



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

# Data science: a survey on the statistical analysis of the latest outbreak of the 2019 pandemic novel coronavirus disease (COVID-19) using ANOVA

R. Mohammed Harun Babu<sup>1</sup>, M. Shebana<sup>1</sup>,  
R. Mohammed Harish<sup>2</sup>, V. Kanimozhi<sup>1</sup>, K. Arun Kumar<sup>1</sup>

<sup>1</sup>INURTURE EDUCATION SOLUTIONS, BANGALORE, KARNATAKA, INDIA; <sup>2</sup>KARUNYA INSTITUTE OF TECHNOLOGY AND SCIENCES, COIMBATORE, TAMIL NADU, INDIA

## 1. Introduction

A submicroscopic infectious agent that kills millions of people is called a virus. It can replicate only inside the cells of living organisms. There are 320,000 different types of viruses that can infect mammals alone. Some deadliest viruses can kill more than 15 million people annually [1]. There are some most dangerous viral diseases that cause a rapid death rate [2]. They are rotavirus, smallpox virus, measles, dengue, and Spanish flu. Rotavirus, also known as the child killer, is an infectious virus that causes the death of more than 1.5 million children every year. It causes diarrheal illness and damage in lungs among kids of age less than 9 years [3]. Smallpox virus is one of the famous and deadliest viruses that have killed more humans than any other viruses. During the 20th century, the spread of this virus was rapid and it infected an average of 300–500 million people worldwide with a mortality rate of 90%. But after a great research, a medicine for this viral infection was discovered and it has been eradicated worldwide now [4].

Measles is responsible for the deaths of 200 million people worldwide in the past 150 years. More than 197,000 people die due to measles every year. It causes rashes and wounds throughout the body [5]. Dengue is one of the dangerous viral diseases that most of the people are aware of. The dengue virus is spread via mosquitoes. This disease has infected around 100 million people worldwide and has killed more than 20,000 people [6]. Influenza is responsible for 500,000 deaths worldwide [7,8]. The most deadly

pandemic disease is Spanish flu that killed 10–100 million people. The transmission rate of this virus is comparatively low but the mortality rate is too high [9]. In the 21st century, the viral disease that affected the human species is coronavirus disease 2019 (COVID-19) [10]. The causative agent of this disease is severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Even though it has a less mortality rate, the number of people infected by this virus is quite high [11].

Coronavirus is not a single virus, it is a viral family, and this is not the first time that these viruses are attacking human beings [12]. It started its attack in the late 1960s. There are four main categories in the coronavirus family: alpha, beta, gamma, and delta [13]. Scientists classified the human-infecting coronaviruses into seven types: 229E (alpha-coronavirus), NL63 (alphacoronavirus), OC43 (betacoronavirus), HKU1 (betacoronavirus), Middle East respiratory syndrome coronavirus (MERS-CoV, betacoronavirus), SARS-CoV (betacoronavirus), and SARS-CoV-2 (novel coronavirus) [14].

Human coronavirus 229E (HCoV-229E) is a coronavirus genus that infects humans and bats. This species is included in the class Alphacoronavirus and the subclass Duvinaco virus. It is one of the first identified coronavirus strains. It is associated with common symptoms of cold in healthy adults [15]. Human coronavirus NL63 (HCoV-NL63) is one of the four human coronavirus prototypes in circulation. HCoV-NL63 infection causes inner and outer respiratory tract disease and is widespread worldwide, particularly among children under the age of 6 years. The production of lower respiratory disease is a dangerous case to young children and elderly people [16].

Human coronavirus OC43 (HCoV-OC43) is a member of the Betacoronavirus family and the subclass Embecovirus, which infects humans and bovines. It has an additional shorter spikelike protein layer called Hem agglutinin esterase and this is a common similarity among all coronaviruses of the beta family [17]. The coronavirus HKU1 is included with community-acquired pneumonia. It was first identified in 2005, has prevalence worldwide, and is found in both children and adults with upper and lower respiratory tract diseases. The abovementioned four coronaviruses have caused only mild effects and infections in human beings. But this statistic does not last long, as in 2002 a new coronavirus was identified and it was named as SARS-CoV [18].

