



## OPEN ACCESS

## EDITED BY

Keith Elder,  
Mississippi College, United States

## REVIEWED BY

Maria Eugenia Jimenez-Corona,  
National Institute of Cardiology Ignacio  
Chavez, Mexico  
Liane S. Canas,  
King's College London, United Kingdom

## \*CORRESPONDENCE

Gustavo Mora-Aguilera  
✉ morag@colpos.mx

†These authors have contributed equally to this work and share first authorship

RECEIVED 16 July 2022

ACCEPTED 14 July 2023

PUBLISHED 07 August 2023

## CITATION

Acevedo-Sánchez G, Mora-Aguilera G,  
Coria-Contreras JJ and Álvarez-Maya I (2023)  
Were metabolic and other chronic diseases the  
driven onset epidemic forces of COVID-19 in  
Mexico? *Front. Public Health* 11:995602.  
doi: 10.3389/fpubh.2023.995602

## COPYRIGHT

© 2023 Acevedo-Sánchez, Mora-Aguilera,  
Coria-Contreras and Álvarez-Maya. This is an  
open-access article distributed under the terms  
of the [Creative Commons Attribution License  
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction  
in other forums is permitted, provided the  
original author(s) and the copyright owner(s)  
are credited and that the original publication in  
this journal is cited, in accordance with  
accepted academic practice. No use,  
distribution or reproduction is permitted which  
does not comply with these terms.

# Were metabolic and other chronic diseases the driven onset epidemic forces of COVID-19 in Mexico?

Gerardo Acevedo-Sánchez<sup>1†</sup>, Gustavo Mora-Aguilera<sup>1\*†</sup>,  
Juan J. Coria-Contreras<sup>1</sup> and Ikuri Álvarez-Maya<sup>2</sup>

<sup>1</sup>Laboratory of Epidemiological Risk Analysis (LANREF), Postgraduate College, Montecillo Campus, Texcoco, State of Mexico, Mexico, <sup>2</sup>Center for Research and Applied Technology in Jalisco (CIATEJ), Guadalajara, Jalisco, Mexico

The underline hypothesis of this study was that SARS-CoV-2 can infect individuals regardless of health condition, sex, and age in opposition to the classical epidemiological assumption of an identifiable susceptible subpopulation for epidemic development. To address this issue, a population cohort with 24.4 million metadata associated with 226,089 official RT-qPCR positive and 283,450 negative cases, including 27,769 deceased, linked putatively to *B.1.* and *B.1.1.* SARS-CoV-2 lineages were analyzed. The analysis baseline was to determine the infection and mortality structure of the diseased cohort at the onset-exponential phase of the first epidemic wave in Mexico under the assumption of limited herd immunity. Individuals with nonchronic diseases (NOCDs) were compared with those exhibiting at least one of 10 chronic diseases (CDs) adjusted by age and sex. Risk factors for infection and mortality were estimated with classification and regression tree (CART) and cluster analysis based on Spearman's matrix of *rho*-values in RStudio<sup>®</sup>, complemented with two proposed *mortality indices*. SARS-CoV-2 infection was independent of health condition (52.8% NOCD vs. 47.2% CDs;  $p = 0.001-0.009$ ) but influenced by age  $>46$  in one risk analysis scenario ( $p < 0.001$ ). Sex contributed 9.7% to the overall risk. The independent effect was supported by the health structure of negative cases with a similar tendency but a higher proportion of NOCDs (61.4%,  $p = 0.007$ ). The infection probability in individuals with one CD was determined by the disease type and age, which was higher in those older individuals ( $\geq 56$  years) exhibiting diabetes (12.3%,  $cp = 0.0006$ ), hypertension (10.1%,  $cp < 0.0001$ ), and obesity (7.8%,  $cp = 0.001$ ). In contrast, the mortality risk was heavily influenced by CD conditioned by sex and age, accounting for 72.3% of total deaths ( $p = 0.001-0.008$ ). Significant mortality risk (48%) was comprised of women and men (w, m) aged  $\geq 56$  years with diabetes (19% w and 27.9% m,  $cp < 0.0004$ ), hypertension (11.5% w,  $cp = 0.0001$ ), and CKD (3.5% w and 5.3% m,  $cp = 0.0009$ ). Older people with diabetes and hypertension comorbidity increased the risk to 60.5% ( $p = 0.001$ ). Based on a *mortality-weighted index*, women were more vulnerable to preexisting metabolic or cardiovascular diseases. These findings support our hypothesis and justify the need for surveillance systems at a communitarian level. This is the first study addressing this fundamental epidemiological question.

## KEYWORDS

SARS-CoV-2, comorbidities, mortality risk, infection risk, first wave

## Introduction

SARS-CoV-2, the most successful zoonotic coronavirus in human history, has caused over 668 million infection cases and more than 6.8 million deaths worldwide through several epidemic waves (1, 2). Since the Wuhan outbreak in China (3), at least 19 variants of the epidemic have emerged and spread rapidly before an effective natural immunological response (4). In infectious epidemic diseases, the classical paradigm behind the susceptible, infected, and recovered (SIR) individuals and any descriptive or predictive epidemiological model imply the preexistence of a susceptible subpopulation due to genetic, epigenetic, clinical, and environmental determinants as the driving forces for contagion (5–10). With COVID-19 epidemics, early findings supported that chronic diseases (CDs), age, and, to a less extent, sex were associated with the success and clinical outcomes of SARS-CoV-2 infection. However, most results were derived at the hospital level, from a small diagnostic dataset, or framed for descriptive epidemiological studies (3, 11, 12). More vital efforts should be addressed from the perspective of mechanistic epidemiology to enhance comprehensive prevention health systems to cope with the increasing risk of emerging and reemerging new human diseases. This study hypothesized that SARS-CoV-2 can infect individuals regardless of their health condition in opposition to the classical epidemiological assumption of an identifiable susceptible subpopulation for epidemic development. It was assumed that fast spreading, limited and unsteady immunological response toward a newly encountered pathogen, constrained clinical knowledge for treatment, and unprepared public health systems were fully expressed during the first wave of the COVID-19 outbreak, thus allowing unrestricted infection scenarios. The first epidemic wave also involved a higher global fatality rate reaching 15.2% (13). The Mexican population, with a high SARS-CoV-2 infection risk due to populated territorial clusters and high incidence of metabolic and cardiovascular chronic diseases in the world, was suitable to address this research (14–16). Previous efforts in Mexico mainly focused on demonstrating the CDs association with COVID-19 clinic course and mortality, thereby lacking a mechanistic epidemiological framework (17–20). This comprehensive study contributes to understanding the epidemiological behavior of new diseases in human populations and provides insights for surveillance and prevention of potential zoonotic outbreaks (21). Moreover, this study was based on big data associated with 509,539 official RT-qPCR test results, comprising 24.4 million metadata (22), which were putatively related to *B.1.* and *B.1.1.* SARS-CoV-2 lineages (23, 24), representing the onset-exponential phase of the first epidemic wave in Mexico (28 February to 30 June 2020). Our approach was to determine the subpopulation structure of infection in ambulatory and hospitalized cases, associated with 10 CDs and nonchronic diseases (NOCDs), considering age and sex as demographic factors in a cohort of 226,089 accumulated positive and 283,450 negative individuals, including 27,769 deaths. Therefore, the objective of this study was to establish the subpopulation attributes toward SARS-CoV-2 infection and the contribution of CDs and baseline demographic factors in shaping population vulnerability under the assumption of unrestricted immunological responses, treatments

availability, and preventive constraints for contagion during the onset of the first epidemic wave.

## Materials and methods

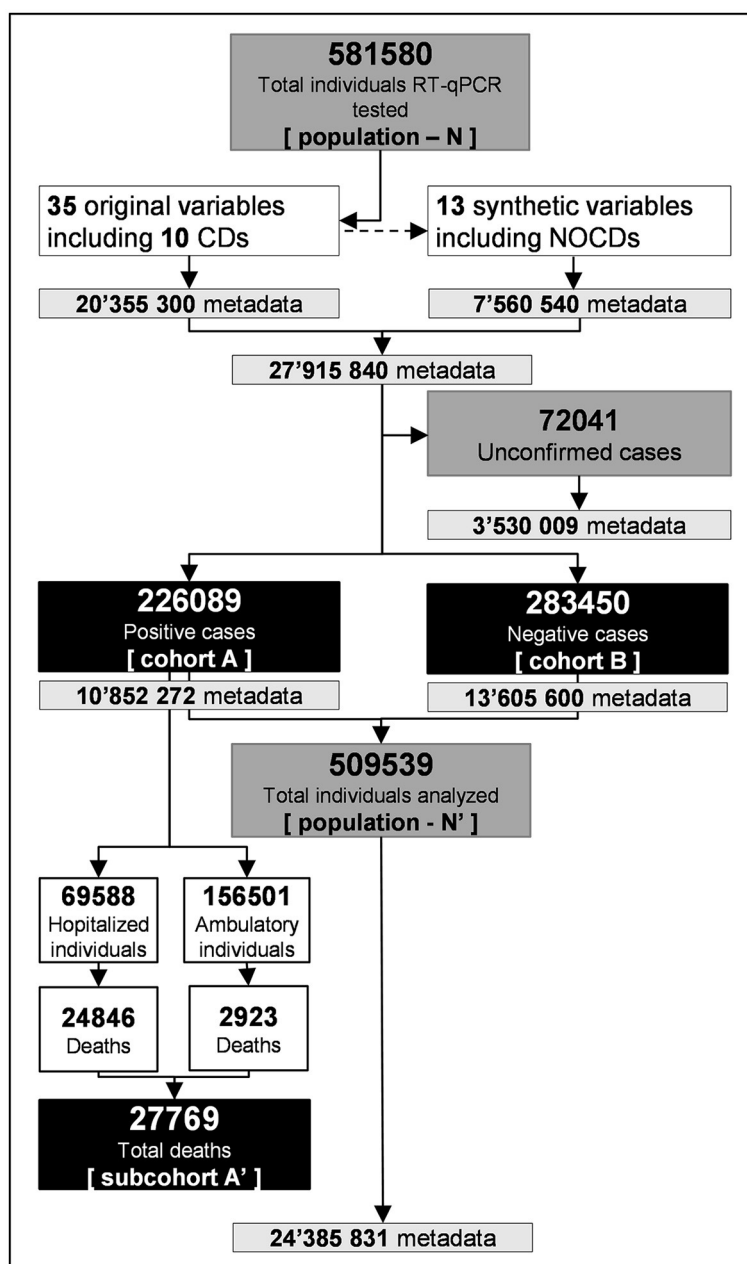
### COVID-19 data source

The first step was to collect the official COVID-19 public databases (MS Excel<sup>®</sup>, dBase-COVID) of the Mexican Ministry of Health (25), from the first positive SARS-CoV-2 reported on 28 February to 30 June 2020, selected for comprising the onset-exponential phase of the first epidemic wave in Mexico. The dBase-COVID, updated daily, had 581,580 individual records (population-N) and 35 variables (20.4 million metadata), including state and municipal locations, diagnosis results, symptoms expression date, death date, sex, age, and 10 CDs, among others (Figure 1 and Supplementary Table 1). All diagnostic tests were officially regulated and conducted with certified protocols based on real-time reverse transcription-polymerase chain reaction (RT-qPCR).

