Temporal analysis of covid-19 in Colombia: associated indicators and modeling

Análisis temporal del covid-19 en Colombia: indicadores asociados y modelización

Hugo Alexander Rondón Quintana¹, Carlos Alfonso Zafra Mejía²

Abstract

Introduction. This study shows statistical information regarding COVID-19 in Colombia up to this date (March 1-2022). Specifically, the daily, monthly and cumulative evolution of infections and deaths, correlated with the distribution of the population according to age and gender. Objective. Show statistical information about COVID-19 that allows help to plan and design, in future Pandemics, public health policy strategies in Colombia. Methods. Daily information since the official declaration of Pandemic in Colombia (March 16 – 2020) was obtained by the National Health Institute (INS) and was organized in a database in order to conduct respective analysis. This information was compared to similar studies obtained based on the bibliographical review. Results and Conclusions. Results and conclusions are similar to those found in the reference literature: most part of those dead by COVID-19 are of senior age and male gender. Regarding Case Fatality Rate (CFR), it notoriously increases with age. The most vulnerable population displays an average age of ≥ 52.8 years. The less vulnerable population are young persons under 30 years of age, but specifically, those within the age range of 10 and 20 years. Gompertz and Logistic models can mathematically simulate the evolution of deaths and the evolution of CFR according to age.

Keywords: COVID-19, case fatality rate (CFR), Colombia, statistic, modeling, pandemic.
Resumen

Introducción. Este estudio muestra información estadística sobre el COVID-19 en Colombia a la fecha (1 de marzo de 2022). Específicamente, la evolución diaria, mensual y acumulada de contagios y defunciones, correlacionada con la distribución de la población según edad y sexo. Objetivo. Mostrar información estadística sobre COVID-19 que permita ayudar a planificar y diseñar, en futuras Pandemias, estrategias de política de salud pública en Colombia. Metodología. La información diaria desde la declaratoria oficial de Pandemia en Colombia (16 de marzo de 2020) fue obtenida del Instituto Nacional de Salud (INS) y fue organizada en una base de datos para realizar los análisis respectivos. Esta información se comparó con estudios similares obtenidos a partir de revisión bibliográfica. Resultados y conclusiones. Los resultados y conclusiones son similares a los encontrados en la literatura de referencia: la mayor parte de los fallecidos por COVID-19 son de edad avanzada y sexo masculino. En cuanto a la tasa de letalidad (CFR), ésta aumenta notoriamente con la edad. La población más vulnerable presenta una edad promedio ≥ 52.8 años. La población menos vulnerable son los jóvenes menores de 30 años, pero específicamente, los que se encuentran en el rango de edad de 10 y 20 años. Los modelos Gompertz y Logistic pueden simular matemáticamente la evolución de las muertes y la evolución de la CFR según la edad.

Palabras clave: COVID-19, tasa de letalidad (TL), Colombia, estadística, modelización, pandemia.

Introduction

COVID-19 is a disease produced by the SARS-CoV-2 coronavirus. Up to this date (March 1 – 2022) throughout the world, 437,792,328 persons have been diagnosed with the disease, 5,978,217 (1.37%) have died, 369,500,307 (84.4%) have recovered and 62,313,804 (14.2%) are under study or in recovery. If this disease is compared to other diseases such as the Middle East Respiratory Syndrome (MERS) and Severe Acute Respiratory Syndrome (SARS) (which are also produced by other coronaviruses), is less lethal. Furthermore, most of COVID-19 cases are light, with a great number of persons that can be asymptomatic, others easily recover and most do not need hospital care (1-3). However, SARS-CoV-2 is spreading faster, generating great-
ter numbers of infections and possible risks of death (reproduction number $R_0 = 3.28, 3.0$ and $1.0$ for COVID-19, SARS and MERS, respectively) (4-5). Estimated Case Fatality Rate – CFR for COVID-19 varies between $2.2$ and $3.0\%$, SARS between $9.6$ and $15.0\%$ and MERS between $34.4$ and $35.8\%$ (2, 6-8).

