Infectious Disease Modelling 8 (2023) 445–457

Contents lists available at ScienceDirect

Infectious Disease Modelling

journal homepage: www.keaipublishing.com/idm

SARS-CoV-2: Air pollution highly correlated to the increase in mortality. The case of Guadalajara, Jalisco, México

Elizabeth Torres-Anguiano [a, b, 1], Itzel Sánchez-López [a, b, 1], Angeles Garduno-Robles [a, 1], Jorge David Rivas-Carrillo [c], Edgar Alfonso Rivera-León [d], Sergio Sánchez-Enríquez [e], Luis Fernando Ornelas-Hernández [a], Fernando Zazueta León-Quintero [a, b], Eduardo Narciso Salazar León-Quintero [h], Guillermo Enrique Juárez-López [a, b], Fernando Antonio Sánchez-Zubieta [f, g], Mariana Ochoa-Bru [a, b], Abraham Zepeda-Moreno [a, b, f, *]

[a] Onkogenetik/Mexicana de Investigación y Biotectogía S.A. de C.V, Guadalajara, Jalisco, Mexico
[b] Unidad de Biología Molecular, Investigación y Diagnóstico, Hospital San Javier, Guadalajara, Jalisco, Mexico
[c] Departamento de Fisiología, Centro Universitario de Ciencias de la Salud, Universidad de Guadalajara, Guadalajara, Jalisco, Mexico
[d] Departamento de Biología Molecular y Genómica, Centro Universitario de Ciencias de la Salud, Universidad de Guadalajara, Guadalajara, Jalisco, Mexico
[e] Departamento de Clínicas, Centro Universitario de los Altos, Universidad de Guadalajara, Tepatitlán de Morelos, Jalisco, Mexico
[f] Departamento de Clínicas Médicas, Centro Universitario de Ciencias de la Salud, Universidad de Guadalajara, Guadalajara, Jalisco, Mexico
[g] Servicio de Hemato-Oncología Pediátrica, Hospital Civil de Guadalajara "Dr. Juan I. Menchaca", Guadalajara, Jalisco, Mexico
[h] Hospital Civil de Guadalajara "Fray Antonio Alcalde", Guadalajara, Jalisco, Mexico

ARTICLE INFO

Article history:
Received 8 November 2022
Received in revised form 5 January 2023
Accepted 10 April 2023
Available online 20 April 2023
Handling Editor: Dr Daihai He

Keywords:
Air pollution
COVID-19
Guadalajara
Mexico
SARS-CoV-2
SARS-CoV-2 lineages

ABSTRACT

Objectives: To determine whether air pollution or changes in SARS-CoV-2 lineages lead to an increase in mortality.

Methods: Descriptive statistics were used to calculate rates of infection (2020–2021). RT–PCR was used to compare viral loads from October 2020 to February 2021. Next-generation sequencing (NGS) (n = 92) was used to examine and phylogenetically map SARS-CoV-2 lineages. A correlative "air pollution/temperature" index (I) was developed using regression analysis. $PM_{2.5}$, PM_{10}, O_3, NO_2, SO_2, and CO concentrations were analyzed and compared to the mortality.

Results: The mortality rate during the last year was ~32%. Relative SARS-CoV-2 viral loads increased in December 2020 and January 2021. NGS revealed that approximately 80% of SARS-CoV-2 linages were B.1.243 (33.7%), B1.1.222 (11.2%), B.1.1 (9%), B.1 (7%), B.1.1.159 (7%), and B.1.2 (7%). Two periods were analyzed, the prehigh- and high-mortality periods and no significant lineage differences or new lineages were found. Positive correlations of air pollution/temperature index values with mortality were found for $IPM_{2.5}$ and IPM_{10}. INO_2. ISO_2, and ICO but not for O_3. Using ICO, we developed a model to predict mortality with an estimated variation of ~±5 deaths per day.

* Corresponding author. Onkogenetik/Mexicana de Investigación y Biotectogía SA de CV Pablo Casals 640. Col. Prados Providencia, Guadalajara, Jalisco, C.P. 44670, Mexico.
E-mail address: abraham.zepeda@mxaib.com (A. Zepeda-Moreno).
Peer review under responsibility of KeAi Communications Co., Ltd.
[1] These authors have contributed equally to this work.

https://doi.org/10.1016/j.idm.2023.04.004

Conclusion: The mortality rate in the MZG was highly correlated with air pollution indices and not with SARS-CoV-2 lineage.

1. Introduction

The metropolitan zone of Guadalajara (MZG) has approximately 5.2 million habitants distributed in ten municipalities that cover approximately 3000 km^2 (Gobierno del Estado de Jalisco, 2021). During the last year of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) pandemic, the rates of infection and mortality remained stable until December 2020 and January 2021. (COVID-19 Tablero México) During these two months, mortality increased markedly, leading to constraints in health care services and supplemental oxygen resources. Many factors, such as preventive actions and restrictions, did not reduce or change the rate of infection during 2020. Additionally, mortality started to increase during the second half of December for no apparent reason. Approximately two months later, the mortality and infection rates were notably reduced, achieving rates even lower than those in the previous six months. (COVID-19 Tablero México).

This mortality pattern in the MZG prompted the idea that environmental factors should be considered targets for preventive actions in the future. Therefore, we examined the rate of infection and the relative quantity of SARS-CoV-2 viral copies in samples collected from the general population without selection criteria, correlated air pollution concentrations with mortality, and comparatively examined SARS-CoV-2 lineages before and during the period of high mortality. With this data, we constructed a predictive model for mortality.

2. Materials and methods

2.1. Total RNA extraction

RNA from nasopharyngeal and oropharyngeal swabs was extracted using the MagaBio plus Virus DNA/RNA Purification Kit (BioFlux cat. #BSC71L1E) and the automated nucleic acid extraction system GenePure Pro 96T (Bioer Technology Co., Ltd.) according to the manufacturer's specifications. The sample input volume was 200 μl, and the elution volume was set at 100 μL for all extracted samples.

2.2. SARS-CoV-2 detection by RT–qPCR

Molecular diagnosis of SARS-CoV-2 was performed by amplification and detection of target sequences in the ORF1ab and N genes using the Detection Kit for 2019 Novel Coronavirus (2019-nCoV) RNA (PCR-Fluorescence Probing) (Da An Gene Co., Ltd. of Sun Yat-sen University. Cat. # D-930) according to the manufacturer's protocol. In brief, 5 μL of extracted RNA was added directly to 20 μL of the reaction mixture. Each reaction mixture contained 17 μL of PCR solution A (specific primers, probes, and TRIS-HCl buffer) and 3 μL of solution B (hot start Taq DNA polymerase, c-MMLV reverse transcriptase). RT-PCR assays were performed using a CFX96 Touch instrument (Bio-Rad, CA, UDA), with the following reaction conditions: a cDNA synthesis step at 50 °C for 15 min, initial denaturation at 95 °C for 15 min, followed by 45 amplification cycles of 94 °C for 15 s and 55 °C for 45 s. The cycle threshold (CT) was determined using the CFX Maestro Software (Bio-Rad). A CT value $\leq$ 40 was considered detected/positive, and a CT value > 40 was considered negative. For each assay, a negative and positive control was included.