### Metadata structure

The second step was to set up the database structure to conform the research objective. The dBase-COVID data were imported into RStudio<sup>®</sup> v1.4.1106 – R Project<sup>®</sup> v4.1.1 and performed in a workstation (HP Z1-G6. IntelCore i7 of 10<sup>th</sup> generation). Data extraction was performed with *readxl*, *base*, *rattle*, and *dplyr* functions of Rstudio<sup>®</sup>. Sixteen numerical variables were transformed into categories, e.g., sex 1 = “female”, sex 2 = “male”, or CD 1 (presence of any chronic disease) = “yes”, CD 2 = “no”. The geo-location variables were transformed using the official nomenclature of the National Institute of Statistics and Geography (26). Additional 13 synthetic variables were created to potentially enhance the analyses, e.g., days with symptoms at testing or days from detection to death in the hospital settings. The final structured and conform database contained a cohort A of 226,089 positive individuals including 27,769 deaths, linked to 48 variables totaling 10'852272 metadata (Figure 1). A total of 72,041 unconfirmed RT-qPCR tests were excluded from the analysis (Figure 1). In this study, the infected cohort A was conform for all positive cases, symptomatic or asymptomatic at testing, including those individuals who eventually died. Death cases were considered subcohort A' of the infected cases (Figure 1).

To properly assess the age effect on infection, this variable was grouped into five categories (age<sub>c</sub>): <29, 30–37, 38–46, 47–56, and >56 years. Similarly, 10 CDs were independently analyzed, as well as by categories (CD<sub>c</sub>) according to clinical typology: metabolic (diabetes, obesity, immunosuppressants, and chronic kidney disease); cardiovascular (hypertension and cardiovascular disease); respiratory (asthma, COPD, and smoking), even though smoking is not a CD, it was considered due to implications on pulmonary diseases; “other-CDs” (this general category was specified as such in original data matrix); and a nonchronic



**FIGURE 1**  
Data extraction flow (black boxes) from 581,580 official database entries accumulated during the onset-exponential phase of the first COVID-19 epidemic wave in México (population-N), from 28 February to 30 June 2020. The final big data matrix was associated with 509,539 total individuals analyzed comprising 24.4 million metadata conform in cohort A with 226,089 RT-qPCR-positive cases including subcohort A' with 27,769 deaths and cohort B with 283,450 negative cases.

disease (NOCD) category for the absence of any reported CD on the dataset.

### Onset-exponential phase modeling

The third step was to confirm and characterize the onset-exponential epidemic phase intensity by fitting it to the exponential model and comparing 10 COVID-19 epidemics selected from

an equal number of countries with the highest reported positive cases at the first wave onset (1). The significant epidemic rate- $r_e$  estimation was fundamental to validate the fastest contagion assumption required to prove the working hypothesis. The comparison among epidemics to depict Mexico's scenario framed the study assumptions' validity. The plotting of all curves characterization was performed with *ggplot* function of RStudio® using cumulative daily ( $x$ ) positive cases from onset ( $y_0$ ) to the inflection curve point. The positive and death data ( $y$ ) were independently fitted in SAS® v9.4 using the nonlinear model:

$\hat{y} = y_0 r_e^{(x)}$ . The  $r_e$ -parameter and  $y_0$  estimated the exponential epidemic rate and positive cases of primary infection, respectively. The goodness-of-fit ( $R^2$ ) and significance level ( $p < 0.0001$ ) were obtained for comparison purposes.

### Probabilistic risk categorization for infection and mortality

The fourth step was to conduct an independent risk categorization analysis for the infection cohort A and mortality subcohort A' by using two approaches: the classification and regression tree (CART) and Spearman's  $\rho$  correlation linked to a clustering analysis. CART allows for identifying and weighting tree-decision rules to generate splitting stratified groups of similar risk toward SARS-CoV-2. These rules were fitted using *rpart*, *rpart.plot*, and *prp* functions and the analysis of variance (ANOVA) among groups in RStudio®. The *rpart* and *rpart.plot* best-fitting function for major splitting generated an overall *complexity parameter* (*cp*) value,  $cp = 0.000003$  ( $p = 0.001$ ) and  $cp = 0.000024$  ( $p = 0.001$ ), for infection and mortality CART, respectively. The splitting stratification process runs  $n$ -iterations for each encountered group until a homogeneity value lower than the *complexity parameter* (*cp*) is reached, thus providing the optimal solution. This parameter estimated and compared the variance homogeneity within groups for the final decision. Each CART was fitted as multiple regression model:  $y_i = x_1 + x_2 + x_3 \dots x_n$ , where  $y_i$  was the infected or death cases as dependent variables, and  $x_1 \dots x_n$  were 10 CDs, NOCD, age, and sex as variable predictors. Finally, with *prp*, a risk tree was built via cross-validation, thus creating stratified groups at the lowest error (27). Only nodes with statistically significant  $p$ -values ( $p \leq 0.05$ ) were plotted. Nodes per quartile of cases number were colored using a bar-scale. The CART procedure was selected because (1) it establishes rules based on multivariate criteria to explain overall variance (28); (2) it does not make any statistical distribution assumptions associated with dependent or independent variables (29, 30); and (3) it stratifies and classifies data based on weighted variables to create high- or low-risk homologous groups (30).

The second approach used was Spearman's correlation matrix based on 10 CDs, NOCD, age, and sex variables for pairings  $\rho$  estimations. Furthermore, a hierarchical cluster analysis was performed using the Euclidian distance of  $\rho$ -values as a dissimilarity measure among clusters and Ward's minimum variance to minimize the within-cluster variance. Independent dendrograms for the infection cohort A and mortality subcohort A' were plotted with the *tanglegram* function of RStudio® for comparison purposes. In addition, per dendrogram, the infection and mortality relative risk ( $r$ ) for tree clusters were estimated with  $r = [y_j / \sum y] / 100$ , where  $y$  is the total infected or death cases and  $\sum y$  is the total infected cohort A or mortality subcohort A' (Figure 1). Spearman's correlation matrix and clustering were selected because 1) it standardize data based on the variables' association level, reducing the effect of sample size and 2) it allow estimating a statistical significance ( $p \leq 0.05$ ).

The fifth step was to perform analogous analyses with cohort B comprising 283,450 negative cases, assuming individual exposure to the SARS-CoV-2 virus by social contact with positive cases.

The purpose was to analyze the whole population-N' structure toward SARS-CoV-2 infection risk. The overall analyses included 509,539 individuals and 24.4 million metadata (population-N', Figure 1).

### A deterministic risk categorization for mortality

To further explain the implication of CD categories on COVID-19 mortality subcohort A' (Figure 1), two relative epidemiological indices were developed to estimate the mortality stratified by age<sub>c</sub> and sex. A *mortality index* (MoI) was calculated with the following equation:

$$MoI = \frac{\sum_{ij}^n Deaths_{ij}}{\sum_j^n Cases_j}$$

where  $Cases_j$  is the number of positive individuals in  $j$ ;  $i$  represents the age<sub>c</sub> category from  $n = 1$  to 5;  $j$  is the CD<sub>c</sub> category from  $n = 1$  to 4; and NOCD.

A *mortality-weighted index* (MWI), weighted by the average ( $\bar{x}$ ) of age in each category, was calculated with the next equation:

$$MWI = \frac{\sum_{ij}^n Deaths_{ij} * \bar{x}_{ij}}{\sum Deaths}$$

where  $i$  and  $j$  are described as beforehand.