A more precise indicator than CFR is infection fatality ration (IFR), given that CFR could overestimate lethality of COVID-19 (9). Its estimation is difficult given that many COVID-19 cases are asymptomatic or develop light symptoms and these are not taken into consideration. Additionally, the standard test for COVID-19 involves polymerase chain reaction testing (PCR) of nasopharyngeal swabs from patients suspected of having contracted the virus, which can produce some false negatives (10-13). Some IFR values obtained by several researchers are $0.68\%$ (13), $1.38\%$ (15), $0.5\%$ (16), $0.8\%$ (17), $0.3\%$ (18), $0.9\%$ (19), $0.64\%$ (20), $1.03\%$ (21), $1.096\%$ (22). These values are dispersed (varying from $0.3\%$ up to $1.38\%$). According to (14), IFR $< 1\%$ makes closing schools and social distancing not profitable. IFR and CFR must be analyzed according to age ranges. According to (9) IFR $= 0.001\%$ for children under age 15 and $1.5\%$ for persons age $\geq 70$ years. These researchers also report an average CFR/IFR ratio of $= 23$. However, this ratio also changed with age: CFR/IFR $= 100$ (children age $< 15$ years), $60$ (15-19 years), $25$ (20-29 years) and between $17-21$ (persons $\geq 30$ years). Based on four studies analyzed and extracted from (22), the average estimated IFR changed according to age range $(8, 0.7, 0.08$ and $0.008\%$ for persons ages $80, 60, 40$ and $20$ years, respectively).

Various studies have been carried out and are daily published throughout the world in relation to COVID-19. Despite this, it is still difficult to obtain reliable information on the matter, given that social networks and communications media on some occasions show an erroneous infodemic, which misinforms. This exposure to infodemic managed to affect mental health of people and impacting on negative responses to the disease (23-27). In general, a fully validated aspect regarding the disease ever since the beginning of the Pandemics is that the main risk factors are associated with people of senior age, male gender and with comorbidities (e.g., arterial hypertension, diabetes mellitus, kidney, cardiovascular, cerebrovascular and chronic obstructive pulmonary diseases, cancer, asthma, acquired immune deficiency syndrome, among others) (21, 26, 28-39). Likewise, malnutrition, (one of the main causes of immune deficiency throughout the world), obesity and tobacco use (which have a strong relation to respiratory diseases), are directly associated to deaths by COVID-19 (26, 36, 40-46). These conclusions are not new. According to (47-72) several epidemiological studies have demonstrated that people with
ages above 65 and with prior comorbidities have a greater risk of developing severe complications and dying due to other diseases catalogued as less lethal, such as influenza, common flu, as well as malnutrition, abuse, abandon, mental health, etc. Combining aspects associated to age and comorbidities, some researchers conclude that biological age, more so than chronological age of affected patients could be the critical factor in order to systematically assess COVID-19 infections and avoiding excess mortality (53). Also, several studies agree that the gravity of this disease is directly associated with low temperature and environmental contamination (54).

There are other aspects regarding the disease that are ambiguous. Some studies mention that the disease poses greater risks in places with more poverty and inequality (35, 38, 55-57). However, others conclude the opposite. Specially, a phenomenon known as the “African Paradox” (58) shows that in African countries death cases were less in comparison to European countries and United States of America (59-62). Some hypotheses suggest that in other countries, a younger population, a warmer and more humid climate, prior exposure to other coronaviruses (presence of antibodies), and the possible quicker accomplishment of herd immunity and prior experience in management of other infectious diseases could have reduced COVID-19 death risks (61, 63-66).