SARS refers to severe acute respiratory syndrome [19]. This is the first coronavirus that started to show severe infection in human beings [3]. The disease spread to more than two dozen countries including North America, South America, Europe, and Asia. The disease infected around 8000 people and killed around 750–800 people worldwide. On July 2003, the WHO [20] informed that the SARS virus was contained worldwide [15,21]. After 10 years of the SARS outbreak, another new coronavirus was identified in 2012 and named as MERS-CoV [22]. MERS refers to Middle East respiratory syndrome, which was first identified in Saudi Arabia [22]. This virus has symptoms like fever, cough, and shortness of breath [22]. When compared with these two viruses, SARS-CoV-2 has the highest mortality rate. The history and transmission phase of the virus has been explained in this chapter. This work is organized as follows. [Section 2](#) is a background of the chapter, which consists of five subcategories. [Section 2.1](#) explains the Evolution of

diseases from animals and their spread, [Section 2.2](#) explains the transformation of COVID-19 from epidemic to pandemic, [Section 2.3](#) explains the Transmission phase of the virus, [Section 2.4](#) explains the precautions need to be taken in COVID-19, and [Section 2.5](#) consists of the statistical analysis, a kick-start to data science. [Section 3](#) consists of the overview of the dataset. [Section 4](#) represents the clear calculations of statistical analysis and it consists of three subcategories: [Section 4.1](#) represents the two-way analysis of first 60 days, [Section 4.2](#) represents the variation analysis of March 2020, and [Section 4.3](#) represents the one-way analysis of April 2020. [Sections 5 and 6](#) explain the latest outbreak of COVID-19 of March 2020 and April 2020, respectively. [Section 7](#) describes the comparison of COVID-19 between March 2020 and April 2020. Finally, [Section 8](#) consists of the Conclusion of the chapter.

## 2. Background

This segment discusses the basics of zoonotic diseases and their spread, from the COVID-19 epidemic to pandemic and statistical analysis. The main motivation of this section is to get an idea of how statistical analysis is helpful in understanding the spread of COVID-19 [20].

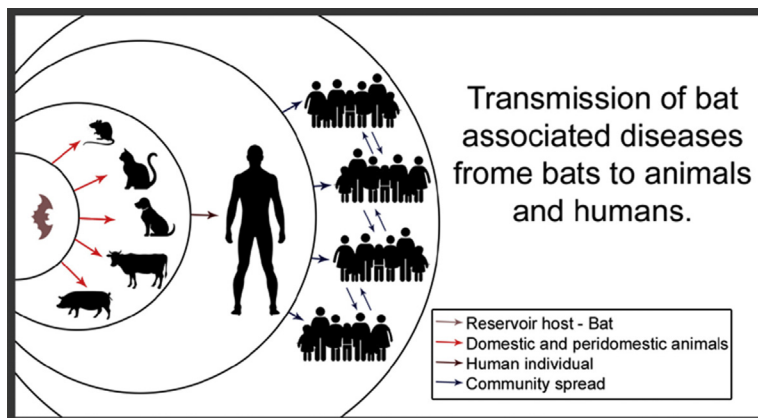
### 2.1 Evolution of diseases from animals and their spread

As the civilization of man develops, the number of diseases and health issues are also getting high. A large number of people die because of many diseases. Not only the small countries but also well-developed countries do not know the strategy to prevent themselves from the diseases. In the world death survey of 2018, around 71% of people died due to diseases caused in the human body but the death rate of natural cause is only 26.7% [13]. In this 70% of diseases, most people die due to heart-related diseases. Infections can affect all age groups but mostly they affect all children and elders [23]. Generally, a disease is an abnormal condition affecting a living organism. There are three main categories in the spread of diseases in the human body: localized diseases, which affect the individual parts of the body; disseminated diseases, which propagate to other areas of the body; and systematic diseases, which spread to the whole body [22]. A disease can occur in people in many forms. The major categories are autoimmune, viral, blood, lung, intestine, skin, nerve or neurologic, sexually transmitted, or thyroid disease. Diseases may be either transmissible or nontransmissible. The diseases that spread to other people while communicating with others are the transmissible ones, and the diseases that cannot be spread while communicating are called nontransmissible [24]. External diseases can be caused by viruses or bacteria, and internal diseases can be caused by autoimmune or hereditary causes. The communicable diseases are further classified into three categories on the basis of modes of transmission: anthroponoses, which can be transmitted from humans to humans; zoonoses, which can be transmitted from animals to humans; and sapronoses, which can be transmitted from the abiotic

resources to humans. In these three modes, the diseases that spread to humans easily and rapidly are zoonotic diseases. They can be caused by germs that spread between animals and humans [25].