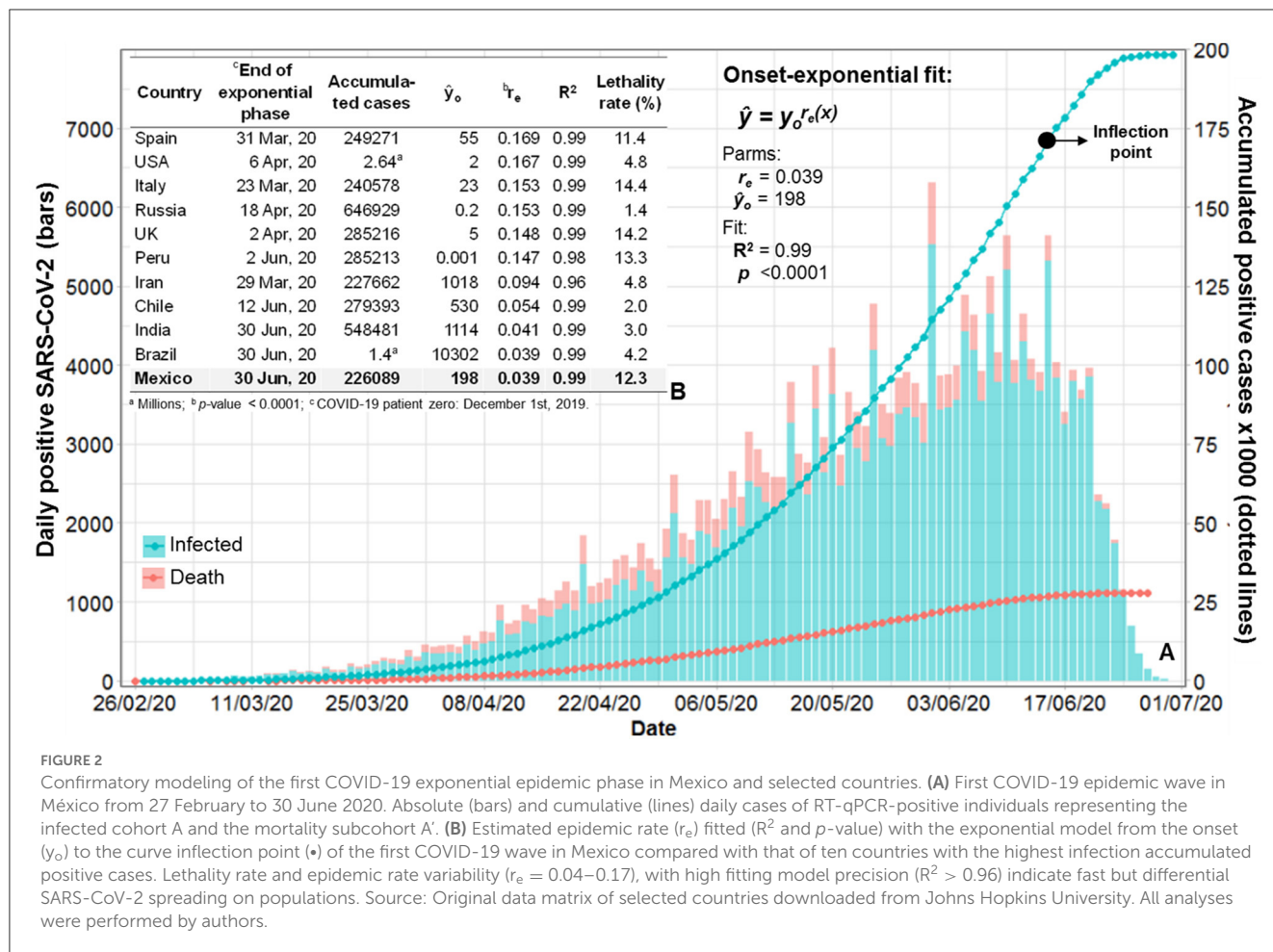
## Results

### SARS-CoV-2 metadata structure

A total of 226,089 individuals infected with SARS-CoV-2 during the onset-exponential phase of the first COVID-19 epidemic wave in Mexico were included in the study (cohort A; Figure 1 and Supplementary Table 1). The mean age was 45.7 years (range: 10–98 years), with 54.7% male cases. The overall positivity was 44.4% [(infected individuals/total individuals tested) x 100], with an official lethality rate of 12.3%. Hospitalized individuals accounted for 30.8% (62.2% of whom were men), with fatalities reaching 35.5% (65.1% of whom were men). The outpatient mortality was 1.8%. The data represented 32 Mexican states, of which 36% was associated with the metropolitan area of Mexico City and Mexico State with a combined 23.1 million habitants and a density of 6,163.3 and 760.2 residents by square kilometer, respectively. The data comprised a well-conform exponential phase as in selected comparative epidemics but with a relatively lower epidemic rate ( $r_e = 0.040$  units day<sup>-1</sup>), in contrast to Spain, USA, Italy, Russia, the UK, and Peru, which ranged from 0.15 to 0.17 (Figure 2A). In all cases, the exponential model fitted with  $R^2 > 0.96$  (Figure 2B). The lethality rate of Mexico was among the highest, in conjunction with Italy, Spain, UK, and Peru.

The infected cohort A, i.e., the total number of positive individuals regardless of the COVID-19 outcome, included 52.8% with NOCDs (53.9% of whom were men). The remaining 47.2%





**FIGURE 2** Confirmatory modeling of the first COVID-19 exponential epidemic phase in Mexico and selected countries. (A) First COVID-19 epidemic wave in México from 27 February to 30 June 2020. Absolute (bars) and cumulative (lines) daily cases of RT-qPCR-positive individuals representing the infected cohort A and the mortality subcohort A'. (B) Estimated epidemic rate ( $r_e$ ) fitted ( $R^2$  and  $p$ -value) with the exponential model from the onset ( $y_0$ ) to the curve inflection point (\*) of the first COVID-19 wave in Mexico compared with that of ten countries with the highest infection accumulated positive cases. Lethality rate and epidemic rate variability ( $r_e = 0.04$ – $0.17$ ), with high fitting model precision ( $R^2 > 0.96$ ) indicate fast but differential SARS-CoV-2 spreading on populations. Source: Original data matrix of selected countries downloaded from Johns Hopkins University. All analyses were performed by authors.

exhibited at least one CD (55.5% of whom were men), representing 27.1, 13.1, and 7% of single CD, comorbidity, and multimorbidity, respectively. The most reported chronic diseases were obesity (20,539 cases, 52.3% men) and hypertension (14,048 cases, 54.1% men). Grouped into categories, metabolic diseases (i.e., diabetes, obesity, immunosuppressants, and CKD) represented 39.7% of CD cases (Table 1). Diabetes-hypertension (4.5%) and diabetes-hypertension-obesity (2%) were the most prevalent comorbidity and multimorbidity, respectively. The control dataset (cohort B), with 283,450 RT-qPCR negative cases (Figure 1), had similar age, gender, and CD structure to cohort A.

### Probabilistic risk categorization for SARS-CoV-2 infection and mortality

As the first probabilistic classificatory approach applied to the infected cohort A (226,089 individuals), tree risk categorization significantly selected age as the primary factor of infection risk, with 46.9% of the explained variance ( $cp$ -value = 0.000003) (Figure 3A). The age cutoff onto two main probabilistic branches, from which a significant classificatory risk node was derived, was 46 years, which

represented 123,047 ( $p = 0.001$ – $0.009$ ) and 103,042 ( $p = 0.001$ – $0.002$ ) for younger and older than the significant age cutoff, respectively. Furthermore, age (29 years) and sex were the second most significant subordinated factors toward infection ( $p = 0.001$ ). Notably, NOCD represented only 6.4% of the explained variance due to restricted probabilistic combinations only within age and sex toward infection. This restricted determination resulted in 58,679 and 28,204 infection cases associated with NOCD in the root branches determined by sex and age (29–46 years), respectively.

A similar low variance contribution was found on infection associated with CDs. The type of disease determined the probability of infection in individuals with one CD, conditioned by age, being higher in those individuals exhibiting diabetes (12.3%,  $cp = 0.0006$ ), hypertension (10.1%,  $cp = 0.0001$ ), and obesity (7.8%,  $cp = 0.001$ ) accounting for a total of 30.2%. Sex contributed 9.7% of the overall infection risk, mainly associated with ages older than 46 years. For instance, for women younger and older than 46 years threshold, 12,829 and 28,702 had SARS-CoV-2 infection, respectively, and exhibited at least one chronic disease (Figure 3A). For the same contrasting risk scenario considering only diabetic women, there were 2,918 and 13,293 positive cases for younger and older than the 46-year cutoff, respectively. Notably, this combinatory effect was even higher in men, with 2,481 and 42,117 cases, indicating a higher infection probability in diabetic older men

**TABLE 1** Structure of 226,089 SARS-CoV-2 infected cohort A, including the subcohort A' with 27,769 mortality cases, and Mol and MWI epidemiological relative indices adjusted by age and sex.

Category	Age	$\bar{x}$ age	Men			Women			Total	
			Death	Cases	Mol <sup>x</sup>	Death	Cases	Mol	Deaths/cases	Mol
Metabolic (diabetes, obesity, Imm, and CKD)	<29	24.6	124	3,077	0.040	126	3,077	0.041	250/6,154	0.041
	29–37	33.7	438	5,381	0.081	195	4,746	0.041	633/10,127	0.063
	37–46	42.3	1,243	9,001	0.138	601	7,534	0.080	1,844/16,535	0.112
	46–56	51.4	2,730	12,267	0.223	1,549	10,589	0.146	4,279/22,856	0.187
	> 56	67	7,004	18,038	0.388	5,137	16,089	0.319	12,141/34,127	0.356
<b>Subtotal</b>			<b>11,539</b>	<b>47,764</b>	<b>0.242</b>	<b>7,608</b>	<b>42,035</b>	<b>0.181</b>	<b>19,147/89,799</b>	<b>0.213</b>
<b>MWI<sup>y</sup></b>			<b>37.22</b>			<b>48.47</b>			<b>25.36</b>	
Cardiovascular (hypertension and CVD)	<29	23.9	44	689	0.064	46	504	0.091	90/1,193	0.075
	29–37	33.9	118	1,506	0.078	62	1,004	0.062	180/2,510	0.072
	37–46	42.6	504	3,569	0.141	260	2,812	0.092	764/6,381	0.120
	46–56	51.6	1,421	6,675	0.213	843	5,760	0.146	2,264/12,435	0.182
	>56	70	5,771	15,045	0.384	4,114	12,991	0.317	9,885/28,036	0.353
<b>Subtotal</b>			<b>7,858</b>	<b>27,484</b>	<b>0.286</b>	<b>5,325</b>	<b>23,071</b>	<b>0.231</b>	<b>13,183/50,555</b>	<b>0.261</b>
<b>MWI</b>			<b>27.49</b>			<b>36.39</b>			<b>18.85</b>	
Respiratory (COPD, asthma, and smoking)	<29	24.4	44	2,681	0.016	16	1,663	0.010	60/4,344	0.014
	29–37	33.5	129	3,102	0.042	35	1,994	0.018	164/5,096	0.032
	37–46	41.6	312	3,292	0.095	95	2,204	0.043	407/5,496	0.074
	46–56	51	569	3,194	0.178	223	2,071	0.108	792/5,265	0.150
	>56	68	2,224	5,705	0.390	976	3,117	0.313	3,200/8,822	0.363
<b>Subtotal</b>			<b>3,278</b>	<b>17,974</b>	<b>0.182</b>	<b>1,345</b>	<b>11,049</b>	<b>0.122</b>	<b>4,623/29,023</b>	<b>0.159</b>
<b>MWI</b>			<b>10.88</b>			<b>8.80</b>			<b>6.28</b>	
Other CD <sup>s</sup>	<29	20.7	16	207	0.077	10	260	0.038	26/467	0.056
	29–37	33.7	7	176	0.040	2	308	0.006	9/484	0.019
	37–46	42	19	224	0.085	19	328	0.058	38/552	0.069
	46–56	51.1	36	189	0.190	23	276	0.083	59/465	0.127
	>56	68	128	276	0.464	60	220	0.273	188/496	0.379
<b>Subtotal</b>			<b>206</b>	<b>1,072</b>	<b>0.192</b>	<b>114</b>	<b>1,392</b>	<b>0.082</b>	<b>320/2,464</b>	<b>0.130</b>
<b>MWI</b>			<b>0.65</b>			<b>0.67</b>			<b>0.41</b>	
Nonchronic disease	<29	22.9	117	13,554	0.009	58	14,070	0.004	175/27,624	0.006
	29–37	33.5	275	13,678	0.020	96	12,917	0.007	371/26,595	0.014
	37–46	41.9	721	14,262	0.051	231	12,136	0.019	952/26,398	0.036
	46–56	51.1	1,468	11,960	0.123	426	9,262	0.046	1,894/21,222	0.089
	>56	66	3,058	10,948	0.279	1,242	6,649	0.187	4,300/17,597	0.244
<b>Subtotal</b>			<b>5,639</b>	<b>64,402</b>	<b>0.088</b>	<b>2,053</b>	<b>55,034</b>	<b>0.037</b>	<b>7,692/119,436</b>	<b>0.064</b>
<b>MWI</b>			<b>17.50</b>			<b>12.50</b>			<b>9.75</b>	
<b>Total</b>			<b>18,289</b>	<b>123,616</b>		<b>9,480</b>	<b>102,473</b>		<b>27,769/226,089</b>	

COVID-19 data of the first onset-exponential epidemic phase in Mexico.