In countries such as Colombia, COVID-19 could produce a greater impact than in others, possibly because: i) its health infrastructure, lack of intensive care units (ICUs) (36-37); ii) high rate of informal employment and social inequality (67); iii) hospital capacity for facing a new infectious disease (28, 38); iv) delay in notifying symptoms in affected patients (average of 9.1 days) (26, 68), along with a delay in the treatment of symptoms; v) prevalence of comorbidities along with chronic malnutrition (36). Furthermore, some studies in Colombia (69) conclude that hospital occupation is normally high (especially in flu seasons) putting the population at risk before Pandemics and events with multiple victims (natural disasters, earthquakes etc.). Additionally, in Colombia, risk factors for a senior population are two or three times greater in comparison to countries of higher incomes, and increase when people live in geographically remote areas and areas that are economically and socially depressed (70). On the contrary, an aspect that could have reduced the evolution of deaths in Colombia is its younger population in comparison to European countries (39).

The foregoing study had the main objective of conducting a statistical analysis on the evolution of COVID-19 in Colombia. This information was associated to infections, deaths, age and gender of people who were infected with the disease. This information can help to plan and design, in future Pan-
demics, public health policy strategies. Similar studies have been conducted on the matter throughout the world, however, no document has consulted, organized summarized and analyzed information such as the one presented in this paper. Additionally, updated information is presented in comparison to that of other documents. The results presented can be a source of consultation for academics, researchers and research groups on the matter, as well as for institutions that dedicate to public health policy planning.

**Methods**

The type of research design in this study was descriptive observational. Initially, a bibliographical review was carried out on the subject, mainly in the ScienDirect scientific database. In this database, today (March 1 – 2022), if you type the word COVID-19, 113,442 results or documents are displayed in order to be consulted. These increase on a daily basis due to the high number of studies that have been carried out on the matter. In order for the analysis, the ones chosen were those that had a greater affinity with the object of study of this paper: statistical information on the evolution of infections and deaths by COVID-19 in the world in order to compare these with the case of Colombia. Additionally, articles published in high impact journals were mainly chosen (Q1 or Q2 according to SCImago Journal Rank and Scopus).

The data that was statistically analyzed for this study mainly comes from the Ministry of Health and Social Protection of Colombia (MinSalud), which daily publishes official information concerning COVID-19 through the National Health Institute (INS). Information obtained from INS was related to new daily confirmed cases (DNC), daily confirmed deaths (DD), age and gender of infected and deceased, comorbidities of deceased, lastly, city and department of the deceased. This information was organized in a database ever since the declaration of Pandemics (March 16 – 2020) to this date (March 1 – 2022). Accumulating DNC and DD values, there was an estimation of the total coronavirus cases (CC) and the total number of deaths (TND), respectively, as well as CFR=TND/CC in percentage. The epidemiological curve was mathematically modelled the evolution of TND in Colombia throughout time. For the case of deceased persons, the sample analyzed was of 138,612 (99.89% of total deaths registered up to the date of March 1 – 2022). The information registered in the database was organized and presented in Tables and Figures in order to facilitate its analysis.

**Results**

The evolution of confirmed daily new cases (DNC) of coronavirus SARS-CoV-2 and the daily number of deceased (DD) across time are shown in Figures 1a and 1b, respectively. If cumulative DNC and DD values
are shown, we would obtain the number of confirmed coronavirus cases (CC) and the total number of deceased (TND), respectively (see Figures 1c and 1d). An exponential growth can be observed in CC and TND until reaching each peak, followed by a brief stabilization phase each time the epidemiological peak dropped. Up to this date (March 1 – 2022), TND is of 138,767 persons of which 60.8% are male and 39.2% female gender, respectively.

Figures 1c and 1d can be mathematically represented using Gompertz [1] and Logistic [2] equations. These equations have been used throughout several studies in the area of biology and medicine. Furthermore, they have even been used to attempt to predict the COVID-19 in some countries (71-76). However, most part of these studies had initial information regarding the evolution of COVID-19 (scarce information), and because of that reason, the constants obtained do not represent the current evolution of the disease. For the case of Figure 1d, parameter a defines the maximum

\[ a \]

\[ b \]

\[ c \]

\[ d \]

Figure 1: a) DNC, b) DD, c) CC, d) TND evolution in the time.