Animals are most beautiful creatures made by God. Nowadays, animals have also become a part of the family. People's love and care toward the animal species are drastically high. More number of institutions and clubs are running for the protection and development of the animal species. Animals provide many benefits to humans by being a part of the house. Millions of people have one or more pets in their houses. In spite of having all these good characteristics, animals can sometimes carry harmful germs in their body and spread them to humans and cause illness. These are known as zoonotic diseases. It can be caused by viruses, fungi, bacteria, and parasites. There are two possibilities in the transmission of zoonoses. First one is the animal remains healthier even when it carries the infection and spreads it to humans where humans can only be affected, and the second one is both animals and humans are affected due to the disease [26]. There are many animals and insects that can transmit diseases to humans. The major diseases transmitted by animals are animal flu, bird flu, anthrax, bovine tuberculosis, brucellosis, cat-scratch fever, dengue, Ebola, glanders, etc. There are many animals that cause these diseases, which can be easily transmitted to humans. In most of the diseases, the carrier of the disease is bats [27]. Bats are the only flying mammals on earth. Their genetic name is Chiroptera, which comes under the class Mammalia in the kingdom Animalia. Bats are more maneuverable than other birds, which have a high capability of sustaining light. Their forelimbs are modified into large wings that are made up by a thin membrane called patagium. There are more than 1200 species of bats worldwide. This creature is very important in the ecosystem because it is significant for pollinating flowers and eating small insects [28].

Even though it has a lot of good characteristics, it is a dangerous carrier of viruses. Bats can carry a total of 61 zoonotic viruses and 137 host viruses. While analyzing the death rate of zoonotic diseases, most of them are caused by bats [26]. The complete cycle of how a disease spreads from bats to animals and humans is explained in Fig. 7.1.



**FIGURE 7.1** Transmission of bat-associated diseases from bats to animals.

## 2.2 COVID-19 epidemic to pandemic

Pandemic is not a new term for humanity. Throughout history, mankind has suffered a lot of infectious diseases starting from the Plague of Justinian up to today's COVID-19 [29]. Humans have adapted themselves to infectious diseases and outbreaks, but sometimes they meet up with some novel infections for which they have no solution. The development in overseas trading has helped the novel diseases to spread far and wide, creating a global pandemic [30].

The novel coronavirus that originated in Wuhan, China, in late December 2019 has taken a pandemic status, affecting over 170 countries around the world in less than 3 months [29]. This disease is now referred to as COVID-19, and its causative virus is called severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). It has a lot of similarities with the other two coronaviruses, the MERS-CoV and SARS-CoV [31,32].

The outbreak of COVID-19 started during early December 2019 in Wuhan, a port city in central Hubei province, China. Many adults were admitted to hospitals with severe pneumonia of unknown cause [8,33]. Respiratory samples were taken from patients and sent to laboratories for testing [7]. On December 31, 2019, China informs the WHO about 41 patients who were suspected of unusual pneumonia. Most of the patients are connected to the Hunan [34] seafood wholesale market [35]. On January 1, 2020, China closes the wholesale market. On January 5, 2020, Chinese authorities informed that they have identified a new virus and named it as the novel coronavirus (nCoV). Samples taken from the Hunan seafood market were tested and results were positive, confirming the virus originated from there [22,36].

The New Year 2020 and the Chinese New Year vacation paved a way for the rapid spread of this virus. The number of cases was increasing tremendously, and human-to-human transmission was also confirmed as some cases that are reported had no connection with Wuhan, the place of origin of the virus. China reported its first death (a 61-year-old man) due to the novel coronavirus on January 11, 2020. Cases were reported in other countries, which include Thailand, Japan, and South Korea, most of which are people who returned from Wuhan. The first case outside of China was reported in Thailand [31].

On January 23, 2020, Wuhan was placed under quarantine followed by other cities in the Hubei province. Other countries also started to screen passengers at the airport, especially those who returned from China. Passengers were tested for the novel coronavirus, and if they tested positive, they were isolated and treated by the health authorities. On February 2, 2020, first death outside of China was reported in the Philippines. On January 30, 2020, a public health emergency of international concern was officially declared by the WHO [37].