<sup>x</sup>Relative mortality index (Mol) associated to CDc and age. <sup>y</sup>Relative mortality-weighted index associated to CDc (MWI).

than in diabetic women. In individuals younger than 29 years, the infection risk associated with those exhibiting at least one CD was 25.6%. The remaining CD<sup>s</sup> cases, independent of sex, were associated with obesity (4,393 cases), smoking (596 cases),

and immunosuppressants (314 cases) with a risk of 14.5%. Other CD<sup>s</sup>, such as kidney (CKD), cardiovascular (CVD), smoking, and immunosuppressants, accounted for 6.9% of the infection risk variance. Lower risk of infection, but significant ( $p = 0.009$ ),

probably due to underrepresentation in cohort A, was found to be associated with individuals with comorbidities, i.e., more than one CD, such as obesity and smoking (1,556 men); diabetes–obesity–smoking (1,009 men); diabetes–CVD (953 women); and obesity–immunosuppressant (60 women) (Figure 3A).

In the restricted analyses of the mortality subcohort A' (27,769 cases), age was again the primary significant risk factor with 72.3% of the explained variance conditioned by the type of CD ( $cp$ -value < 0.000024), from which a significant tree classification risk was derived upon a 56-year cutoff ( $p = 0.001$ – $0.008$ ) (Figure 3B). NOCD accounted only for 3.7% of the variance, representing 26.9% of all death cases ( $cp = 0.0006$ ,  $p = 0.001$ ). A robust significant risk, representing 48% of the cases, was composed of women and men (w, m) older than the 56-year cutoff who mainly exhibited diabetes (19% women, 27.9% men;  $cp = 0.0004$ ), hypertension (11.5% women, nonsignificant in men,  $cp = 0.0001$ ), and CKD (3.5% women, 5.3% men;  $cp = 0.0009$ ). Deaths with comorbidity combinations involving diabetes were significantly associated with CKD among patients aged 56 years, regardless of sex (401 cases). Conversely, in those individuals older than 56 years, diabetes was significantly combined with hypertension (1,748 women and 2,261 men) (Figure 3B). Multimorbidity disease significant combinations, regardless of the age category, included diabetes–CKD–CVD (10 cases), diabetes–hypertension–immunosuppressants (14 women), and diabetes–hypertension–CVD (8 men) (Figure 3B). For individuals younger than 29 years, mortality was independent of sex and mainly associated with obesity and hypertension (60 and 32 deaths, respectively). However, this node was not significant.

The second probabilistic associative approach applied to the infected cohort A and subcohort A', based on the matrix of the Spearman's  $\rho$ -values (Figures 4A1, B1), confirmed that CD, age, and sex did not fully explain infection risk toward SARS-CoV-2. NOCD, with 52.8% of infection probability and conform for 119,436 positive cases, represented a well-separated independent cluster ( $p = 0.05$ ) at a Euclidean distance of 1.4 cutoff. The CDs and demographic factors formed four risk clusters with 47.2% infection probability (Figure 4A2) ( $p = 0.04$ – $0.06$ ). Age and sex conform a cluster with diabetes and hypertension, and obesity and smoking, respectively ( $p = 0.05$ ). The infection risk increased to 25.4% and 15.3%, respectively, for individuals who presented the two diseases (i.e., comorbidity). Age influenced the vulnerability of older people with diabetes or hypertension to developing SARS-CoV-2 infection ( $p < 0.00001$ ). Age-related associations with diabetes and hypertension had the highest positive  $\rho$ -values of 0.39 and 0.33, respectively (Figure 4A1).

Contrary to the infection scenario and targeting only the subcohort, the higher probability for mortality was associated with CDs and age totaling 72.2% (Figure 4B2) ( $p = 0.05$ ;  $\rho = -0.51$ – $0.31$ ). NOCD and sex defined a well-distant risk cluster of 27.8% ( $p = 0.05$ ). Older people with comorbidity of diabetes–hypertension ( $p = 0.04$ ) had an increased risk of death at 60.5%, whereas those with a single CD accounted for only 18.8%. These conditions were more determinant over the threshold of 56 years (Figure 4B2). The comparison between infection and mortality dendrograms showed a slight displacement of risk-cluster location with an estimated 66% similarity, thus indicating

differences in influencing health factors toward SARS-CoV-2 outcome (Figures 4A2, B2).

In the negative cases of cohort B, the variance structure was similar to cohort A (Supplementary Figures 1A, 2 and Supplementary Table 2). The primary statistically significant age cutoff was also 46 years ( $p = 0.001$ – $0.007$ ) (Supplementary Figure 1A). For those older than 46 years (51,206), smokers were the first cutoff linked to obesity. Meanwhile, nonsmokers were associated with diabetes, obesity, and hypertension. For those under 46 years (58,337), obesity was the leading cutoff, but linked to diabetes and smokers. In NOCD-negative individuals (173,907), the population structure variance was determined only by sex and age as expected (Supplementary Figure 1 and Supplementary Table 2). The cluster structure was also similar to positive SARS-CoV-2 in cases of cohort A. The cross-dendrogram correlation revealed associativity of  $r^2 = 0.93$  among cohorts. Notably, asthma was included in the sex–obesity–smoking cluster (Supplementary Figure 2).

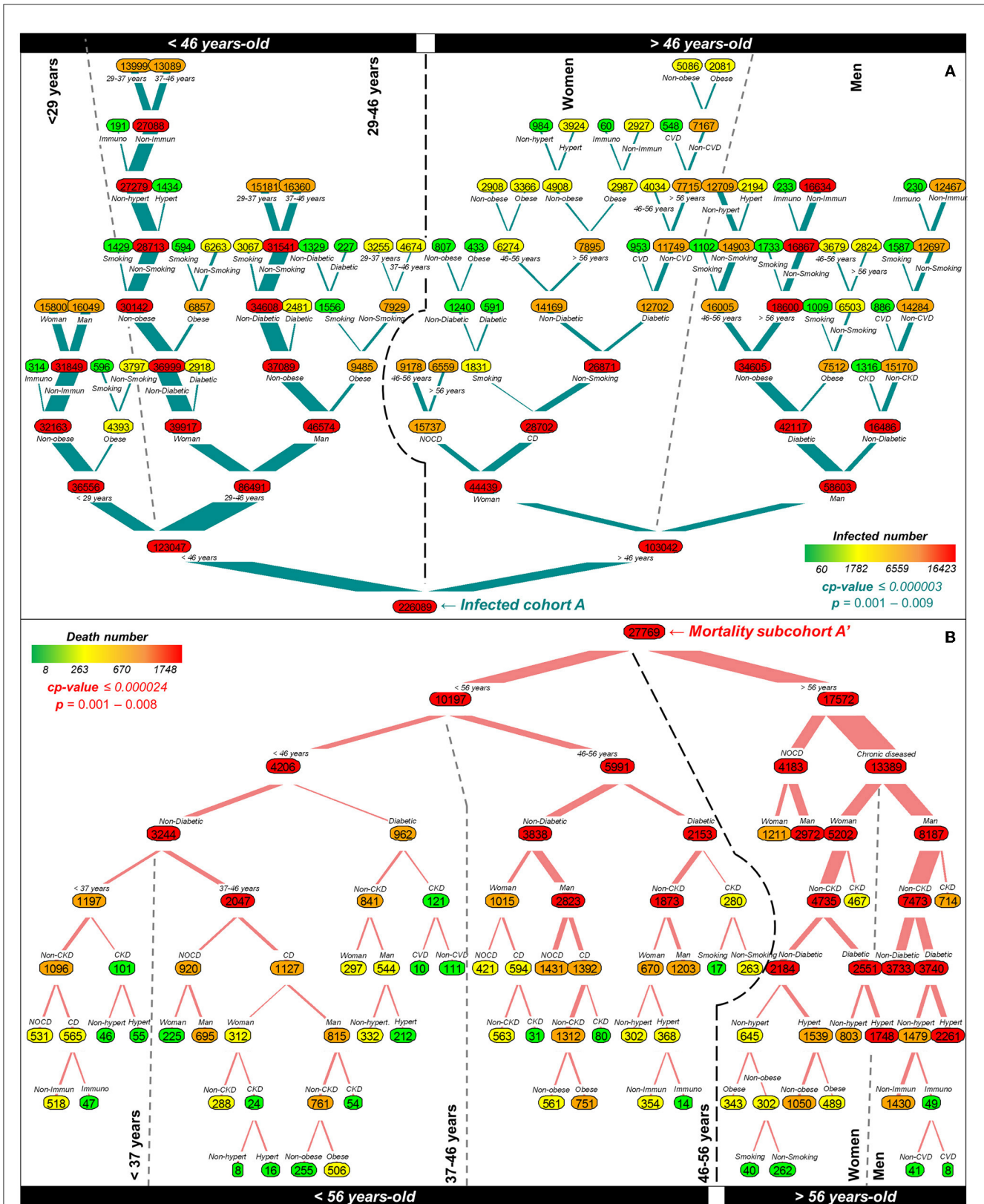
## SARS-CoV-2 relative mortality indices

The relative *mortality index* (MoI) stratified by age confirmed the differential effect of CD category ( $CD_c$ ) and NOCD on mortality (Figure 5A). Cardiovascular and metabolic diseases represented the higher index with 0.26 and 0.21, respectively, whereas NOCD was the lowest with 0.06 (Table 1). MoI values increased by age category ( $age_c$ ) and were higher, but similar, for patients older than 56 years among  $CD_c$  (0.35–0.37) compared to NOCD (0.24), thus indicating a significant conditional age effect on mortality (Figure 5A). Conversely, for ages less than 56 years, the MoI did not exhibit clear differences between  $CD_c$  and NOCD. As for sex, the MoI was consistently higher among men than women, independent of age,  $CD_c$ , or NOCD (Table 1).