Source: Hugo Alexander Rondón Quintana, Carlos Alfonso Zafra Mejía.
quasi-asymptotic value of TND, whereas $b$ moves the curve and defines its inflection. $D$ is the number of days from the first case of death. In Equation [1] (Gompertz model), based on the projected trend and obtaining $r^2=0.987$, the values $a=1.617 \times 10^5$, $b=1.81$ and $c=5.39 \times 10^{-3}$ were expected. In other words, values close to 162 thousand COVID-19 deaths could be expected. For the case of Equation [2] (Logistic model), for $r^2=0.990$ the values $a=1.434 \times 10^5$, $b=45.9$ and $c=1.012 \times 10^{-2}$ would be expected. In other words, values close to 145 thousand COVID-19 deaths were expected. In both equations with these parameters, correlations start to become strong from the fifth month of simulation onward. Simulations using both equations are shown on Figure 2.

\begin{align*}
TND &= ae^{-e^{b+cD}} \quad [1] \\
TND &= \frac{a}{1+b e^{-cD}} \quad [2]
\end{align*}

The months with greatest cases of infections and deaths in each peak of the disease were August 2020, January 2021, June 2021 and January 2022 (Figure 3). In these four months 38.5% of infections were produced and 30.6% of deaths occurred. On average, the number of daily infections (DAAC) in these months was $10,311.6 \pm 1,598$, $14,584.2 \pm 3,246$, $27,817.5 \pm 3,358$ and $23,542.6 \pm 7,359.7$ respectively (Figure 3a), while the monthly accumulated number of infections (CC-m) was $319,660$, $452,109$, $834,526$ and $729,821$ respectively (Figure 3b). On average, the total number of daily deaths (DATND) in these months was $308.3 \pm 42$, $347.4 \pm 54$, $592.3 \pm 63$ and $140.6 \pm 88$, respectively (Figure 3c), while the number of monthly accumulated number of (TND-m) was $9,558$, $10,770$, $17,770$ and $4,358$, respectively (Figure 3d).
If we divide DNC by DD, we obtain daily CFR (see Figure 4a). The evolution of daily CFR is not clear and is scattered (values constantly change from 1.5% up to 5%). On the contrary, cumulative CFR (percentage relation between CC and TND) has a clearer trend (Figure 4b), reaching a quasi-constant value of approximately 2.55% from January 2021 until December 2021. Then, it declines until reaching approximate values of 2.29% to date. This behavior is also observed when graphing TND against CC (Figure 5, with an average slope of approximately 2.5% or average CFR). This average CFR is distributed across ages in Table 1 and increases with age. This information was obtained from March 2020 to October 2021 and agrees with several stu-
dies consulted on the matter which conclude that the most important risk factor for COVID-19 is increase of age (3, 39, 77-78). The evolution of CFR with age (A in years) can be mathematically represented by using Logistic and Gompertz Equations [3] and [4]; $r^2=0.9995)$. $a=40.33$, $b=3411.8$ and $c=0.1$ for Equation [3] and $a=56.3$, $b=3.58$ and $c=0.045$ for Equation [4].
Table 1. Distribution of CC and CFR across age range.

<table>
<thead>
<tr>
<th>Age range (years)</th>
<th>CC</th>
<th>CFR (%)</th>
<th>Deviation (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>≥90</td>
<td>20360</td>
<td>34.12</td>
<td>2.18</td>
</tr>
<tr>
<td>80-89</td>
<td>92760</td>
<td>26.11</td>
<td>2.12</td>
</tr>
<tr>
<td>70-79</td>
<td>202132</td>
<td>16.57</td>
<td>1.49</td>
</tr>
<tr>
<td>60-69</td>
<td>413735</td>
<td>7.99</td>
<td>1.35</td>
</tr>
<tr>
<td>50-59</td>
<td>688396</td>
<td>2.81</td>
<td>0.40</td>
</tr>
<tr>
<td>40-49</td>
<td>847674</td>
<td>1.16</td>
<td>0.29</td>
</tr>
<tr>
<td>30-39</td>
<td>1121350</td>
<td>0.36</td>
<td>0.062</td>
</tr>
<tr>
<td>20-29</td>
<td>1063334</td>
<td>0.16</td>
<td>0.060</td>
</tr>
<tr>
<td>10-19</td>
<td>380415</td>
<td>0.05</td>
<td>0.020</td>
</tr>
<tr>
<td>0-9</td>
<td>173820</td>
<td>0.17</td>
<td>0.196</td>
</tr>
</tbody>
</table>

Source: Hugo Alexander Rondón Quintana, Carlos Alfonso Zafra Mejía.

