Perspectives

Open Access

Implementation of the Pareto principle in focus group generation based on global coronavirus disease morbidity and mortality rates

Mostafa Essam Eissa¹, Engy Refaat Rashed² and Dalia Essam Eissa³

Abstract

The recent pandemic that has hit the world has affected humanity in all aspects of life. Since the outbreak of this worldwide epidemic, a huge amount of data has been generated. In this article, we have provided a new simplified insight into the 2019 coronavirus disease (COVID -19) using the Pareto principle to highlight the main contributors to morbidity and mortality. A time series database of confirmed cumulative cases and deaths for all countries was processed from the Humanitarian Data Exchange website provided by the United Nations Office for the Coordination of Humanitarian Affairs. More than 85% of the incidents recorded worldwide were from the AMRO, EURO, and SEARO WHO regions, and the United States, Russia, and India were found to account for the largest proportion of cases and deaths in these affected areas. The application of Pareto analysis is useful in finding focus groups for further study and modeling.

Keywords: COVID-19, Morbidity, Mortality, Pareto principle, WHO

Introduction

The new coronavirus disease (COVID -19), which has been causing a worldwide pandemic for about two years, has affected every aspect of human life, and the world has changed significantly since then [1]. A large amount of data has been compiled showing morbidity and mortality rates as cumulative numbers of cases and deaths on a daily basis for each affected country and/or area [2]. The database was processed based on the World Health Organization classification (WHO) and then subdivided according to the daily records of the political region.

Implementation of the Pareto principle on big data of COVID-19 morbidities and mortalities

Before assessing the distribution of COVID -19 deaths and morbidities, a brief look at the distribution of the world population based on the census should be taken to find a match with the reported cases and deaths according to the 60/40 and 80/20 rule using Minitab version 17.0.1. This software was implemented as it has been used in other research. [3-6]. The Asian population accounts for about 60% of the world population, followed by Africa and Europe (Figure 1A ) with a cumulative total of over 85% of the world population. However, the COVID -19 cases and deaths show a different descending order, with the American regional office (AMRO) zone ranking first, followed by the European regional office (EURO) and the Southeastern regional office (SEARO) with 0.40, 0.32, and 0.17 cases and 0.48, 0.30, and 0.12 deaths, respectively, as a proportion of the world population (Figure 1B and 1C ).

Contribution of countries by cumulative cases and deaths in WHO regions

An examination of contributing countries and/or areas with cumulative COVID -19 mortality and morbidity rates based on the WHO classification revealed that few regions are particularly affected. The main contributing countries in the AMRO region were, in descending order of cases, the United States of America (USA), Brazil, Argentina, and Colombia, with a cumulative share of more than 80%. For deaths, the same order was maintained for the first two countries, followed by Mexico and Peru with a cumulative contribution of about 82% [7].
Both the USA and Brazil had a cumulative contribution of more than 50% of the total number of people affected in the AMRO region, as shown in Figure 2A and 2B. Another area WHO - which made a relatively small contribution - was studied. The West Pacific Regional Office (WPRO) showed that the Philippines accounted for about one-third of the cumulative daily morbidity and mortality rates. However, the order varied for the other nations that followed, as Japan and Malaysia showed an alternating order in cases and deaths. Interestingly, the most populous country in the world, from which the first epidemic cases originated that triggered the global outbreak, contributed only marginally (≈ 11% in deaths and fourth in cumulative rate) because of rapid and rigorous government and public health interventions [8]. A detailed visualization of the sequence of records is shown in Fig. 3 Figure 2C and 2D.

Figure 1. Pareto chart showing major continents in population census (A) and WHO main contributors in morbidities (B) and mortalities (C).

A look at a third important region, SEARO, showed that India had the highest number of deaths and morbidities at daily rates in the WHO region. This is evident in Figure 2E and 2F with a proportion of over 70%. Adding Indonesia and India to the cumulative deaths gives a proportion of about 0.90 from the SEARO region. Another marginally affected WHO area, the Eastern Mediterranean Regional Office (EMRO), showed a different pattern, where only the first country was strongly affected by COVID-19, followed by a group of seven and six countries showing relatively small variations among them. For the cumulative case rate, the descending order was the Islamic Republic of Iran, Iraq, Pakistan, Morocco, Jordan, the United Arab Emirates (UAE), Saudi Arabia, and Lebanon, with contribution factors of approximately 0.32, 0.13, 0.09, 0.06, 0.06, 0.05, 0.05, and 0.04, respectively (Figure 3). For the cumulative mortality rate, the descending order was Iran, Pakistan, Iraq, Egypt, Tunisia, Morocco, and Saudi Arabia with percentage contributions of 42.1, 9.7, 8.8, 7.1, 6.7, 4.8, and 3.9%, respectively (Figure 3A and 3B). Data from the African Regional Office (AFRO) - which is also a minority - show similar behavior to EMRO, but with a greater contrast between the first country affected and the other nations that follow, as shown in Figure 3C and 3D. Moreover, the reported cumulative cases show a broader range of joint contributions. South Africa alone accounted for approximately 50% and 60% of the total cumulative cases and deaths, respectively. In terms of morbidity and mortality for the remaining countries, the order was as follows: Ethiopia (6.0%), Kenya (4.2%), Nigeria (4.1%), Algeria (3.0%), Zambia (3.6%), Ghana (2.5%), Mozambique (2.1%), Botswana...
(2.1%), Namibia (1.8%) and Zimbabwe (1.8%) and Algeria (4.4%)
Ethiopia (4.0%), Kenya (3.3%), Zimbabwe (2.4%), Nigeria
(2.3%) and Zambia (2.1%), respectively. EURO Lands showed
a unique pattern in **Figure 3E and 3F**. While this region was
considered one of the three main zones affected by daily death
and disease rates from COVID-19, much less variation was
shown with broader contributing countries, resulting in an ob-
servable low declining ramp in the Pareto diagram. Thus, 15
and 12 nations shared the top contributing countries (80%) in
morbidity and mortality data, respectively. The first two coun-
tries (the Russian Federation and the United Kingdom (UK)) ac-
counted for only about one-fifth and one-quarter, respectively,
of the cumulative cases and deaths in the data set. The descend-
ning list of remaining countries in the third and fourth morbid-
ity categories included France (with about the same contribution
as the United Kingdom), Turkey, Spain, Italy, Germany, Poland,
Ukraine, the Netherlands, the Czech Republic, Belgium, and Ro-
mania, with contribution ratios of 0.10, 0.09, 0.07, 0.07, 0.06,
0.04, 0.04, 0.03, 0.03, and 0.02, respectively. The descending
order of the remaining nations for the fourth three cumulative
deaths was as follows: Italy, France, Spain, Germany, Poland,
Turkey, Ukraine, and Romania, with proportions of 0.11, 0.09,
0.08, 0.07, 0.05, 0.04, 0.04, and 0.03, respectively.