The relative *mortality-weighted index* (MWI) showed that mortality was also influenced by  $CD_c$  and sex (Figure 5B). Again, the individuals with metabolic or cardiovascular diseases were associated with a higher mortality risk index of 25.4 and 18.9, respectively, compared to NOCD (9.8). However, contrary to MoI, women were notably the most vulnerable in metabolic and cardiovascular categories with 48.5 and 37.3, respectively. Furthermore, men had a higher risk associated with respiratory diseases and NOCD (Table 1; Figure 5B). These indices were calculated relative to each  $CD_c$  to avoid biases due to sample underrepresentation of specific chronic disease category in cohort A.

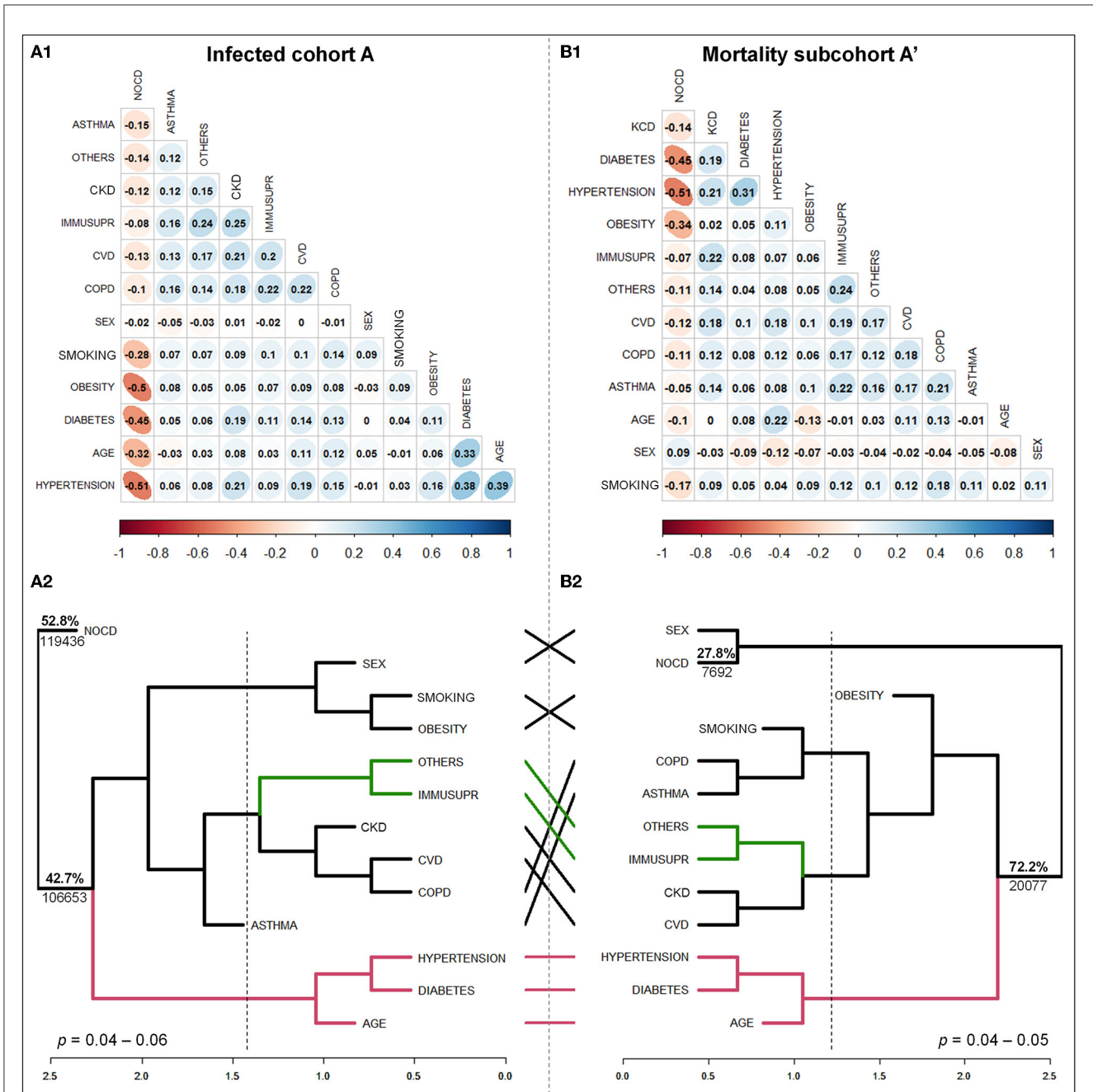
## Discussion

Despite massive vaccination and lethality reduction, the recent COVID-19 pandemic, which was characterized by fast virus contagion, a dynamic prevalence of variants, and a reduction of the age threshold for infection, raises questions about our mechanistic comprehension of SARS-CoV-2 epidemiology at the communitarian level (6). Most studies continue to focus on an understanding of the infection clinical outcome,



**FIGURE 3**  
 Tree risk categorization of infection and mortality due to SARS-CoV-2 during the onset-exponential phase of the first COVID-19 epidemic wave in Mexico based on 226,089 positive cases and 10'857,272 metadata records comprising 13 variables including NOCD and CDs. Branch thickness represents the main root of significant risk. The colored bar scale represents the number range of positive cases applied to nodes. (A) The major virus infection risk implicated four main branches, highlighted by upper black boxes and vertical dotted lines, determined primarily by age, followed by sex, with a  $cp\text{-value} \leq 0.000003$  ( $p < 0.009$ ). The infection risk for individuals with NOCD represented 52.5%. (B) The mortality risk was also influenced by age ( $cp\text{-value} \leq 0.000024$ ;  $p < 0.008$ ) but conditioned by chronic diseases with a higher association in older to 56 years (48%) and splitting the risk by sex (women = 38.9%; men = 61.1%). NOCD accounted for 26.9% of the mortality risk. Main and secondary tree branches are highlighted in bottom black boxes and vertical dotted lines, respectively.

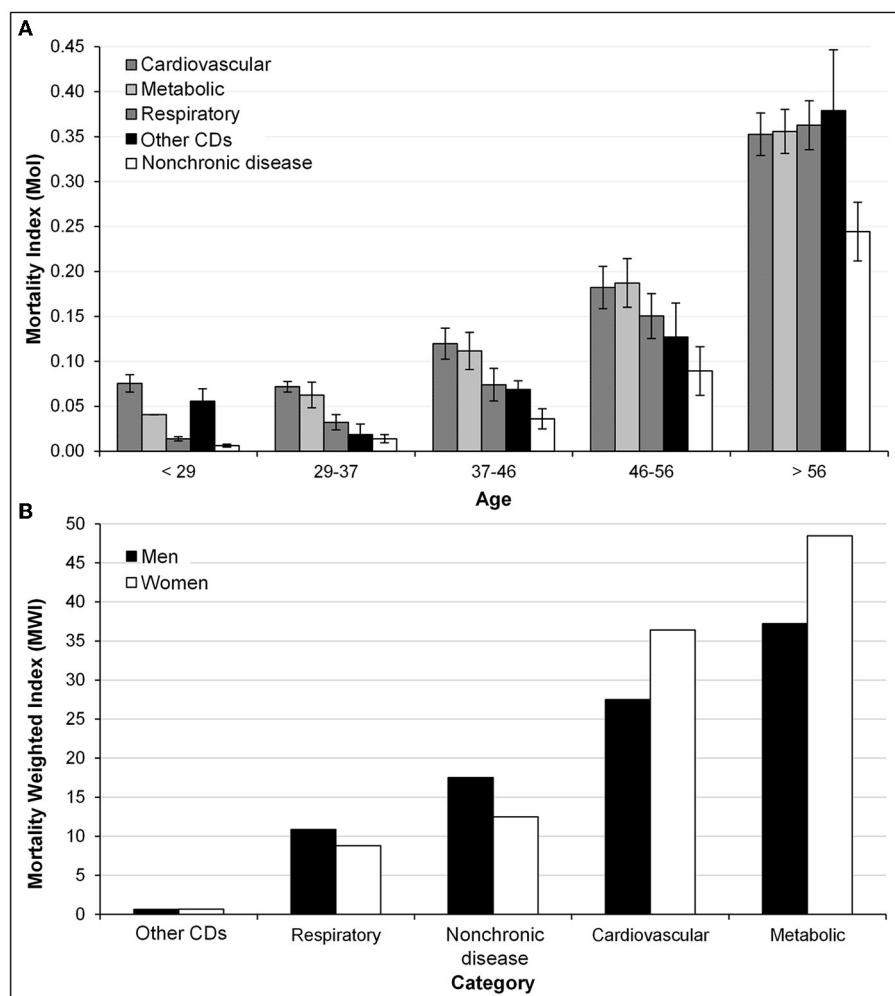




**FIGURE 4** Differential risk structure toward infection (A1, A2) and mortality (B1, B2) due to SARS-CoV-2 based on Spearman's rho-values estimated with 13 variables, comprising sex, age, nonchronic (NOCD), and 10 chronic (CDs) non-infectious diseases associated with 226,089 infected individuals during the onset-exponential phase of the first COVID-19 epidemic wave in Mexico. (A1, B1): correlation matrix for the infected cohort A and mortality subcohort A', respectively. The colored bar-scale represents the rho-value. If closer to ± 1 indicates a higher correlation between variables. (A2, B2): Dendrogram of rho, linked to cluster analyses for the infected cohort A and mortality subcohort A', shows respectively, a clear independent and dependent risk effect on CD, age, and sex, respectively. The scale at the bottom represents the dissimilarity of Euclidean distance. The dotted line represents the cutoff for risk-cluster conformation, and the percentage is the estimated risk based on positive cases associated with a specific branch ( $p = 0.04-0.06$ ). Lines connecting dendrograms identify the clustering variables. Others. Other CDs.

particularly the post-COVID condition, the development of cure treatments, and the enhancement of vaccines to include children (7, 8, 11, 31-33). However, there is still a strong need for comprehensive studies associated with virus behavior at the ambulatory population level for surveillance

and prevention purposes (21). Current forecasting relies on limited clinical and hospital settings data (34-38). Moreover, current data availability and quality of detection and monitoring have been strongly compromised based on the worldwide expectation of immunization coverage to cope with the disease.