CFR = \( \frac{a}{1+be^{-cA}} \) \[3\]

CFR = \( ae^{-bA} \) \[4\]

The average age of COVID-19 deceased in Colombia each month (AA-m) and cumulative (AAA) age are presented in Figures 6a and 6b, respectively. The monthly evolution of MVP (most vulnerable population; difference between average age and standard deviation) for each month (MVP-m) as well as cumulative (MVPA) is also depicted in Figures 7a and 7b, respectively.

Source: Hugo Alexander Rondón Quintana, Carlos Alfonso Zafra Mejía.
The mode in ages of deceased each month (Mode-m) and cumulative mode (Mode-A) is shown on Figures 8a and 8b, respectively. To date, the mode is 70 years.

The median for ages of deceased each month (Median-m) and cumulative median (Median-A) is shown in Figures 9a and 9b, respectively. To date, the median is 70 years.
The percentage of deceased according to age range is depicted in Figure 10, and the respective percentiles are shown in Figure 11. These data are coherent to the results of AAA, Mode-A and Median-A.
Discussion

Four peaks of the disease can be observed in which it had the greatest values of DNC and DD. These peaks coincide with the traditional flu peaks in Colombia (September to December and April to June, (79)).

It is observable that most death persons are of senior age (≥ 60 years). Up to this date (March 1 – 2022), the average age of deceased is of 68.2±15.4 years and the MVP has an average age of ≥ 52.8 years. An increase in age average of deceased can be observed from the beginning of the Pandemics (March – 2020) until January 2021. Then it drops until June 2021, possibly because of the following reasons: i) the youngest population was the one with most infections (see Table 1) given that it is the most active. On average, 74.5±0.69% of infections occurred in persons with ages between 20 and 60 years and 10.3 ± 0.79% under the age of 20 (in total, this population added up to almost 85% of infections), whereas only 15.3±0.71% had an age above 60 years; ii) flexibilization in mitigation measures, especially for younger persons; iii) between the months of April and August 2021, important protests against the National Government emerged, in which young persons were more exposed to infections; iv) in Colombia, young population predominates (population with ages 60 years is 13.1% approximately); v) vaccination began in Colombia on March 2021, beginning with the oldest population. However, from June 2021, the age of deceased once again increased mainly because of: i) there is still a great part of the population above 60 years of aye that has not been vaccinated in the country; ii) vaccines are less effective in older persons with comorbidities in comparison to younger population; iii) vaccina-
tion in persons below age 60 began from May 2021; iv) a great part of young people in Colombia have been infected and could have possibly reached a high level of herd immunity.

It is observable that the monthly mode of death persons has always been superior or equal to 64 years, except in the month of June 2021 (59 years). However, based on accumulated data, the mode has always been above 65 years and furthermore, it can be observed that between the months of August 2020 and April 2021, maintained at 80 years.

It is observable that the monthly median of death persons has always been superior or equal to 65 years (with a maximum peak of 80 years), except in the month of June 2021 (63 years). However, based on accumulated data, the median has always been above 66 years and furthermore, it is observable that the greatest values (80 years) were reached in the last two months (January and February 2022). The drop in mode and median from January 2021 until June 2021 could be explained by the same reasons previously mentioned for the case of AA-m.