On February 5, 2020, the Diamond Princess Cruise ship in Yokohama, Japan, was quarantined and all passengers were tested for COVID-19 by the officials. During the SARS epidemic, it took 6 months to exceed 5000 cases in mainland China, but the novel coronavirus just took 1 month. On February 9, 2020, the death toll in China surpassed

that of the 2002–03 SARS epidemic, with 811 deaths recorded. On February 11, 2020, the WHO announced that the new coronavirus infection would be called as COVID-19.

Outside China, Italy was the most affected country. On March 8, 2020, Italy places all its 60 million residents on lock down, as the spread of the virus was uncontrollable in the country. On March 11, the WHO declares the outbreak as a pandemic. As of March 19, 2020, globally 240,000 cases were reported, out of which 145,000 were active cases and roughly 85,000 recoveries and 10,000 deaths. While the number of cases was increasing at an alarming rate in other parts of the world, China was slowly recovering from the impact of COVID-19 [38].

As of March 19, 2020, China has reported no new local cases of the coronavirus for the first time since the pandemic began [39]. The spread of disease in the country has slowed down dramatically. More than 65,000 people have recovered from the disease. But still, other countries especially Italy, Spain, Iran, the United States, and a few other countries are struggling to control the spread of this virus [21].

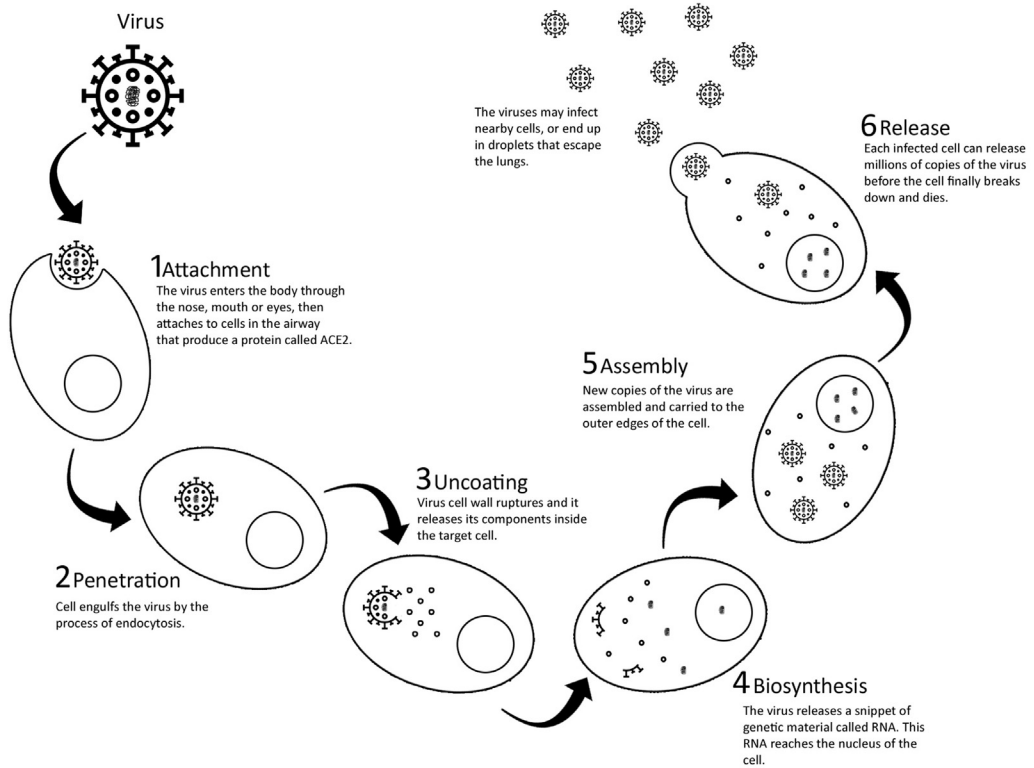
### 2.3 Transmission phase

It is important to understand how coronavirus affects or enters into a human cell, which will contribute to the development of a vaccine or drug to control this contagious virus [28]. The coronavirus got its name from the crownlike spikes that protrude from its surface [18]. These spikes or spike protein or S-protein acts as a key and attaches itself to the surface of respiratory cells [34]. The virus uses angiotensin-converting enzyme 2 (ACE2) receptors to enter into the target cells [40]. Once the virus enters the target cell, it releases its genetic material, RNA. The infection develops and the cell organelles begin to form new spikes and proteins that will form new copies of the coronavirus [8]. Each infected cell can produce millions of cells before they rupture and die. These new viruses that are produced will spill out and affect other cells or settle in droplets that escape the lungs [41]. COVID-19 causes fever as the immune system tries to fight against the virus. In the majority of cases, the immune system overreacts and starts to attack the lung cells [4]. Due to this, lungs will be filled with fluid and dead cells [19]. This causes difficulty in breathing and can lead to respiratory diseases and ultimately death. The complete transmission phase of the virus is shown in Fig. 7.2 [22,26].