2.3. Next-generation sequencing of SARS-CoV-2

To investigate whether the SARS-CoV-2-related mortality increase was due to the presence of (new) variants with higher rates of transmission, sequencing of the viral genome was carried out for 92 SARS-CoV-2 RNA samples using the AmpliSeq SARS-CoV-2 panel for Illumina in combination with the AmpliSeq Plus Library for Illumina (Cat. # 20019103). In brief, 60 ng of total RNA was converted into cDNA using an AmpliSeq cDNA Synthesis for Illumina Kit (Cat. # 20022654), according to the manufacturer's instructions, followed by reactions to amplify target regions of the cDNA sample using an AmpliSeq Custom RNA Panel for Illumina kit (Cat. # 20020496). The reaction mixture contained 5 μL of cDNA, 4.5 μL of 5X AmpliSeq HiFi Mix, and 3.5 μL of nuclease-free water. The amplification mix was distributed into two wells, and 2 μl each of 5X AmpliSeq Custom RNA Panel Pool 1 and Pool 2 were added. The amplification conditions were 99 °C for 2 min, followed by 18 cycles at 99 °C for 15 s and 60 °C for 4 min. Then, the amplicons were digested with 2 μl of FuPa Reagent. Dual indexes and adapters were attached to the digested amplicons by DNA ligase. The libraries were cleaned using Agencourt AMPure XP beads (Beckman Coulter) and then amplified by PCR. The amplification master mix contained 45 μL of 1X Lib Amp Mix and 5 μL of 10X Library Amp Primers. The PCR thermal cycling protocol used was 98 °C for 2 min, followed by seven cycles at 98 °C for 15 s and 64 °C

for 1 min. To ensure optimal cluster density in the flow cell, the libraries' quality and quantity were assessed using an Agilent High Sensitivity DNA Kit (Agilent cat. # 5067–4626) and KAPA SYBR FAST qPCR Master Mix (Kapa Biosystems cat. # KK4600). Based on the quality and quantity analyses, the library pool was normalized to a concentration of 30 pM. Sequencing was performed on a MiSeq System using the MiSeq Reagent Kit v3 (Illumina Cat. # MS-102-3001). The Illumina DRAGEN COVID Lineage App was used to assign lineages as well as PANGO.

3. Calculations

3.1. Data collection

The study data included the daily number of deaths caused by SARS-CoV-2 in the Guadalajara Metropolitan area reported by the Government of Mexico. The study period was from October 1, 2020, to May 17, 2021. (COVID-19 Tablero México).

The meteorological data included mean temperature data from the state of Jalisco and data obtained from the Comisión Nacional del Agua (CONAGUA). The average temperatures of each month were used (October 2020–May 2021). (Resúmenes Mensuales de Temperaturas y).

Daily average concentrations of air pollutants ($PM_{2.5}$, PM_{10}, O_3, CO, SO_2, and NO_2) were obtained from the following locations: Tlaquepaque, Zona Centro, Santa Fe, Las Águilas, Las Pintas, Atemajac, Loma Dorada, and Miravalle. These data were obtained from Sistema Nacional de Información de la Calidad del Aire (SINAICA). The study period was from October 1, 2020, to May 17, 2021 (Instituto nacional de ecología y, 2021).

3.2. Rates of positive tests

Results of SARS-CoV-2 PCR tests were collected from the ONKOGENETIK molecular biology laboratory.

Samples were collected from October 1, 2020, to May 17, 2021. A total of 2994 samples were collected, of which 1061 were positive for SARS-CoV-2. The analyzed RNA was obtained from people located in the MZG.

The daily test positivity rate was obtained as follows:

$$rate\ (+) = (\#Detected * 100)\ /\ (\#\ Not\ detected). \quad \text{Eq. 1}$$

3.3. Particle-temperature ratio

This ratio was calculated using pollutant particle concentrations ($PM_{2.5}$, PM_{10}, O_3, CO, SO_2, and NO_2) and temperature. No direct correlation between pollutant concentration and deaths was used for this study since the relation between pollutants concentration and temperature may contribute to substantial increase of air pollution, therefore a ratio that represents this phenomenon was used (Wine et al., 2022). During our analysis, we found that it is possible to calculate this ratio using either the monthly average temperature or the daily average temperature since there was no significant difference between them (daily $r = 0.7101$, $P < 0.0001$; monthly $r = 0.7386$, $P < 0.0001$). However, if the daily average temperature is used, it is necessary to pair the first day of I_p data with the death count on the fifth day (lag). This is due to two main reasons: first, the highest correlation coefficients were found by pairing the data in this way, and second, in a substantial number of patients, the environmental conditions on one day aggravate the symptoms on the following days, in other words, the effect of air pollution on the clinical horizon might be observed after five days. On the other hand, using the monthly average temperature simplifies the calculations. Herein, we calculated this ratio using the daily particle concentration divided by the monthly average temperature.

$$I_p = (Daily\ particle\ concentration)\ /\ (Average\ monthly\ temperature) \quad \text{Eq. 2}$$

3.4. UV index

The monthly UV radiation index (kWh/m^2) was obtained for the city of Guadalajara. The maximum and minimum values of UV radiation were averaged. (POWER data access viewer).

All twelve months were averaged to obtain an annual value I_{UV}.

For each month, a percentage was obtained with respect to the annual value, as follows:

$$\%Monthly_{UV} = ((UV\ radiation\ index\ for\ each\ month) \times (100))\ /\ I_{UV} \quad \text{Eq. 3}$$

Subsequently, the percentage difference between the $\%\ Monthly_{UV}$ and 100% for each month was obtained as follows:

$$Monthly\ Difference_{UV} = 100\% - \%\ Monthly_{UV} \quad \text{Eq. 4}$$

The correlation between the daily average UV radiation index in the MZG and the 5-day lag number of deaths was calculated as a positive control.

3.5. Linear correlation

Linear correlation analyses were carried out to analyze whether there was a relationship between I_p and the number of deaths caused by SARS-CoV-2. Correlation analysis was performed for each pollutant ($PM_{2.5}$, PM_{10}, O_3, CO, SO_2, and NO_2). The correlation analyses were performed in Origin software.

In addition, linear correlation analysis between *ICO* (air pollution CO/temperature index) and *rate (+)* and between absolute positive cases and deaths with a 5-day lag was calculated. Finally, correlations between each pollutant and absolute number of positive cases, temperature, and UV index were analyzed as well.

3.6. Multiple linear regression

Based on the results of the linear correlation analyses, we performed a multiple correlation analysis to construct a predictive model of expected deaths based on the *ICO* and *rate (+)*.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 \quad \text{Eq. 5}$$

Where:

y = *deaths*.
β_0 = *y intercept* (*constant term*).
β_1 = *slope coefficient for* I_{CO}.
β_2 = *slope coefficient for rate* (+).
β_3 = *Monthly Difference*$_{UV}$.
x_1 = I_{CO}.
x_2 = *rate* (+).

There are two ways for calculating the monthly expected deaths. One is using the *Daily Difference*$_{UV}$ and the other one is using the *Monthly Difference*$_{uv}$ (β_3) on Eq. (5).

4. Results

The SARS-CoV-2 positivity rate-based RT–PCR testing was remarkably stable during the first year of the pandemic in Guadalajara, Jalisco, Mexico.

4.1. Absolute count of cases

During the first year of the SARS-CoV-2 pandemic, we identified three peaks for absolute incident case numbers; the first was between July and August 2020, the second was in October 2020, and the third was from December 2020 to January 2021, with notable reductions in February and May 2021 (Fig. 1A).