The greater part of deceased were persons with ages ≥ 60 years (Figure 10a). Between March 2020 and April 2021, more than 72% of deceased had that age, showing the highest peaks (79.5% to 81.5%) in the months of October 2020 and March 2021. Between May 2021 and June 2021, decay can be observed as a product of the drop reported in the case of AAA, Mode-A and Median-A. However, even this age range continued being the one with greater prevalence of deaths (60%), and the trend within the last months has been once again increasing this percentage (89% in last two months). The percentile 28.5% is equivalent to 60 years, in other words, 28.5% of deceased had an age below 60 years. Additionally, if age range is analyzed from 50 years onward (Figure 10b), it can be seen that these persons contributed with between 80% and 95% of deceased, which is coherent with MVPA. On average, this age range (≥ 50 years) contributed with 88.9 ± 3.8% of deceased. The percentiles 12.5% and 5% are equivalent for 50 and 40 years, respectively. Between March 2020 and April 2021, deaths varied between 6.3% and 11% approximately for the age range of 30 to 50 years. From May 2021 to July 2021 this percentage increased up to 18% approximately, however, within the latest months, a drop is reported once again. On average, this age range (30 to 50 years) contributed with 9.5±3.5% of deaths. Regarding persons with ages under 30 years, the virus was less lethal. Between 20 and 30 years of age, deaths varied between 0.63% and 1.73% (average of 1.2±0.4%). For the case of persons under 20 years of age, deaths varied between 0.17% and 0.5% (average of 0.39 ±
0.23%), with the less vulnerable population being between the ages of 10 and 20 years (0.12±0.06%). For the case of 10, 20 and 30 years of age, respective percentiles are 0.13%, 0.28% and 1.5%.

**Conclusions**

Based on the results obtained the following is concluded:

Up to this date (March 1 – 2022), the epidemiological evolution in Colombia has gone through four peaks, which match the traditional flu seasons. The months of greatest cases of infections and deaths in each peak of the disease were January 2021, June 2021 and January 2022, representing 38.5% of infections and 30.6% of deaths. In the last peak, the age of the dead increased markedly.

The daily evolution of TND and the growth of CFR with the increase of age can be mathematically simulated using Gompertz and Logistical models. The current average CFR is around 2.5%, however, since it is distributed in ages, it notoriously increases with its increase.

Most COVID-19 deaths occur in people of senior age (the average age of deceased = 68.2±15.4 years; mode = 70 years; median = 70 years) and of male gender (60.8%). The most vulnerable population is the one that has an average age of ≥ 52.8 years, contributing to 85% of deaths (percentile 15). The least vulnerable population were young persons between the ages of 10 and 20 years, but in general terms, young population under the age of 30 has presented low mortality rates. For 10, 20, 20 and 40 years, the respective percentiles were 0.11%, 0.26%, 1.5% and 5%.

Based on the bibliographical reviewed, additional conclusions are the following: i) most persons that are infected with COVID-19 are asymptomatic, or do not experience serious symptoms; ii) most people diagnosed recover; iii) lethality of COVID-19 in terms of CFR and IFR increase with age, number and type of comorbidities in people and are greater among the male gender; iv) lethality in the disease is strongly associated to malnutrition, obesity and tobacco use.

Future studies must approach matters related to socio-economic, public health, cultural, political and environmental impacts that the COVID-19 has generated and will generate in Colombia. Additionally, it is important to compare COVID-19 lethality with regards to other diseases that are known as less lethal (e.g. influenza, flu etc.) All of the above, geared towards preparing and helping the country during a future Pandemics crisis.
Referencias


8. World Health Organization. Middle East respiratory syndrome coronavirus (MERS-CoV). Retrieved from https://applications.emro.who.int/docs/EMRPUB-CSR-241-2019-EN.pdf?ua=1%26ua=1%26ua=1%26ua=1%26ua=1%26ua=1 [Accessed 20 September 2021]


70. Guerrero N, Yépez-Ch M C. Factores asociados a la vulnerabilidad del adulto mayor con alteraciones de salud [Factors associated with the vulnerability of the elderly with health disorders]. Universidad y Salud. 2015;17(1):121-31.