**FIGURE 5** Association of SARS-CoV-2 positive individuals exhibiting nonchronic diseases (NOCD) or any CD within five categorized chronic diseases (CD<sub>c</sub>) with COVID-19 mortality at the first exponential epidemic phase in Mexico. **(A)** Differential increase of relative mortality index (Mol) values on individuals with NOCD and CD<sub>c</sub> upon age category increase (age<sub>c</sub>). Bars represent the standard deviation. **(B)** Differential effect on relative mortality-weighted index (MWI) values of men versus women on respiratory and NOCD (higher) and on cardiovascular and metabolic category (lower).

The recent endemic and seasonal statement may even more discourage keeping epidemiological studies at the communitarian level (39).

This study deals with a fundamental epidemiological assumption of the preexistence of a susceptible population as a driving force for SARS-CoV-2 epidemics. Our findings challenge the presence of such a subpopulation. The analyses of 226,089 positive individuals and 10'852272 metadata records representing the specific onset-exponential first wave in Mexico (Figure 1) suggest that infection at the communitarian level relies more on infectious sources in the proximity of individuals independently of their health conditions, sex, or age as has been commonly implied (40–42). Rather than ‘choosing’ vulnerable subpopulation(s), this random infection was supported by the fact that baseline chronic diseases, extensively associated with COVID-19, did not condition infection. In one probability scenario, our structural risk analyses showed that individuals with NOCD have a slightly higher infection probability (52.8%) than those exhibiting any

CD, including comorbidities, without age and sex influence ( $p = 0.05$ ). In a second scenario, a cutoff of 46-year individuals was conditioned to diverse risk categories of virus infection ( $p = 0.001–0.009$ ). However, although age and sex have been extensively associated with COVID-19 severity and always associated with CDs under our analytical scenarios, age standalone was a significant factor in shaping the infection risk structure in the population but decreased the age threshold with respect to most reports, wherein older people appear to be more vulnerable. In such reports, the focus on the clinical evolution of inpatients may explain this discrepancy (11, 33, 38, 40, 41, 43). The independent effect of infection regarding CDs toward COVID-19 was supported by the health structure of negative cases with a similar tendency but a higher proportion of NOCDs (61.4%,  $p < 0.007$ ) (Supplementary Figures 1, 2 and Supplementary Table 2).

After restricting the analysis to the mortality subcohort, the results are in agreement with extensive studies suggesting that CD, age, and sex are implicated in COVID-19 severity (33, 40, 41,

44–47). Nonetheless, severity is the outcome of the pathogenesis process beyond infection. This subpopulation included 89.5% of inpatients (age: 24–98 years; men and women 1.9:1). However, in our findings, CD risk categories were conditioned explicitly by age, with an age threshold of 56 years ( $p = 0.001–0.008$ ). Moreover, an age cutoff at 46 years associated with sex was determined as a second significant risk level with some chronic diseases. Similar to other reports, mortality associated with CD increased with age (40, 44), and individuals with hypertension and diabetes, adjusted by their implication on comorbidities, had a differential increase in infection and mortality risk (44, 45, 48, 49). Moreover, sex implication for CD and COVID-19 fatality outcomes are recognized, but not a clear-cut specific association (11, 41, 49). In our results, women exhibited a higher vulnerability to death associated with metabolic diseases (i.e., diabetes, obesity, immunosuppressant, and CKD). In contrast, men showed a higher vulnerability to respiratory diseases (i.e., COPD, asthma, and smoking), even though infection in diabetic individuals was more than 3-fold concerning women.

The CD factor in our research framework was based on the suitability of the Mexican population due to the high incidence of metabolic and cardiovascular chronic diseases (14, 15). However, the analyzed metadata ( $N = 581,580$ ) accounted for 16.2, 12.5, and 16.3% of obesity, diabetes, and hypertension, respectively, which is in contrast with the 40.2, 10.6, and 13.4% of the last official survey specifically designed to estimate the status of CD ( $n = 120,843$ ) (50). When conceding that the slightly lower prevalence of diabetes and hypertension, and higher prevalence of obesity in the official data were the proper estimations, such values may not change our fundamental findings. Specifically, the independent SARS-CoV-2 infection probability and age are significant factors in shaping the infection risk.

These findings shape the classical paradigm of the preexistence of a specific susceptible population for the occurrence of epidemics. This may be true for diseases framed by long host–pathogen coevolutionary processes and endemicity but not for pathogens encountering a new host. The SARS-CoV-2 strain diversity and mutational patterns through time and space (51, 52), as well as the parasitic fitness switch from aggressivity to spreading survival, appear to be indicators of an early evolutionary process involving a pathogen obligated to survive on the host (53). In this development, vaccination as a massive host intervention has played a minor role in comparison to host genetics and health attributes of the population itself, as inferred from this study and many clinical studies (8, 9, 51, 54, 55).

The spread of SARS-CoV-2 and pathogenicity support the rationality of these findings. The airborne virus spreading, which is the main contagious mechanism through respiratory droplets and, to a lesser extent, via aerosols (56, 57), is not host-target specific, which allows the virus acquisition by any individual upon inoculum exposure (58). Primary infection requires upper respiratory tissues for rapid multiplication before host internalization (59–61). This pathway is mediated through high angiotensin-converting enzyme 2 (ACE2) receptor expression in epithelial cells lining salivary gland ducts (60, 62), and other respiratory tissues, heart, and gastrointestinal tracts but with lower expression and infectivity (8, 55, 63). The coding gene of ACE2 is constitutive to the human

genome with low protein-coding variability and no differential expression due to sex, age, or population (55, 64, 65).

Therefore, we postulate that infection with SARS-CoV-2 originates from random virus exposures rather than a specific health condition. Infection is the first stage of pathogenicity involving virus–host recognition and entry into epithelial cells to initiate virus multiplication (61). Infection may not lead to disease, as asymptomatic conditions imply (21). This scenario departs from the general usage of infection as equivalent to disease or severity [e.g., (51)]. Once the virus infection is established, health, genetics, and other determinants may play a role in the COVID-19 outcome, including asymptomatic and severe courses with acute respiratory distress syndrome, multiorgan involvement, and death (9). However, at least at the early virus replication stage, it follows an evolutionarily conserved path common to viruses, thus allowing for unrestricted multiplication (61). Current epigenetic studies have shown that ACE2 hypomethylation in the nasal epithelium can lead to increased SARS-CoV-2 infectivity and COVID-19 severity via a greater abundance of ACE2 receptors (7, 8). A meta-analysis of plasma ACE2 also demonstrated that elevated ACE2 levels had a causal relationship with COVID-19 infection, severity, and hospitalization and that a solid X-linked locus associated with ACE2 may explain sex differences in ACE2 expression across various tissues (51).

Although the framework of this extensive study was the high occurrence of obesity/overweight (33–60%), hypertension (32–45%), and diabetes (3.1–10.6%) in the Mexican population (14, 15, 50), as well as one of the highest lethality rates (12.3%), further epidemiological studies may be needed to unveil the driving question of this research. The inclusion of diverse core populations, as implied by contrasting fatalities and epidemic rates of selected countries in this study (Figure 2), may provide advanced insights when considering ethnicity and geographical disparities, coupled with significant genomic data and health determinants. However, these results encourage the imperative need for communitarian approaches to develop preventive surveillance systems. The development of algorithms to address ambulatory populations may improve COVID-19 management and cope with zoonotic threats, without assuming a specific susceptible subpopulation that is reached through clinical or hospital settings (21). Our results may also support the benefit of massive ambulatory SARS-CoV-2 testing conducted for several countries during the critical contagious stage (58), rather than using digital risk assessment or directing tests on individuals upon presumptive COVID-19 symptoms to assist disease control treatment (66–70). It is well known that asymptomatic individuals, estimated at 22.1% under lockdown conditions (58), may exhibit a comparable virus titer to those with symptoms and thus could play a significant role in transmission chains (21). A web-app surveillance platform, linked to testing at clustering labor, social, and household environments, may overcome the cost-time factors of massive testing and effectively accomplish the confinement strategy and clinical monitoring at the community level (21). Although WHO and many countries have recently declared the end of COVID-19 as public health emergency (2), the risk of new variants and emerging diseases should encourage us to continue our comprehension of this epidemic to enhance local and global preventive health systems.

## Conclusion

Based on 24.4 million metadata records associated with 509,539 official RT-qPCR cases accumulated during the onset-exponential phase of the first epidemic wave in Mexico, we provided robust epidemiological evidence to support our hypothesis that SARS-CoV-2, a novel pathogen to the human population, did not encounter a susceptible subpopulation with a specific set of health condition for the infection establishment and epidemic development. However, the clinical evolution of COVID-19, such as disease severity and mortality, was associated with vulnerability factors explicitly conditioned by age and sex, as has been extensively published. The differentiation of infection, as the process of the successful virus, entering and early multiplication in the host, independent of the disease outcome, was fundamental in this research to primarily account for an ambulatory and hospitalized cohort. The specific selection of the onset-exponential phase of the first epidemic wave was also essential to assess the cohort risk structure based on the assumptions of random population exposure to the virus due to the fast spreading of the virus (lethality rate = 12.3%,  $R_0 > 1$ ), limited and unsteady immunological response, pathogen capabilities to evade or subvert host defense mechanisms, constrained clinical knowledge for treatment, and unprepared health systems. These findings encourage the addressing of communitarian approaches to develop preventive surveillance systems to target ambulatory populations. Such systems may complement conventional and specific surveillance platforms, such as SUIVE (<https://sinave.gob.mx/>) or SISVER (<https://sisver.sinave.gob.mx/influenza/>), respectively, that are currently in operation in Mexico. This view may effectively intervene in COVID-19, which remains a global health risk, and potential zoonotic threat without assuming a specific susceptible subpopulation targeted by new pathogens with no signals at the human coevolutionary microbiological core. To our knowledge, this is the first work addressing this fundamental epidemiological question.