4.2. Relative proportions of cases

In the first two and last two months of this study, we identified the largest changes in the relative frequencies of samples that were positive for SARS-CoV-2 genetic material. In May 2020, the proportion of positive SARS-CoV-2 test results was

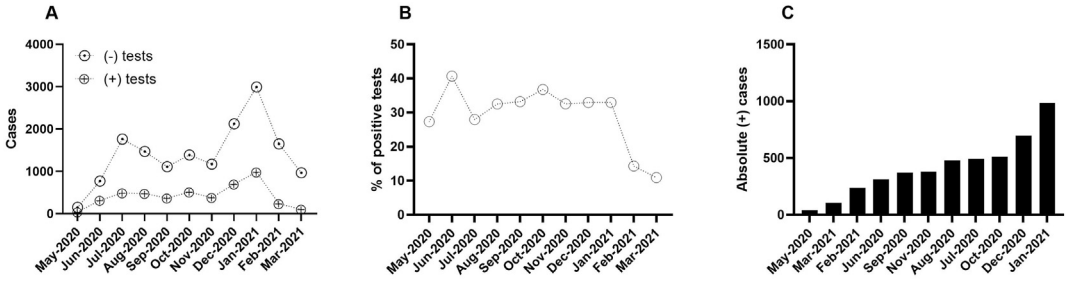


Fig. 1. SARS-CoV-2 infection spread during the first year of the pandemic in the MZG. A) Absolute count of SARS-CoV-2-positive and negative cases in the general population. B) Sustained incidence of SARS-CoV-2 positive tests during the whole study period. Hypothetical penetration of SARS-CoV-2 in the population. The entire year maintained a constant contagion pattern, even from the winter period 2020–2021 until February 2021 C) Order of months considering the absolute number of cases.

approximately 27%, followed by an increase to 40% in June 2020. for the next seven months (July 2020 to January 2021), the average proportion of SARS-CoV-2-positive tests was approximately 32.68%, with a standard deviation of 2.5. A notable decrease was observed in February (14%) and March (10.96%) 2021 (Fig. 1A–C). No increase in the incidence was observed during the whole period of SARS-CoV-2 virus transmission, except in June 2020 (Fig. 1A).

4.3. SARS-CoV-2 viral loads in samples

Monthly comparative viral loads were studied from October 2020 to February 2021. Viral loads of samples collected in December 2020 and January 2021 showed significantly higher relative viral counts than those collected in October 2020, November 2020, and February 2021. No significant differences in viral loads were found between December 2020 and January 2021 or October 2020, November 2020, and February 2021 (Fig. 2).

4.4. The air pollution/temperature index was highly correlated with mortality in Guadalajara, Jalisco, Mexico

The examined particles were $PM_{2.5}$ and PM_{10}, and the polluting gases were NO_2, O_3, SO_2, and CO. Air pollution and temperature were converted into an index (I_p) and showed a high correlation with mortality. In general, all air pollutant concentrations evaluated, except for O_3 and SO_2, followed the pattern of the daily mortality rates (Fig. 3). For O_3, there was a sudden increase in the index that might have been due to forest fires during March 2021 (Fig. 3E). This might also be the case for SO_2 in May and April 2021 (Fig. 3I). The most relevant correlation was that of the CO index (ICO), with an $r = 0.7386$, a standard deviation of 0.0284 and a P value of <0.0001. The second air pollutant with the highest correlation was NO_2, with an $r = 0.7060$, a standard deviation of 0.0006, and a P value of <0.0001. The graphs in Fig. 3 were separated into two periods for a deeper understanding of the influence of the seasons (temperature, pollution, UV index) on the SARS-CoV-2 mortality (Fig. 4). The first one is in autumn, from October 1st to December 21st, symbolizing the pre-high mortality period. The other one involves winter and spring, from December 22nd to May 17th, and it shows the mortality and post-mortality periods.

On the other hand, there was no correlation between *ICO* and *rate (+)* ($r = 0.1230$, $P = 0.0636$). but we found a high correlation between positive cases and absolute number of deaths ($r = 0.7985$, $P < 0.0001$). No relevant correlations were found between pollutant concentration and positive cases (Fig. 5). However, a negative correlation was found between the UV Index and absolute number of deaths (Fig. 6), which could mean that the more UV radiation, the less mortality ($r = -0.6607$, $P < 0.0001$). The correlation coefficients between each pollutant and temperature and UV Index are shown in Table 1.

4.5. The most prevalent SARS-CoV-2 lineages in Guadalajara, Jalisco, Mexico

Since the average SARS-CoV-2 incidence was approximately 32% during the last year, with increased mortality in December 2020 and February 2021, we divided the samples by collection date into two periods: from October–December 15, 2020 (when we considered that mortality started to increase) and from December 16, 2020, to February 2021. These periods were established to detect lineage changes or new prevalent lineages that might have increased the mortality rate during the second period. Nevertheless, sequencing results showed that there were no significant changes or new lineages between the periods (Fig. 7A). Analysis of the entire period from October 2020 to February 2021 showed that in approximately 80% of SARS-CoV-2-positive cases, the most prevalent SARS-CoV-2 lineages were B.1.243 (33.7%), B1.1.222 (11.2%), B.1.1 (9%), B.1 (7%), B.1.1.159 (7%), and B.1.2 (7%) (Fig. 7B and C for phylogeny and Table 2).

4.6. It is possible to predict mortality by considering air pollution concentrations, percentages of positive tests, temperatures, and UV radiation

We calculated the expected deaths using the *Daily Difference*$_{UV}$ and the *Monthly Difference*$_{UV}$. On Table 3, we show that there is no significant difference when using either one. Nevertheless, the values obtained when using *Monthly Difference*$_{UV}$ are closer to the real deaths than when using daily data.

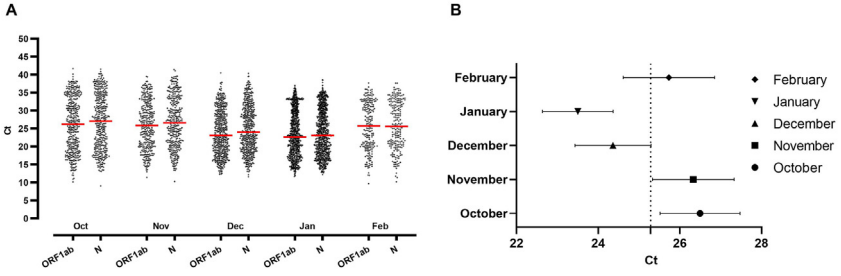


Fig. 2. A) Relative viral loads in samples positive for SARS-CoV-2 using RT–PCR. During the winter months (December 2020–January 2021), a reduction in CT was observed. ANOVA ($n = 4847$, $P = 0.0074$, SD 6.564).

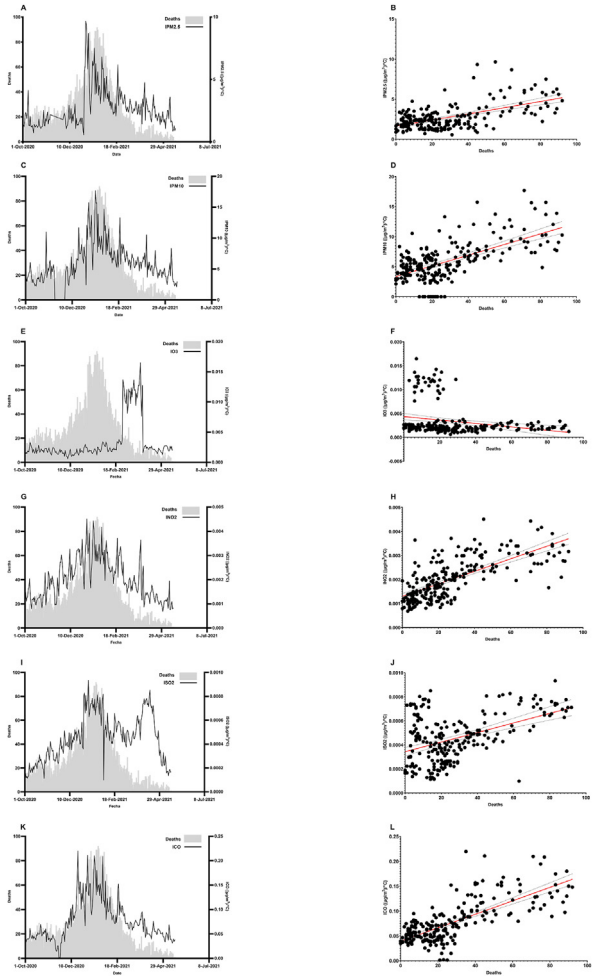


Fig. 3. Air pollution and mortality. All pollution particles or gases presented similar patterns to the mortality curve when adjusted for temperature (index). A) $IPM_{2.5}$ over the mortality curve. B) $IPM_{2.5}$ regression graph (n = 229, r = 0.5770, SD 1.299, P<0.0001). C) IPM_{10} over the mortality curve. D) IPM_{10} regression graph (n = 229, r = 0.6252, SD 2.548, P<0.0001). E) IO_3 over the mortality curve. F) IO_3 regression graph (n = 229, r = 0.2427, SD = 0.0033, P = 0.0002). G) INO_2 over the mortality curve. H) INO_2 regression graph (n = 229, r = 0.7060, SD = 0.0006, P<0.0001). I) ISO_2 over the mortality curve. J) ISO_2 regression graph (n = 229, r = 0.4687, SD = 0.0002, P<0.0001). K) ICO over the mortality curve. L) ICO regression graph (n = 229, r = 0.7386, SD = 0.0284, P<0.0001).