**Figure 3.** Pareto chart showing major countries in mortalities and morbidities
from COVID-19 from EMRO (A and B), AFRO (C and D), EURO (E and F).

**Findings and limitations**

The previously analyzed report is time-bound, as updating
the dataset with new inputs may show some discrepancies [9,
10]. In particular, the distribution of COVID-19 across the globe
seemed to be independent of the climatic nature of the regions,
as seen in **Figure 4**, and showed a wide geographic dispersion
within different temperature and humidity ranges. Nevertheless,
**Table 1** showed that countries near the equator had a higher pro-
portion of daily cumulative morbidity and mortality rates than
those farther away, where AMRO and SEARO countries had
about 45% and 43% of cumulative morbidity and mortality rates,
respectively.

![Figure 4.](image)

**World Major Contributors of COVID-19 Mortalities and Morbidities Based on
WHO Regions**

Since other researchers have concluded that hot and humid
climates can halt the spread of the disease to some degree, this
suggests that other factors played a greater role in these polit-
ica regions, including but not limited to public health policy
and agency actions, public perception and awareness of outbreak
control measures, general health status of citizens, racial factors,
extent and rate of vaccination, and population size and density,
in addition to political and economic conditions [11-14]. An-
other important factor that should not be underestimated con-
cerns the government and public health officials, as well as the
routine system for recording, monitoring, and tracking affected
individuals, as this could affect the reliability and quality of the
derived databases.
### Table 1. Major countries in the global distribution of morbidities and mortalities cases by WHO regions.

<table>
<thead>
<tr>
<th>WHO Region</th>
<th>Regional Cumulative * Rate Contribution (%)</th>
<th>Morbidity Country</th>
<th>Cases Contribution (%) *</th>
<th>Mortality Country</th>
<th>Death Contribution (%) *</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMRO B</td>
<td>Cases: 40.1 Deaths: 48.2</td>
<td>USA</td>
<td>19.93</td>
<td>USA</td>
<td>16.39</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Brazil</td>
<td>9.38</td>
<td>Brazil</td>
<td>11.86</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Argentina</td>
<td>2.13</td>
<td>Mexico</td>
<td>6.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mexico</td>
<td>6.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEARO C</td>
<td>Cases: 16.8 Deaths: 11.6</td>
<td>India</td>
<td>13.73</td>
<td>India</td>
<td>8.41</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EURO D</td>
<td>Cases: 31.8 Deaths: 30.0</td>
<td>Russia</td>
<td>3.50</td>
<td>Russia</td>
<td>3.90</td>
</tr>
<tr>
<td></td>
<td></td>
<td>UK</td>
<td>3.18</td>
<td>UK</td>
<td>3.60</td>
</tr>
<tr>
<td></td>
<td></td>
<td>France</td>
<td>3.18</td>
<td>Italy</td>
<td>3.30</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Turkey</td>
<td>2.86</td>
<td>France</td>
<td>2.70</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Spain</td>
<td>2.23</td>
<td>Spain</td>
<td>2.40</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Italy</td>
<td>2.23</td>
<td>Germany</td>
<td>2.10</td>
</tr>
</tbody>
</table>

A World Health Organization, B American Regional Office, C Southeast Asia Regional Office, D European Regional Office. * Based on cumulative daily record.

### Conclusion

In the world of Big Data, it will be useful to analyze a large amount of information quickly, easily and effectively to derive a useful conclusion. From the cumulative daily cases and deaths in the coronavirus pandemic database, a focus study group could be identified using Pareto analysis. These segregated countries will be subjected to further quantitative study using a modeling approach for morbidity and mortality that could serve as the basis for further outbreak investigations.

### Abbreviations

- **AFRO**: African Regional Office
- **AMRO**: American Regional Office
- **COVID-19**: Coronavirus disease of 2019
- **EMRO**: Eastern Mediterranean Region
- **EURO**: European Regional Office
- **SEARO**: Southeast Asia Regional Office
- **WHO**: World Health Organization
- **WPRO**: West Pacific Regional Office

### References