## Limitations

The limitation of this research was derived from SARS-CoV-2 diagnostic data upon presumptive COVID-19 symptoms or associations with infected individuals. Therefore, the database does not represent an entirely random sampling of the ambulatory population. Despite the high lethality rate observed during the addressed epidemic phase, the epidemic rate was lower compared to many countries, thus restricting the sampling size and health structure of the studied population. Data on social, behavioral, and environmental determinants and cases with asymptomatic conditions were unavailable. Although confinement was not mandatory in Mexico, restricted activities limited the children and young people's movements, thus preventing data of these cohorts despite reports of less susceptibility (58).

## Data availability statement

The original contributions presented in the study are included in the article/[Supplementary material](#), further inquiries can be directed to the corresponding author.

## Author contributions

GM-A was responsible for the conception, conceptualization, and design of the study. JC-C and GA-S were responsible for the data acquisition. GA-S and GM-A were responsible for the statistical analyses. GM-A, GA-S, and IÁ-M were responsible for the preparation of the manuscript. All authors contributed to the article and approved the submitted version.

## Funding

Postgraduate College, Montecillo Campus, supported this research.

## Acknowledgments

The authors acknowledge the team of the Laboratory of Epidemiological Risk Analysis (CP-LANREF) of Postgraduate College for their technical support and the Mexican Ministry of Health for making possible the use of official data. The authors thank Miranda Mora-Gutiérrez for proofreading the manuscript.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpubh.2023.995602/full#supplementary-material>



## References

- John Hopkins Coronavirus Resource Center. (2023). Available online at: <https://coronavirus.jhu.edu/us-map>. (accessed March 10, 2023).
- World Health Organization. (2023). Available online at: <https://www.who.int/about>. (accessed May 17, 2023).
- Wang D, Hu B, Hu C, Zhu F, Liu X, Zhang J, et al. Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China. *JAMA*. (2020) 323:1061–9. doi: 10.1001/jama.2020.1585
- Nextstrain. (2022). *Real-Time Tracking of Pathogen Evolution*. Available online at: <https://nextstrain.org/> (accessed July 2022).
- Phuong J, Riches NO, Madlock-Brown C, Duran D, Calzoni L, Espinoza JC, et al. Social determinants of health factors for gene-environment COVID-19 research: challenges and opportunities. *Adv Genet*. (2022) 3:2100056. doi: 10.1002/ggn.2.202100056
- Moein S, Nickaeen N, Rooitani A, Borhani N, Heidary Z, Javanmard SH, et al. Inefficiency of SIR models in forecasting COVID-19 epidemic: a case study of Isfahan. *Sci Rep*. (2021) 11:1–9. doi: 10.1038/s41598-021-84055-6
- Kianmehr A, Faraoni I, Kucuk O, Mahrooz A. Epigenetic alterations and genetic variations of angiotensin-converting enzyme 2 (ACE2) as a functional receptor for SARS-CoV-2: potential clinical implications. *Eur J Clin Microbiol Infect Dis*. (2021) 40:1587–98. doi: 10.1007/s10096-021-04264-9
- Cardenas A, Rifas-Shiman SL, Sordillo JE, DeMeo DL, Baccarelli AA, Hivert M-F, et al. DNA methylation architecture of the ACE2 gene in nasal cells of children. *Sci Rep*. (2021) 11:7107. doi: 10.1038/s41598-021-86494-7
- Velavan TP, Pallerla SR, Rüter J, Augustin Y, Kremsner PG, Krishna S, et al. Host genetic factors determining COVID-19 susceptibility and severity. *EBioMedicine*. (2021) 72:103629. doi: 10.1016/j.ebiom.2021.103629
- Alanazi SA, Kamruzzaman MM, Alruwaili M, Alshammari N, Alqahtani SA, Karime A. Measuring and preventing COVID-19 using the SIR model and machine learning in smart health care. *J Healthc Eng*. (2020) 8857346. doi: 10.1155/2020/8857346
- Cummings MJ, Baldwin MR, Abrams D, Jacobson SD, Meyer BJ, Balough EM, et al. Epidemiology, clinical course, and outcomes of critically ill adults with COVID-19 in New York City: a prospective cohort study. *Lancet*. (2020) 395:1763–70. doi: 10.1016/S0140-6736(20)31189-2
- Lei S, Jiang F, Su W, Chen C, Chen J, Mei W, et al. Clinical characteristics and outcomes of patients undergoing surgeries during the incubation period of COVID-19 infection. *EClinicalMedicine*. (2020) 21:100331. doi: 10.1016/j.eclinm.2020.100331
- Ritchie H, Mathieu E, Rodés-Guirao L, Appel C, Giattino C, Ortiz-Ospina E, et al. *Coronavirus Pandemic (COVID-19)*. Published online at OurWorldInData.org (2020). Available online at: <https://ourworldindata.org/coronavirus> (accessed July 09, 2022).
- Organization for Economic Co-operation and Development. (2021). Available online at: [www.oecd.org/health/obesity-update.htm](http://www.oecd.org/health/obesity-update.htm) (accessed June 13, 2021).
- General Directorate of Epidemiology (2019). Available online at: <https://www.gob.mx/cms/uploads/attachment/file/541112/InformeDMT23erTrim2019.pdf> (accessed June 13, 2021).
- Denova-Gutiérrez E, Vargas-Chanes D, Hernández S, Muñoz-Aguirre P, Napier D, Barquera S. Linking socioeconomic inequalities and type 2 diabetes through obesity and lifestyle factors among Mexican adults: a structural equations modeling approach. *Salud Publ Mex*. (2020) 62:192–202. doi: 10.21149/10819
- Ayón-Aguilar J, Méndez-Martínez S, Toledo-Tapia R, García-Flores MA, Mayoral-Ortiz A, Tlecuitl-Mendoza N, et al. Influence of risk factors on mortality from COVID-19. *Rev Med Inst Mex Seguro Soc*. (2022) 4:433–9.
- Suárez V, Suarez QM, Oros RS, Ronquillo de JE. Epidemiology of COVID-19 in Mexico: from the 27th of February to the 30th of April 2020. *Rev Clin Esp*. (2021) 220:463–71. doi: 10.1016/j.rce.2020.05.007
- Parra-Bracamonte GM, Lopez-Villalobos N, Parra-Bracamonte FE. Clinical characteristics and risk factors for mortality of patients with COVID-19 in a large data set from Mexico. *Ann Epidemiol*. (2020) 52:93–8. doi: 10.1016/j.annepidem.2020.08.005
- Bello-Chavolla OY, Bahena-López JP, Antonio-Villa NE, Vargas-Vázquez A, González-Díaz A, Márquez-Salinas A, et al. Predicting mortality due to SARS-CoV-2: a mechanistic score relating obesity and diabetes to COVID-19 outcomes in Mexico. *J Clin Endocrinol Metab*. (2020) 105:346. doi: 10.1101/2020.04.20.20072223
- Mora-Aguilera G, Martínez-Bustamante V, Acevedo-Sánchez G, Coria-Contreras JJ, Guzmán-Hernández E, Flores-Colorado OE, et al. Surveillance web system and mouthwash-saliva qPCR for labor ambulatory SARS-CoV-2 detection and prevention. *Int J Environ Res Public Health*. (2022) 19:1271. doi: 10.3390/ijerph19031271
- Ministry of Health Mexico. *Daily Technical Communication COVID-19*. (2020). Available online at: <https://coronavirus.gob.mx/> (accessed June 30, 2020).
- O'Toole Á, Hill V, Pybus OG, Watts A, Bogoch II, Khan K, et al. Tracking the international spread of SARS-CoV-2 lineages B.1.1.7 and B.1.351/501Y-V2 (version 1; peer review: 3 approved). *Wellcome Open Res*. (2021) 6:121. doi: 10.12688/wellcomeopenres.16661.1
- Taboada B, Vazquez-Perez JA, Muñoz-Medina JE, Ramos-Cervantes P, Escalera-Zamudio M, Boukadida C, et al. Genomic analysis of early SARS-CoV-2 variants introduced in Mexico. *J Virol*. (2020) 31:e101056–20. doi: 10.1128/JVI.101056-20
- General Directorate of Epidemiology. (2021). Available online at: <https://coronavirus.gob.mx/datos/> (accessed June 13, 2021).
- The National Institute of Statistics and Geography. (2019). *Press Release No. 475/19*. Available online at: [https://www.inegi.org.mx/contenidos/saladeprensa/aproposito/2019/edad2019\\_Nal.pdf](https://www.inegi.org.mx/contenidos/saladeprensa/aproposito/2019/edad2019_Nal.pdf) (accessed June 13, 2021).
- Lewis JR. *An Introduction to Classification and Regression Tree (CART) Analysis. Annual Meeting of the Society for Academic Emergency Medicine in San Francisco, California*. (2000). Available online at: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.95.4103&rep=rep1&type=pdf> (accessed July 09, 2022).
- Fonarow GC, Adams KF, Abraham WT, Yancy CW, Boscardin WJ, et al. Risk stratification for in-hospital mortality in acutely decompensated heart failure: classification and regression tree analysis. *JAMA*. (2005) 293:572–80. doi: 10.1001/jama.293.5.572
- Low M, Joy KM, Makan T. Using regression trees to predict patterns of male provisioning in the stitchbird (hihi). *Anim Behav*. (2006) 71:1057–68. doi: 10.1016/j.anbehav.2005.07.021
- Takahashi O, Cook EF, Nakamura T, Saito J, Ikawa F, Fukui T. Risk stratification for in-hospital mortality in spontaneous intracerebral haemorrhage: a classification and regression tree analysis. *QJM-Int J Med*. (2006) 99:743–50. doi: 10.1093/qjmed/hcl107
- Subramanian A, Nirantharakumar K, Hughes S, Myles P, Williams T, Gokhale KM, et al. Symptoms and risk factors for long COVID in non-hospitalized adults. *Nat Med*. (2022) 28:1706–14. doi: 10.1038/s41591-022-01909-w
- Alharbi A, Rabbani SI, Halim Mohamed AA, Almushayti BK, Aldhwayan NI, Almohaimed AT, et al. Analysis of potential risk factors associated with COVID-19 and hospitalization. *Front Public Health*. (2022) 10:921953. doi: 10.3389/fpubh.2022.921953
- Goodall JW, Reed TAN, Ardissino M, Bassett P, Whittington AM, Cohen DL, et al. Risk factors for severe disease in patients admitted with COVID-19 to a hospital in London, England: a retrospective cohort study. *Epidemiol Infect*. (2020) 148:1–8. doi: 10.1017/S0950268820002472
- Khan DM, Ali M, Iqbal N, Khalil U, Aljohani HM, Alharthi AS, et al. Short-term prediction of COVID-19 using novel hybrid ensemble empirical mode decomposition and error trend seasonal model. *Front Public Health*. (2022) 10:922795. doi: 10.3389/fpubh.2022.922795
- Somyanonthanakul R, Warin K, Amasiri W, Mairiang K, Mingmalairak C, Panichkitkosolkul W, et al. Forecasting COVID-19 cases using time series modeling and association rule mining. *BMC Med Res Methodol*. (2022) 22:281. doi: 10.1186/s12874-022-01755-x
- European Covid-19 Forecast Hub. *Forecasts*. (2022). Available online at: <https://covid19forecasthub.eu/visualisation.html> (accessed July 09, 2022).
- Bartolomeo N, Trerotoli P, Serio P. Short-term forecast in the early stage of the COVID-19 outbreak in Italy. Application of a weighted and cumulative average daily growth rate to an exponential decay model. *Infect Dis Model*. (2021) 6:212–21. doi: 10.1016/j.idm.2020.12.007
- Jain V, Yuan JM. Predictive symptoms and comorbidities for severe COVID-19 and intensive care unit admission: a systematic review and meta-analysis. *Int J Public Health*. (2020) 65:533–46. doi: 10.1007/s00038-020-01390-7
- World Health Organization. (2023). Available online at: [https://www.who.int/news/item/05-05-2023-statement-on-the-fifteenth-meeting-of-the-international-health-regulations-\(2005\)-emergency-committee-regarding-the-coronavirus-disease-\(covid-19\)-pandemic](https://www.who.int/news/item/05-05-2023-statement-on-the-fifteenth-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-coronavirus-disease-(covid-19)-pandemic) (accessed May 10, 2023).
- Kim J, Blaum C, Ferris R, Arcila-Mesa M, Do H, Pulgarin C, et al. Factors associated with hospital admission and severe outcomes for older patients with COVID-19. *J Am Geriatr Soc*. (2022) 1:12. doi: 10.1111/jgs.17718
- Covino M, De Matteis G, Burzo ML, Russo A, Forte E, Carnicelli A, et al. Predicting in-hospital mortality in COVID-19 older patients with specifically developed scores. *J Am Geriatr Soc*. (2021) 69:37–43. doi: 10.1111/jgs.16956
- Shahid Z, Kalayanamitra R, McClafferty B, Kepko D, Ramgobin D, Patel R, et al. COVID-19 and older adults: what we know. *J Am Geriatr Soc*. (2020) 68:926–9. doi: 10.1111/jgs.16472
- Wu C, Chen X, Cai Y, Xia J, Zhou X, Xu S, et al. Risk factors associated with acute respiratory distress syndrome and death in patients with coronavirus disease 2019 pneumonia in Wuhan, China. *JAMA Intern Med*. (2020) 180:934–43. doi: 10.1001/jamainternmed.2020.0994