Using this simple model, we obtained very similar counts of daily deaths summarized per month. For example, in October 2020, an average of 18.58 deaths per day occurred during this period. Our predictive model predicted an average of 20.86 deaths per day when similar air pollution conditions, positive test percentages, temperatures, and UV radiation (average) are present. This prediction may vary from the calculated average by plus or minus five points (see Fig. 8: distance of units). We noticed that the largest differences between actual and predicted values occurred especially in the months characterized by climate transition between seasons (December).

5. Discussion

5.1. The apparent increase in the number of SARS-CoV-2-positive cases during the winter months in the MZG

According to our collected data (Fig. 1), considering only positive tests, there were two increases in the absolute count of incident cases: one from May to November 2020 and another from December 2020 to January 2021. Nevertheless, throughout 2020, the rate of SARS-CoV-2-positive cases was approximately 32% considering the total number of tests performed. Thus, during the apparent increase in the number of cases in December 2020 and January 2021, the rate of positive tests was also approximately 32%, indicating that there was an increase in the number of tests performed despite the same percentage of samples infected. To the best of our knowledge, this means that the SARS-CoV-2 infection rate in the MZG was very similar during the whole year of the pandemic and that there were no “infection waves”. Anti-infection precautionary measures seemed to be insufficient or ineffective in reducing the spread of the contagion and the probability of hospitalization or the

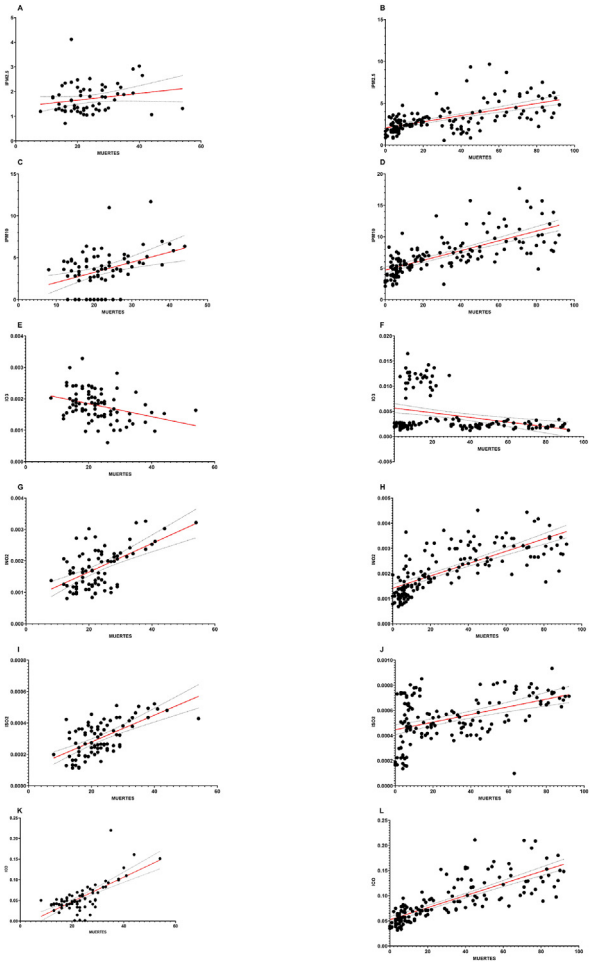


Fig. 4. Air pollution and mortality divided into two periods simulating seasons in MZG, period 1 from October 1 to December 21, and period 2 from December 22 to May 17. A) $IPM_{2.5}$ regression graph for period 1 (n = 82, r = 0.2006, SD = 0.5823, P = 0.1120). B) $IPM_{2.5}$ regression graph for period 2 (n = 147, r = 0.6004, SD = 1.347, P < 0.0001). C) IPM_{10} regression graph for period 1 (n = 82, r = 0.3764, SD = 2.183, P = 0.0005). D) IPM_{10} regression graph for period 2 (n = 147, r = 0.7068, SD = 2.150, P < 0.0001). E) IO_3 regression graph for period 1 (n = 82, r = 0.3410, SD = 0.0004, P = 0.0017). F) IO_3 regression graph for period 2 (n = 147, r = 0.3236, SD = 0.0037, P < 0.0001). G) INO_2 regression graph for period 1 (n = 82, r = 0.5797, SD = 0.0005, P < 0.0001). H) INO_2 regression graph for period 2 (n = 147, r = 0.7365, SD = 0.0006, P < 0.0001). I) ISO_2 regression graph for period 1 (n = 82, r = 0.6342, SD = 8.42×10^{-5}, P < 0.0001). J) ISO_2 regression graph for period 2 (n = 147, r = 0.4639, SD = 0.0002, P < 0.0001). K) ICO regression graph for period 1 (n = 82, r = 0.7075, SD = 0.0239, P < 0.0001). L) ICO regression graph for period 2 (n = 147, r = 0.7950, SD = 0.0253, P < 0.0001).

need for supplemental oxygen. However, supposed protection by such preventive measures might lead to an increase in mortality during the winter months because the preventive measures ignore other contributing factors, such as air pollution and low temperatures.

5.2. The correlation between air pollution/temperature and mortality

It is well known that during the winter season, due to low temperatures in the MZG, there is an increase in suspended pollution particles and gases in the air. The general population is exposed to these particles, especially during the morning and evening. Particle stagnation and lack of air movement coupled with higher concentrations of combustion pollution from vehicles and industries due to December population dynamics could have played a key role in the increase in pollution concentrations. Air pollution and low temperatures are known to increase the sensitivity of the respiratory tract to infective agents as well as damage *per se*. Exposure to air pollution and low temperature during COVID-19 infection could play a crucial role in its progression, causing fatal results, therefore reduced exposure to high concentration of pollutants and low temperature would help improve the clinical horizon (Tian et al., 1001).