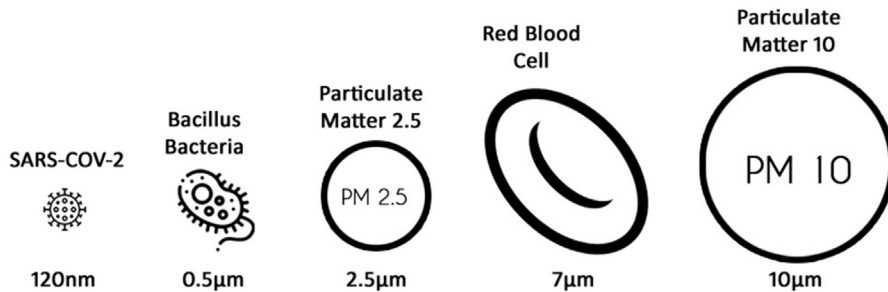
### 2.4 Precautions against COVID-19

Currently, no vaccine is available for COVID-19. So the best cure available is to avoid being exposed to the virus. This virus spreads from people to people who are in close contact with each other [41]. It spreads in the form of droplets when an infected person sneezes or coughs. A single cough can produce up to 3000 droplets. These particles can settle on other people, their clothes, or any other surface around them [10]. The SARS-CoV-2 virus is so small compared with other particles, with its size ranging from 100 to 120 nm, as shown in Fig. 7.3. As these viruses are smaller in size, they can remain in





**FIGURE 7.2** Replication and spread of coronavirus after entering the human cell. ACE2, angiotensin-converting enzyme 2.



**FIGURE 7.3** Comparing coronavirus with other particles based on their size. SARS-CoV-2, severe acute respiratory syndrome coronavirus 2.

the air for quite a long time [33]. These viruses can also remain intact for a longer time in fecal matter, so it is important to wash hands after visiting the toilet [42].

SARS-CoV-2 has a genetic material RNA surrounded by a lipid membrane with crownlike spikes protruding from its surface [18]. The WHO and other health authorities



are emphasizing frequent handwashing to prevent the spread of COVID-19. The reason for this is soap is effective in breaking the lipid membrane of the virus and eventually it kills the virus. One should avoid touching their eyes, nose, and mouth with unwashed hands. Other precautions include avoiding contact with sick people, covering the mouth and nose while sneezing and coughing, throwing used tissues in a closed dustbin, wearing face mask if a person is sick or when he/she visits a sick person, and cleaning the surfaces with disinfectants [33]. As the spread of COVID-19 is rapid the best preventive measure is to stay at home and avoid infection.

## 2.5 Statistical analysis—kick-start to data science

Everyone knows the serious issue that is going on all over the world and that is COVID-19. Many scientists, researchers, and government officers were struggling to find a solution to control the spread of the virus. Although control measures are taken, researchers are still working on finding a vaccine for the disease. The rate of spread of COVID-19 is rapid and comparatively higher than the previous coronavirus diseases. The only way to be protected from the virus is social distancing; doctors and medical professionals confirm that social distancing is mandatory to protect oneself from the disease. To control the spread of the virus, one must know the level of vulnerability of the disease. For this process, two components are required: past analysis and future prediction of the virus. With the help of these two components, the number of people got affected in the past and the number of people that are going to be affected in the future can be easily predicted [43]. For the past analysis of the disease, many statistical tools are used to calculate the effectiveness of the disease. In this analysis, analysis of variance (ANOVA) is used to calculate the result. ANOVA is a statistical tool that is used to calculate the significant relationship among various groups. Here, more than 194 countries are involved in the analysis of COVID-19. From this analysis, the variance level of the number of cases in those countries can be calculated day by day. With this result, the effect of disease in the future can be easily calculated. For the future prediction of the disease, three main statistical factors have to be identified:

- (1) growth factor,
- (2) mortality rate,
- (3) doubling time.