44. Asai Y, Nomoto H, Hayakawa K, Matsunaga N, Tsuzuki S, Terada M, et al. Comorbidities as risk factors for severe disease in hospitalized elderly COVID-19 patients by different age-groups in Japan. *Gerontology*. (2022) 68:1027–37. doi: 10.1159/000521000
45. Guan WJ, Liang WH, Zhao Y, Liang H, Chen Z, Li Y, et al. Comorbidity and its impact on 1590 patients with COVID-19 in China: a nationwide analysis. *Eur Respir J*. (2020) 55:2000547. doi: 10.1183/13993003.01227-2020
46. Jordan RE, Adab P, Cheng KK. COVID-19: risk factors for severe disease and death. *BMJ*. (2020) 368:m1198. doi: 10.1136/bmj.m1198
47. Li B, Yang J, Zhao F, Zhi L, Wang X, Liu L, et al. Prevalence and impact of cardiovascular metabolic diseases on COVID-19 in China. *Clin Res Cardiol*. (2020) 109:531–8. doi: 10.1007/s00392-020-01626-9
48. Terada M, Ohtsu H, Saito S, Hayakawa K, Tsuzuki S, Asai Y, et al. Risk factors for severity on admission and the disease progression during hospitalisation in a large cohort of patients with COVID-19 in Japan. *BMJ Open*. (2021) 11:e047007. doi: 10.1136/bmjopen-2020-047007
49. Yang J, Zheng Y, Gou X, Pu K, Chen C, Guo Q, et al. Prevalence of comorbidities and its effects in patients infected with SARS-CoV-2: a systematic review and meta-analysis. *Int J Infect Dis*. (2020) 94:91–5. doi: 10.1016/j.ijid.2020.03.017
50. Shamah-Levy T, Romero-Martínez M, Barrientos-Gutiérrez T, Cuevas-Nasu L, Bautista-Arredondo S, Colchero MA, et al. *Encuesta Nacional de Salud y Nutrición 2020 sobre Covid-19. Resultados nacionales*. Cuernavaca: Instituto Nacional de Salud Pública (2021). ISBN 978-607-511-212-1.
51. Yang Z, Macdonald-Dunlop E, Chen J, Zhai R, Li T, Richmond A, et al. Genetic landscape of the ACE2 coronavirus receptor. *Circulation* 3;145. (2022) 1398–411.
52. Lythgoe KA, Hall M, Ferretti L, de Cesare M, MacIntyre-Cockett G, Trebes A, et al. SARS-CoV-2 within-host diversity and transmission. *Science*. (2021) 16:eabg0821. doi: 10.1126/science.abg0821
53. Mora-Aguilera G, Acevedo-Sánchez G. A retrospective analysis of plant and human epidemics for COVID-19 comprehension. *Mex J Phytopathol*. (2021) 39:1–91. doi: 10.18781/R.MEX.FIT.2021-27
54. Xia X. Extreme genomic CpG deficiency in SARS-CoV-2 and evasion of host antiviral defense. *Mol Biol Evol*. (2020) 37:2699–705. doi: 10.1093/molbev/msaa094
55. Li MY, Li L, Zhang Y, Wang XS. Expression of the SARS-CoV-2 cell receptor gene ACE2 in a wide variety of human tissues. *Infect Dis Poverty*. (2020) 9:45. doi: 10.1186/s40249-020-00662-x
56. Meyerowitz EA, Richterman A, Gandhi RT, Sax PE. Transmission of SARS-CoV-2: a review of viral, host, and environmental factors. *Ann Intern Med*. (2021) 174:69–79. doi: 10.7326/M20-5008
57. Zhou L, Aye SK, Chidambaram V, Karakousis CP. Modes of transmission of SARS-CoV-2 and evidence for preventive behavioral interventions. *BMC Infect Dis*. (2021) 21:496. doi: 10.1186/s12879-021-06222-4
58. Sun K, Wang W, Gao L, Wang Y, Luo K, Ren L, et al. Transmission heterogeneities, kinetics, and controllability of SARS-CoV-2. *Science*. (2021) 371:eabe2424. doi: 10.1126/science.abe2424
59. To KK-W, Tsang OT-Y, Yip CC-Y, Chan K-H, Wu T-C, Chan JM-C, et al. Consistent detection of 2019 novel coronavirus in saliva. *Clin Infect Dis*. (2020) 71:841–3. doi: 10.1093/cid/ciaa149
60. Wang W-K, Chen S-Y, Liu I-J, Chen Y-C, Chen H-L, Yang C-F, et al. Detection of SARS-associated coronavirus in throat wash and saliva in early diagnosis. *Emerg Infect Dis*. (2004) 10:1213–9. doi: 10.3201/eid1007.031113
61. Sanyal S. How SARS-CoV-2 (COVID-19) spreads within infected hosts - what we know so far. *Emerg Top Life Sci*. (2020) 4:371–8. doi: 10.1042/ETLS20200165
62. Güçlü E, Koroglu M, Yürümez Y, Toptan H, Kose E, Güneysu F, et al. Comparison of saliva and oro-nasopharyngeal swab sample in the molecular diagnosis of COVID-19. *Rev Assoc Med Bras*. (2020) 66:1116–21. doi: 10.1590/1806-9282.66.8.1116
63. Zou L, Ruan F, Huang M, Liang L, Huang H, Hong Z, et al. SARS-CoV-2 viral load in upper respiratory specimens of infected patients. *N Engl J Med*. (2020) 382:1177–9. doi: 10.1056/NEJMc2001737
64. Cao Y, Li L, Feng Z, Wan S, Huang P, Sun X, et al. Comparative genetic analysis of the novel coronavirus (2019-nCoV/SARS-CoV-2) receptor ACE2 in different populations. *Cell Discov*. (2020) 6:11. doi: 10.1038/s41421-020-0147-1
65. Stawiski EW, Diwanji D, Suryamohan K, Gupta R, Fellouse FA, et al. Human ACE2 receptor polymorphisms predict SARS-CoV-2 susceptibility. *Commun Biol*. (2020) 4:475. doi: 10.1101/2020.04.07.024752
66. Canas LS, Sudre CH, Capdevila-Pujol J, Polidori L, Murray B, Molteni E, et al. Early detection of COVID-19 in the UK using self-reported symptoms: a large-scale, prospective, epidemiological surveillance study. *Lancet Digit Heal*. (2021) 3:e587–98. doi: 10.1016/S2589-7500(21)00131-X
67. Drew DA, Nguyen LH, Steves CJ, Menni C, Freydin M, Varsavsky T, et al. Rapid implementation of mobile technology for real-time epidemiology of COVID-19. *Science*. (2020) 368:1362–7. doi: 10.1101/2020.04.02.20051334
68. Menni C, Valdes AM, Freydin MB, et al. Real-time tracking of self-reported symptoms to predict potential COVID-19. *Nat Med*. (2020) 26:1037–40. doi: 10.1038/s41591-020-0916-2
69. Khaloufi H, Abouelmehdi K, Rustam F, Jurcut AD, Lee E, Ashraf I. Deep learning based early detection framework for preliminary diagnosis of COVID-19 via onboard smartphone sensors. *Sensors*. (2021) 21:6853. doi: 10.3390/s212106853
70. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet*. (2020) 395:1054–62. doi: 10.1016/S0140-6736(20)30566-3