Air pollution and low temperatures may also lead to higher viral loads in infected samples, as shown in Fig. 2, which demonstrates that during the winter period, there was an increase in the relative SARS-CoV-2 viral load due to an environmental change and not an intrinsic change in the characteristics of the viral particles. It is worth noting that there were no

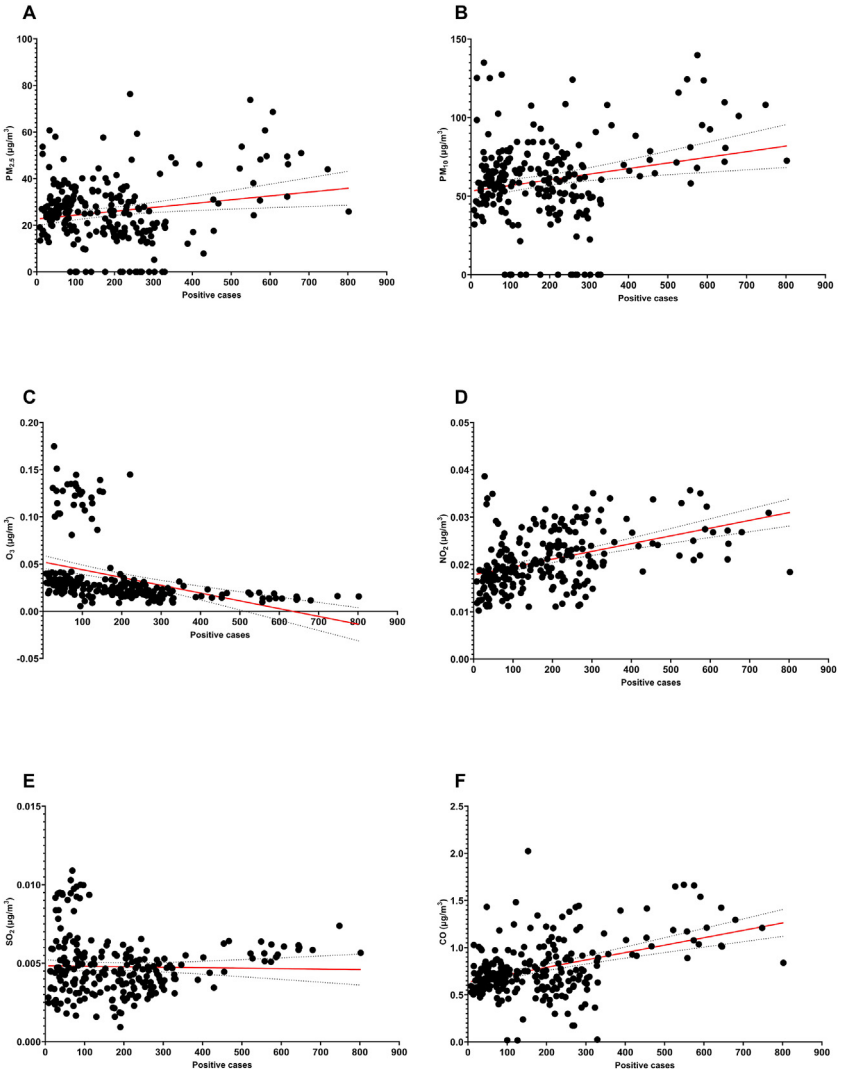


Fig. 5. Pollutant concentration and absolute positive cases. A) $PM_{2.5}$ regression graph (n = 229, r = 0.1834, SD = 13.79, P = 0.0054). B) PM_{10} regression graph (n = 229, r = 0.2112, SD = 25.91, P = 0.0013). C) O_3 regression graph (n = 229, r = −0.3609, SD = 0.0334, P < 0.0001). D) NO_2 regression graph (n = 229, r = 4288, SD = 0.0054, P < 0.0001). E) SO_2 regression graph (n = 229, r = −0.0249, SD = 0.0019, P = 0.7078). F) CO regression graph (n = 229, r = 0.4132, SD = 0.2710, P < 0.0001).

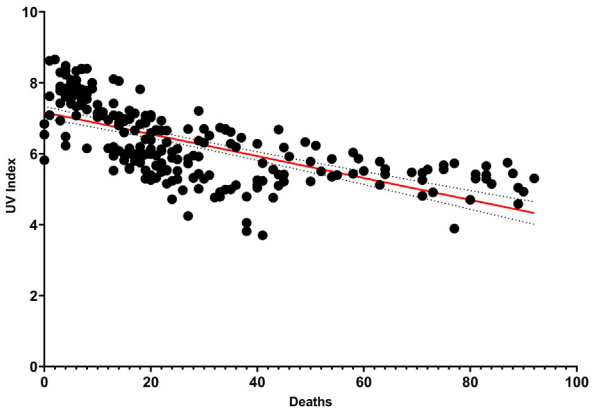


Fig. 6. Regression graph between UV Index and absolute number of deaths (n = 224, r = −0.6607, SD = 0.8118, P < 0.0001).

Table 1
Correlation coefficients between pollutant concentration and temperature and pollutant concentration and UV Index.

Environmental Factor	Particle	R^2	Pearson	P Value	Significant
Temperature	CO	0.1371	−0.3702	<0.0001	Yes
	$PM_{2.5}$	0.0039	−0.0622	0.3690	No
	PM_{10}	0.0371	−0.1926	0.0048	Yes
	O_3	0.1845	−0.4296	<0.0001	Yes
	NO_2	0.0146	0.1209	0.0679	No
	SO_2	0.0782	0.2796	<0.0001	Yes
UV Index	CO	0.0523	−0.2886	0.0005	Yes
	$PM_{2.5}$	0.0001	0.0107	0.8770	No
	PM_{10}	0.0054	−0.0734	0.2863	No
	O_3	0.0972	−0.3117	<0.0001	Yes
	NO_2	0.1239	0.3519	<0.0001	Yes
	SO_2	0.2100	0.4583	<0.0001	Yes

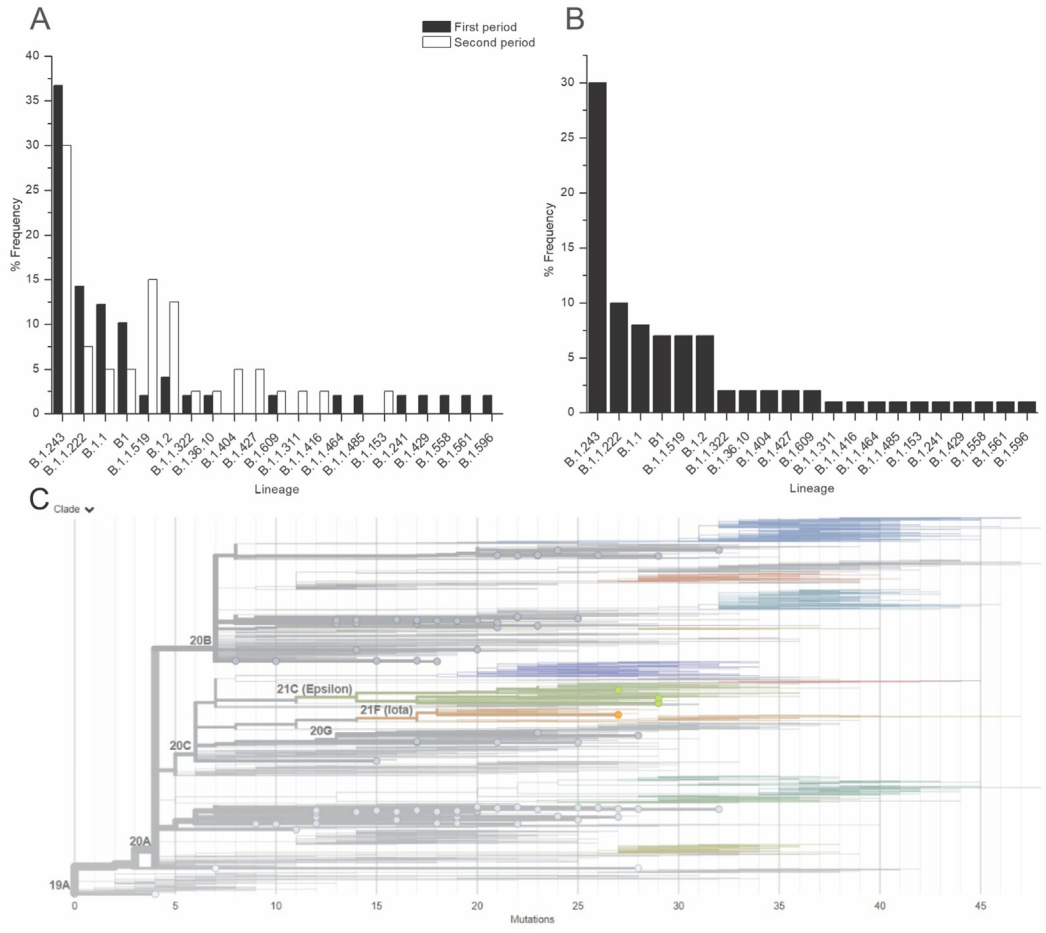


Fig. 7. SARS-CoV-2 lineages and their distribution frequencies from October 2020 to February 2021. A) To determine whether different lineages of SARS-CoV-2 were responsible for the mortality rate increase during December 2020 and January 2021, NGS was performed, and two periods were studied. The first period was from October 2020 to December 15, 2020. The second period (cut) was from December 16 to February 2021. Similar frequencies of dominant SARS-CoV-2 lineages were found in both periods. B) Dominant lineages of the whole period from October 2020 to February 2021. C) SARS-CoV-2 phylogeny of the analyzed isolated RNAs.

important changes in the rate of dominant lineages between the prehigh-mortality period and high-mortality period (see results Fig. 7).