With the help of these factors, the number of people that are going to be affected with COVID-19 in the future can be easily predicted.

Growth factor is the element by which a quantity multiplies itself over time, and growth rate is the factor added by which a quantity increases or decreases over a period. In this process the number of people who are going to be affected by the virus can be calculated with the help of the variance level of already-infected people. The equation to calculate the number of people who will be affected due to COVID-19 is given as

$$P = NGX \tag{7.1}$$

Here,  $P$  denotes the number of people who are going to be affected by the disease,  $N$  represents the number of people affected on the day of prediction,  $G$  represents the growth factor, and  $X$  represents the number of days of prediction [44]. Mortality rate is also known as the death rate, which represents the number of deaths occurred within a given period in a particular population. The mortality rate of coronavirus is 3.8% in China and 3.4% worldwide. This percentage will be used to find the comparison of death cases countrywise. The equation for calculating the mortality rate of COVID-19 is given as

$$M = (D * 10N) / S \quad (7.2)$$

Here,  $M$  denotes the mortality rate,  $D$  represents the deaths occurred in a specific period, and  $S$  denotes the total size of the population [45].

Doubling time is the period in which the number of disease cases reaches twice as the current value. According to the result on April 20, 2020, the doubling rate of COVID-19 in worldwide was 10 days. It means that it takes 10 days to reach twice the number of cases of the current day [46]. The use of this process is to predict the future outcome of the people affected by the coronavirus. It will be very useful for the prevention of the disease.

### 3. Overview of dataset

In this chapter a dataset is prepared based on the reports of the WHO [37,47] and daily update report of the Worldometers [48]. The dataset has been divided into two parts. The first part contains countrywise data and is used for one-way ANOVA analysis, and the second part contains country-day wise data and is used for two-way ANOVA. For the two-way analysis the dataset is prepared for a total of 60 days (January 20, 2020 to March 19, 2020) and the prepared dataset is split into two sets. The first set contains data of 26 countries and its confirmed cases in the first 30 days (January 20, 2020 to February 19, 2020). The second set contains data of 175 countries and its confirmed cases in the next 30 days (February 20, 2020 to March 19, 2020).

For the one-way analysis the datasets are prepared for a total of 100 days (January 20, 2020 to April 25, 2020). This dataset consists the number of daily and total confirmed cases, the number of daily versus total recovered cases, the number of daily versus total death cases, and the number of daily versus total active cases. These things are explained clearly in this dataset. The dataset and details of the whole survey is available on <https://github.com/mhbharun/COVID-19.git>.

### 4. Statistical analysis

In this chapter, statistical analysis is performed in two ways. In the first section, two-way analysis is performed with two trial of experiments. For the first trial of experiment, first 30 days' data was used and for the second trial of experiment the next 30 days' data was used. In the first section, cases till March 19, 2020, have been covered, and in the second section, one-way analysis is performed for a total of 100 days, which is till April 25, 2020.

#### 4.1 Two-way analysis: January 20, 2020 to March 19, 2020

Mathematic models are developed to describe a process, its effective usage, and also to study the properties of materials and products [40]. Different statistical methods are used for precise identification and development of processes [49]. Since the outbreak of COVID-19 in Wuhan, China, the spread of disease was rapid and it traveled to various countries in a short span of time. The number of people infected with the disease varied among countries. As per the reports by the WHO, the disease spread around 26 countries in the first 30 days. China, the place of origin of the virus, was the worst affected country in the first 30 days. With just 248 confirmed cases on day 1 the count reached 1,42,823 cases on day 30. Next to China, Singapore, Japan, and Korea were the mostly affected countries. The countries that are affected due to covid-19 in the first 30 days with total number of confirmed cases are shown in Fig. 7.4 [50].

To analyze the spread of disease among various countries, two-way ANOVA is used [51]. The analysis is performed by comparing the number of confirmed cases in each country and their increase in each day. For analysis, at least two factors are required; in this case, days and countries are taken as factors and denoted as A and B, respectively. The mean of confirmed cases in each day and in each country is calculated and denoted as  $X_A$  and  $X_B$ , respectively.

