characteristics of the COVID-19 epidemic in Brazil. *Nat Hum Behav* 4(8):856–865. <https://doi.org/10.1038/s41562-020-0928-4>
70. Price-Haywood EG, Burton H, Fort D, Seoane L (2020) Hospitalization and mortality among Black patients and White patients with COVID-19. *N Engl J Med* 382(26):2534–2543. <https://doi.org/10.1056/NEJMsa2011686>
 71. Rozenfeld Y, Beam J, Maier H, Haggerson W, Boudreau K, Carlson J, Meadows R (2020) A model of disparities: risk factors associated with COVID-19 infection. *Int J Equity Health*. 19(1):126. <https://doi.org/10.1186/s12939-020-01242-z>
 72. Zakery R, Bendayan R, Ashworth M, Bean DM, Dodbaba H, Durbaba S, O’Gallagher K, Palmmer C, Curcin V, Aitken E, Bernal W, Barker RD, Norton S, Gulliford M, Teo JTH, Galloway J, Dobson RJB, Shah AM (2020) A case-control and cohort study to determine the relationship between ethnic background and severe COVID-19. *EClinicalMedicine* 28:100574. <https://doi.org/10.1016/j.eclinm.2020.100574>
 73. Villeneuve PJ, Goldberg MS (2020) Methodological considerations for epidemiological studies of air pollution and the SARS and COVID-19 coronavirus outbreaks. *Environ Health Perspect* 128(9):95001. <https://doi.org/10.1289/EHP7411>
 74. Hunter Kerr G, Badr HS, Gardner LM, Pérez-Saez J, Zaitchik BF (2021) Associations between meteorology and COVID-19 in early studies: Inconsistencies, uncertainties, and recommendations. *One Health* 12:100225. <https://doi.org/10.1016/j.onehlt.2021.100225>
 75. Bashir MF, Ma BB, Komal B, Bashir MA, Tan D, Bashir M (2020) Correlation between climate indicators and COVID-19 pandemic in New York, USA. *Sci Total Environ* 728:138835. <https://doi.org/10.1016/j.scitotenv.2020.138835>
 76. Fattorini D, Regoli F (2020) Role of the chronic air pollution levels in the COVID-19 outbreak risk in Italy. *Environ Pollut* 264:114732. <https://doi.org/10.1016/j.envpol.2020.114732>
 77. Miyashita L, Foley G, Semple S, Grigg J (2020) Traffic-derived particulate matter and angiotensin-converting enzyme 2 expression in human airway epithelial cells. *bioRxiv*. <https://doi.org/10.1101/2020.05.15.097501v2>
 78. Frontera A, Cianfanelli L, Vlachos K, Landoni G, Cremona G (2020) Severe air pollution links to higher mortality in COVID-19 patients: the “double-hit” hypothesis. *J Infect* 81:255–259. <https://doi.org/10.1016/j.jinf.2020.05.031>
 79. Dey T, Dominici F (2021) COVID-19, air pollution, and racial inequity: Connecting the dots. *Chem Res Toxicol* 34(3):669–671. <https://doi.org/10.1021/acs.chemrestox.0c00432>
 80. Saez M, López-Casasnovas G (2019) Assessing the effects on health inequalities of differential exposure and differential susceptibility of air pollution and environmental noise in Barcelona, 2007–2014. *Int J Environ Res Public Health* 16(18):3470. <https://doi.org/10.3390/ijerph16183470>
 81. Martorell-Marugán J, Villatoro-García JA, et al. DatAC: a visual analytics platform to explore climate and air quality indicators associated with the COVID-19 pandemic in Spain. *Sci Total Environ* 2020; 750: 141424. <https://doi.org/10.1016/j.scitotenv.2020.141424>
 82. DatAC: Data against COVID-19 [in Spanish]. <https://covid19.genyo.es>. Accessed 23 Apr 2021.

Publisher’s Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