The Environmental Protection Agency has identified 6 pollutants present in the air ($PM_{2.5}$, PM_{10}, NO_2, SO_2, CO, and Pb) (Grzywa-Celińska et al., 2020). In our investigation of how pollution influenced the increase in deaths from SARS-CoV-2

Table 2
Principal lineages identified during the period from October 2020 to February 2021. (POWER data access viewer) Most of the identified lineages that affected the MZG, were identified firstly in the United States of America. This suggests that the SARS-CoV-2 outbreak in this region first and constantly imported from the USA. The MZG is an important "hub" for Mexico-USA immigration.

Lineage	Most common countries	Earliest date	Description
B.1.243	United States of America 96.0%, Mexico 2.0%, Canada 1.0%, Northern Mariana Islands 0.0%, Sweden 0.0%	23/03/2020	USA lineage
B.1.1.222	United States of America 79.0%, Mexico 1.0%, Canada 0.0%, Spain 0.0%, Germany 0.0%	04/04/2020	USA/Mexico
B.1.1	United Kingdom 35.0%, United States of America 12.0%, Japan 9.0%, Germany 4.0%, Russia 4.0%	01/01/2020	European lineage with 3 clear SNPs '28881 GA', '288882 GA', '28883 GC'
B.1	United States of America 47.0%, United Kingdom 11.0%, Germany 4.0%, Canada 3.0%, Spain 3.0%	24/01/2020	A large European lineage the origin of which roughly corresponds to the Northern Italian outbreak early in 2020
B.1.1.519	United States of America 60.0%, Mexico 33.0%, Canada 2.0%, Germany 1.0%, Denmark 1.0%	01/11/2020	USA/Mexico lineage
B.1.2	United States of America 94.0%, Canada 4.0%, Mexico 0.0%, Germany 0.0%, Denmark 0.0%	01/01/2020	USA
B.1.1.322	United States of America 75.0%, Mexico 24.0%, Canada 1.0%	13/04/2020	Mexico/Southwest USA
B.1.36.10	United States of America 49.0%, Jordan 10.0%, India 10.0%, Turkey 10.0%, Saudi Arabia 5.0%	08/03/2020	Jordan/Saudi Arabia (single recent record from USA)
B.1.404	United States of America 96.0%, Mexico 2.0%, Canada 1.0%, Zambia 0.0%, Germany 0.0%	16/03/2020	USA lineage
B.1.427	United States of America 97.0%, Mexico 2.0%, Aruba 0.0%, Argentina 0.0%, Canada 0.0%	11/04/2020	USA lineage (CA)
B.1.609	United States of America 84.0%, Mexico 14.0%, Canada 0.0%, United Kingdom 0.0%, France 0.0%	10/03/2020	USA/Mexico lineage
B.1.1.311	United Kingdom 99.0%, Denmark 0.0%, Ireland 0.0%, United States of America 0.0%, Norway 0.0%	21/07/2020	Huge UK lineage
B.1.1.416	United States of America 93.0%, Australia 4.0%, Chile 1.0%, Canada 0.0%, Japan 0.0%	19/03/2020	USA lineage
B.1.1.464	United States of America 97.0%, Switzerland 1.0%, Israel 1.0%, Portugal 0.0%, United Kingdom 0.0%	05/04/2020	US lineage
B.1.1.485	Ghana 42.0%, United States of America 19.0%, Peru 19.0%, Russia 4.0%, Germany 2.0%	09/04/2020	Ghanan lineage
B.1.153	Turkey 48.0%, United Kingdom 19.0%, United States of America 6.0%, Germany 4.0%, India 4.0%	01/03/2020	English lineage with some Australian, New Zealand, and European sequences
B.1.241	United States of America 91.0%, Mexico 7.0%, Malaysia 0.0%, Australia 0.0%, Japan 0.0%	20/04/2020	USA lineage
B.1.429	United States of America 98.0%, Mexico 1.0%, Canada 1.0%, South Korea 0.0%, Denmark 0.0%	15/04/2020	A lineage predominantly circulating in California but with exports to other countries. Characterized by the spike L452R mutation
B.1.558	United States of America 73.0%, Mexico 24.0%, United Kingdom 2.0%, Kenya 1.0%, South Korea 1.0%	06/04/2020	USA/Mexico lineage
B.1.561	United States of America 89.0%, Mexico 6.0%, Canada 4.0%, Aruba 0.0%, Peru 0.0%	18/06/2020	USA – contains previous B.1.404
B.1.596	United States of America 99.0%, Mexico 0.0%, Canada 0.0%, Spain 0.0%, Australia 0.0%	31/03/2020	USA lineage

infection, we chose to include the following six environmental pollutants ($PM_{2.5}$, PM_{10}, O_3, NO_2, SO_2, and CO), and we also included two important environmental factors: temperature and UV radiation, which symbolize the seasons.

The studied contaminants have some common origins and some different origins. $PM_{2.5}$ particles are produced mainly by industrial processes, automobile pollution, and the burning of coal. It is composed of heavy metals, different ions, sulfate, ammonium, nitrate, and carbon sources. The physical characteristics of the particles depend on the season of the year, the region, and climate change (Cho et al., 2018). PM_{10} particles are composed of dust from construction sites, landfills, forest fires, burning waste, industrial sources, windblown dust from open land, pollen, and fragments of bacteria. (Inhalable Particulate Matter and Health) CO is produced by burning; for example, the burning of fuels by cars, trucks, and machinery. Additionally, some household appliances that use fossil fuels are a source of CO (US EPA O; Chen et al., 2007). NO_2 is produced by fuel burning and is naturally generated by organic decomposition, volcanoes, and lightning (Copat et al., 2020). SO_2 is generated mainly by the burning of fossil fuels in industrial processes (Chen & Yan, 2020).

Pollution exposure causes negative effects on health. It can affect different vital parts of the human body, such as the skin, the heart, the nervous system, the immune system, and the respiratory system (Hajirasouliha and Zabiegaj, 2552; Glencross et al., 2020). Some studies have shown that there is a relationship between air pollution exposure and increased mortality from respiratory diseases such as pneumonia, asthma, COPD, and pulmonary embolism (Ling & van Eeden, 2009). Airborne pollution particles that lodge in the respiratory system induce inflammation, as demonstrated in studies carried out on humans and animals. Those studies showed increases in lung inflammatory markers, affecting the immune system, and

Table 3
Expected deaths calculated using different methods, the daily method and the monthly method, compared to the average real deaths on each month. Averages for each month and method are shown.

Month	Deaths	Daily Method	Monthly Method
October	Real	18.58	
	Expected	17.38	20.86
November	Real	20.73	
	Expected	14.20	16.20
December	Real	35.10	
	Expected	44.32	51.36
January	Real	73.23	
	Expected	57.65	66.73
February	Real	39.27	
	Expected	25.23	30.16
March	Real	13.16	
	Expected	12.70	16.66
April	Real	6.93	
	Expected	7.50	10.26
May	Real	3.29	
	Expected	6.63	6.47
Daily method average			23.2
Monthly method average			27.34
Real average			26.29

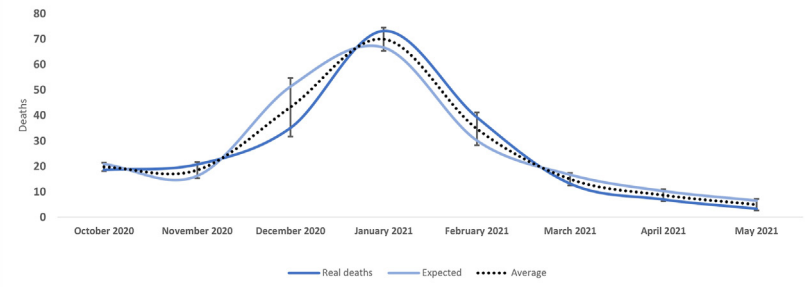
damaging cells such as macrophages, neutrophils, lymphocytes, and dendritic cells since the polluting particles stimulate a pro-inflammatory immune response (Glencross et al., 2020; Ling & van Eeden, 2009).

In 2020, the Center for Research on Energy and Clean Air (CREA) published an article stating that there is a relationship between air pollution exposure and deaths due to coronavirus disease 2019 (COVID-19). They mention that when exposed to pollutants, the body's natural defenses decrease, making the human body more vulnerable to invasion by viruses, such as SARS-CoV-2, and can cause COVID-19 patients to require hospitalization and artificial ventilation. They also mention that there is a lack of scientific studies to prove these theories (Centre for Research on Energy and Clean Air, 2020).

5.3. A predictive model for mortality in Guadalajara, Jalisco, Mexico

As shown in Fig. 8, we were able to predict mortality by considering air pollutant (CO) concentrations, the percentage of SARS-CoV-2 positive tests, the average temperature, and UV radiation. This predictive model was developed using data applicable to the MZG. Other factors not considered may affect the use of this simple method in other cities. This system does

A



B

Month	Real deaths	Expected	Stand. Dev.	Difference	Distance of units	ICO	% (+) tests
October	18.58	20.86	1.61	2.28	2.28	0.05	36.00
November	20.73	16.20	3.21	-4.53	3.82	0.03	34.09
December	35.10	51.36	11.50	16.26	16.26	0.11	33.06
January	73.23	66.73	4.59	-6.50	-6.50	0.15	32.94
February	39.27	30.16	6.44	-9.11	6.49	0.10	15.83
March	13.16	16.66	2.47	3.50	3.50	0.07	11.34
April	6.93	10.26	2.35	3.33	3.33	0.05	12.67
May	3.29	6.47	2.25	3.18	3.18	0.04	15.08
Total average	**26.29**	**27.34**	**3.30**	**1.05**	**4.04**	**0.07**	**23.88**

Fig. 8. Mortality predictive model.

not consider the altitude, vaccination status, or different weather characteristics. This model was developed to calculate daily maximum mortality and can be used to predict necessary hospital capacity via the analysis of historic data.

It is important to note that according to this model, reductions in industrial and vehicle pollution are the most effective ways to decrease mortality in the population who is sensitive to the exacerbation of respiratory symptoms, requires supplemental oxygen, or needs intensive care.

The stagnation of pollution particles and gases due to low temperatures is a well-described phenomenon during the winter season. In other words, if the temperature decreases, it is more likely that the general population will be exposed to air pollution particles and gases, potentially resulting in inflammatory effects and damage. When the temperature increases, the particles tend to move to higher altitudes. In addition, as shown in Table 1, when analyzed individually, the environmental factors show little correlation to pollutant concentration, which does not represent the real effect they have (Tian et al., 1001). To understand the effect of environmental factors on the pollutants' behavior, they must be analyzed together, for which we generated a more complex model. To generate this model, we used the daily CO index as a "template" since it had a higher r-value and greater statistical significance than the other pollutants, suggesting that the *ICO* is the most accurate indicator related to mortality. The monthly UV index for the MZG was used to represent adverse environmental conditions that may reduce the rate of infection since UV light is able to inactivate RNA viruses on surfaces and skin and in the air, as shown in Fig. 6. Using this model as well as the SARS-CoV-2 lineage information, we can infer that the substantial increase in mortality during December 2020 and January 2021 was not due to a change in or a new dominant SARS-CoV-2 lineage but to environmental conditions that play a significant role in increasing mortality.

Research groups from different regions of the world have studied the effect of environmental pollution on increases in the number of cases of SARS-CoV-2. They have studied different pollutants, mainly $PM_{2.5}$ and PM_{10}, in different periods. In our study, we included $PM_{2.5}$, PM_{10}, O_3, NO_2, SO_2, and CO, and we also included environmental factors such as temperature and the solar radiation index. Our study period covered October 2020 to May 2021, since December 2020 and January 2021 were when MZG had the highest number of deaths.

López-Feldman et al. studied the short- and long-term effects of $PM_{2.5}$ exposure on deaths caused by SARS-CoV-2 in Mexico City (Centre for Research on Energy and Clean Air, 2020). They constructed three models to observe the effects of pollution on the probability of death due to SARS-CoV-2 infection and found that there was a positive relationship between the factors. Moreover, they found that the effect was stronger for long-term exposures to $PM_{2.5}$ and that age was another important factor influencing mortality due to SARS-CoV-2 infection (López-Feldman et al., 2021). Another study reported positive relationships between the incidence of and mortality due to SARS-CoV-2 and $PM_{2.5}$ exposure in 24 districts of Lima, Peru, using linear correlation analysis (Vasquez-Apestegui et al., 2020).

PM_{10} particles have also been shown to have a negative effect on deaths caused by COVID-19, which was demonstrated by Isphording and Pestel, who studied the short-term effects of PM_{10} and O_3 exposure on deaths caused by SARS-CoV-2 in Germany through a mathematical model in which deaths due to SARS-CoV-2 infection were the dependent variable. Their results showed that increased particle exposure between four days before and ten days after the onset of symptoms increased the risk of death, as the number of confirmed SARS-CoV-2 cases per day increased (Isphording & Pestel, 2021).

Another research group studied the effects of six pollutants ($PM_{2.5}$, PM_{10}, SO_2, CO, NO_2, and O_3) in confirmed SARS-CoV-2 infection patients using a generalized additive model (GAM) (Zhu et al., 2020). They found that there were positive relationships between $PM_{2.5}$, PM_{10}, NO_2, and O_3 exposure and confirmed SARS-CoV-2 infection (Zhu et al., 2020). Jiang et al. performed a similar study and included some meteorological variants (Jiang et al., 2020). They studied the incidence of SARS-CoV-2 infection and its relationship with exposure to some pollutants ($PM_{2.5}$, PM_{10}, SO_2, NO_2, CO, and O_3) in China as well as some meteorological variants (wind, temperature, and humidity). They used multivariate Poisson regression and reported a positive relationship between SARS-CoV-2 infection development and $PM_{2.5}$ exposure and humidity (Jiang et al., 2020). Li et al. used simple linear regression analyses to investigate the effects of some pollutants ($PM_{2.5}$, PM_{10}, NO_2, and CO) on the incidence of SARS-CoV-2 infection (Li et al., 2020). They also analyzed some meteorological variants (daily temperature, lowest and highest temperatures, and the difference between the two and sunshine duration). They reported a positive relationship between exposure to $PM_{2.5}$ and NO_2 and SARS-CoV-2 infection development; they also found that temperature is an important factor in the incidence of SARS-CoV-2 (Li et al., 2020).

In our research work using linear regression, we found a positive relationship between deaths caused by SARS-CoV-2 and pollution caused by $PM_{2.5}$, PM_{10}, SO_2, NO_2, and CO. However, as seen in Fig. 3E, F and 5C, there is a negative relationship between deaths and ozone concentration. Although more studies are necessary, Ran et al. have shown interesting data where they found that a negative association between ozone and COVID-19 infectivity could be due to ozone virucidal activity, broad-spectrum disinfection properties, and sterilization properties against highly similar respiratory infections (Ran et al., 2020). Through multiple correlations, we developed a predictive model that allowed us to estimate the deaths caused by SARS-CoV-2 based on contamination by CO, temperature and the amount of UV radiation.

6. Conclusions

SARS-CoV-2 infection is still spreading among populations worldwide, as is the case in the MZG. New infections will continue to occur, even with the implementation of preventive actions; therefore, the control of other factors, such as air pollution, might be considered to reduce mortality. Herein, we demonstrated that air pollution played a key role in the increase in mortality in the MZG. Preventive actions should be taken to address this environmental problem. We did not find

any evidence that the increased rates of mortality were due to an intrinsic change in the SARS-CoV-2 variant. Using our predictive model, it is possible to forecast the patterns of mortality related to SARS-CoV-2 in the MZG.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank San Javier Hospital. This research was funded by Mexicana de Investigación y Biotectogía, S.A. de C.V.

References

Centre for Research on Energy and Clean Air. (2020). 11,000 air pollution-related deaths avoided in Europe as coal, oil consumption plummet. https://energyandcleanair.org/air-pollution-deaths-avoided-in-europe-as-coal-oil-plummet/.

Chen, T.-M., Kuschner, W. G., Gokhale, J., & Shofer, S. (2007). Outdoor air pollution: Nitrogen dioxide, sulfur dioxide, and carbon monoxide health effects. *The American Journal of the Medical Sciences, 333*, 249–256. https://doi.org/10.1097/maj.0b013e31803b900f

Chen, W., & Yan, W. (2020). Impact of internet electronic commerce on SO2 pollution: Evidence from China. *Environmental Science & Pollution Research, 27*, 25801–25812. https://doi.org/10.1007/s11356-020-09027-1

Cho, C.-C., Hsieh, W.-Y., Tsai, C.-H., Chen, C.-Y., Chang, H.-F., & Lin, C.-S. (2018). In vitro and in vivo experimental studies of PM(2.5) on disease progression. *International Journal of Environmental Research and Public Health, 15*, 1380. https://doi.org/10.3390/ijerph15071380

Copat, C., Cristaldi, A., Fiore, M., Grasso, A., Zuccarello, P., Signorelli, S. S., et al. (2020). The role of air pollution (PM and NO(2)) in COVID-19 spread and lethality: A systematic review. *Environmental Research, 191*, Article 110129. https://doi.org/10.1016/j.envres.2020.110129

COVID-19 Tablero México. https://datos.covid-19.conacyt.mx/index.php.

Glencross, D. A., Ho, T.-R., Camiña, N., Hawrylowicz, C. M., & Pfeffer, P. E. (2020). Air pollution and its effects on the immune system. *Free Radical Biology and Medicine, 151*, 56–68. https://doi.org/10.1016/j.freeradbiomed.2020.01.179

Gobierno del Estado de Jalisco. https://www.jalisco.gob.mx/es/jalisco/guadalajara. (Accessed 26 July 2021).

Grzywa-Celińska, A., Krusiński, A., & Milanowski, J. (2020). 'Smoging kills' – effects of air pollution on human respiratory system. *Annals of Agricultural and Environmental Medicine, 27*, 1–5. https://doi.org/10.26444/aaem/110477

Hajirasouliha F, Zabiegaj D. Effects of environmental emissions on the respiratory system: secrets and consequences. Environmental Emissions. https://www.intechopen.com/chapters/72552.

Inhalable particulate matter and health (PM2.5 and PM10). California air resources board. https://ww2.arb.ca.gov/es/resources/inhalable-particulate-matter-and-health.

Instituto nacional de ecología y cambio climático. https://sinaica.inecc.gob.mx/, (2021).

Isphording, I. E., & Pestel, N. (2021). Pandemic meets pollution: Poor air quality increases deaths by COVID-19. *Journal of Environmental Economics and Management, 108*, Article 102448. https://doi.org/10.1016/j.jeem.2021.102448

Jiang, Y., Wu, X.-J., & Guan, Y.-J. (2020). Effect of ambient air pollutants and meteorological variables on COVID-19 incidence. *Infection Control and Hospital Epidemiology*, 1–5. https://doi.org/10.1017/ice.2020.222

Ling, S. H., & van Eeden, S. F. (2009). Particulate matter air pollution exposure: Role in the development and exacerbation of chronic obstructive pulmonary disease. *International Journal of Chronic Obstructive Pulmonary Disease, 4*, 233–243. https://doi.org/10.2147/copd.s5098

Li, H., Xu, X.-L., Dai, D.-W., Huang, Z.-Y., Ma, Z., & Guan, Y.-J. (2020). Air pollution and temperature are associated with increased COVID-19 incidence: A time series study. *International Journal of Infectious Diseases, 97*, 278–282. https://doi.org/10.1016/j.ijid.2020.05.076

López-Feldman, A., Heres, D., & Marquez-Padilla, F. (2021). Air pollution exposure and COVID-19: A look at mortality in Mexico city using individual-level data. *Science of the Total Environment, 756*, Article 143929. https://doi.org/10.1016/j.scitotenv.2020.143929

O'Toole, Á., Scher, E., Underwood, A., Jackson, B., Hill, V., McCrone, J. T., et al. (2021). Assignment of epidemiological lineages in an emerging pandemic using the pangolin tool. *Virus Evol, 7*, Article veab064. https://doi.org/10.1093/ve/veab064

POWER data access viewer. https://power.larc.nasa.gov/data-access-viewer/.

Ran, J., Zhao, S., Han, L., Chen, D., Yang, Z., Yang, L., et al. (2020). The ambient ozone and COVID-19 transmissibility in China: A data-driven ecological study of 154 cities. *Journal of Infection, 81*(3), e9–e11. https://doi.org/10.1016/j.jinf.2020.07.011

Resúmenes Mensuales de Temperaturas y Lluvia. https://smn.conagua.gob.mx/es/climatologia/temperaturas-y-lluvias/resumenes-mensuales-de-temperaturas-y-lluvias.

Tian, F., Liu, X., Chao, Q., Qian, Z-, Zhang, S., Qi, L., et al. Ambient air pollution and low temperature associated with case fatality of COVID-19: A nationwide retrospective cohort study in China. Innovation, 2(3), 100139. https://doi.org/10.1016/j.xinn.2021.100139.

US EPA O. Basic information about carbon monoxide (CO) outdoor air pollution. https://www.epa.gov/co-pollution/basic-information-about-carbon-monoxide-co-outdoor-air-pollution.

US EPA O. (2014). Particle pollution and respiratory effects. https://www.epa.gov/particle-pollution-and-your-patients-health/health-effects-pm-patients-lung-disease.

Vasquez-Apestegui, V., Parras-Garrido, E., Tapia, V., Paz-Aparicio, V. M., Rojas, J. P., Sánchez-Ccoyllo, O. R., et al. (2020). Association between air pollution in Lima and the high incidence of COVID-19: Findings from a post hoc analysis. *Res Sq, 3*. https://doi.org/10.21203/rs.3.rs-39404/v1. rs-39404.

Wine, O., Osornio Vargas, A., Campbell, S. M., Hosseini, V., Koch, C. R., & Shahbakhti, M. (2022). Cold climate impact on air-pollution-related health outcomes: A scoping review. *International Journal of Environmental Research and Public Health, 19*(3), 1473. https://doi.org/10.3390/ijerph19031473

Zhu, Y., Xie, J., Huang, F., & Cao, L. (2020). Association between short-term exposure to air pollution and COVID-19 infection: Evidence from China. *Science of the Total Environment, 727*, Article 138704. https://doi.org/10.1016/j.scitotenv.2020.138704