As the first step of ANOVA, the null hypothesis and alternative hypothesis should be declared [52]. Null hypothesis is a statement or a claim that is need to be tested and is denoted as  $H_0$ . For a two-way analysis the possible null hypotheses are (1) there is no difference in means of factor A, (2) there is no difference in means of factor B, and (3) there is no interaction between the factors A and B. For this analysis, the null hypothesis is declared as

$$H_0: X_{A1} = X_{A2} = X_{A3} \dots X_{An} \text{ (means of factor A is equal)} \quad (7.3)$$

$$H_0: X_{B1} = X_{B2} = X_{B3} \dots X_{Bn} \text{ (means of factor B is equal)} \quad (7.4)$$

$$H_0: X_{AB1} = X_{AB2} = X_{AB3} \dots X_{ABn} \text{ (means of both factors A and B are equal)} \quad (7.5)$$

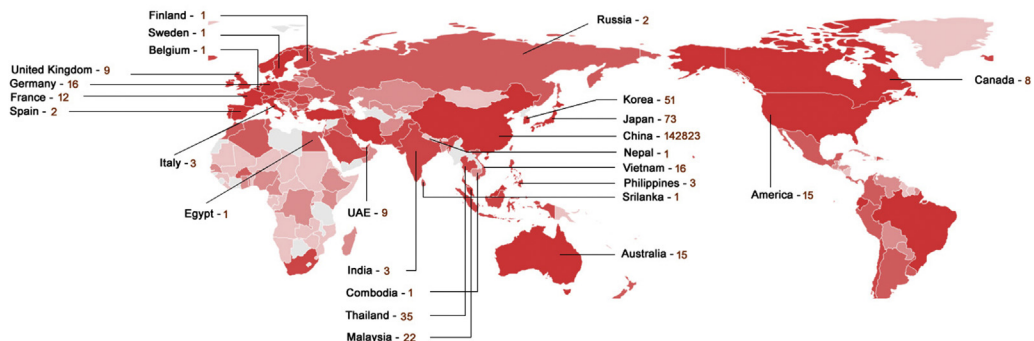


FIGURE 7.4 Distribution of COVID-19 cases in different countries from January 20, 2020 to February 19, 2020.

Alternative hypothesis is a statement that proves the problem statistically where it is the negation of null hypothesis. Alternative hypothesis is denoted as  $H_A$ . For a two-way ANOVA, possible alternative hypothesis are (1) there is a difference in means of factor A, (2) there is a difference in means of factor B, and (3) there is an interaction between the factors A and B. For this analysis, the alternative hypothesis is declared as

$$H_A: X_{A1} \neq X_{A2} \neq X_{A3} \dots X_{An} \text{ (means of factor A is not equal)} \quad (7.6)$$

$$H_0: X_{B1} \neq X_{B2} \neq X_{B3} \dots X_{Bn} \text{ (means of factor B is not equal)} \quad (7.7)$$

$$H_0: X_{AB1} \neq X_{AB2} \neq X_{AB3} \dots X_{ABn} \text{ (means of both factors A and B are unequal)} \quad (7.8)$$

After declaring the null hypothesis and alternative hypothesis, there are four important components that are needed for the calculation of two-way ANOVA: sum of squares ( $SS$ ), mean squares ( $MS$ ), degrees of freedom ( $DF$ ), and  $F$ -value. Sum of squares is used to find the variability in a data [53]. Here, the sum of squares were declared as three types: sum of square for factor A ( $SS_A$ ), sum of square for factor B ( $SS_B$ ), and sum of square for both factors A and B ( $SS_{AB}$ ). The equations of sum of squares for factor A, factor B, and factor AB were represented as

$$SS_A = Nx \ b(X_A X)^2 \text{ (factor A)} \quad (7.9)$$

$$SS_B = Nx \ a(X_B X)^2 \text{ (factor B)} \quad (7.10)$$

$$SS_{AB} = SS_{total} - SS_{error} - SS_B - SS_A \text{ (factor AB)} \quad (7.11)$$

where  $a$  represents the number of values in factor A,  $b$  represents the number of values in factor B, and  $N$  represents the total number of values in a table.

Mean squares are used to find the difference between predicted values and actual values. In this analysis, mean squares were declared in four categories: mean square for factor A ( $MS_A$ ), mean square for factor B ( $MS_B$ ), mean square for both factors A and B ( $MS_{AB}$ ), and mean square of error value ( $MS_{error}$ ). The formulas of mean squares are defined as

$$MS(A) = SS(A) / DF(A) \quad (7.12)$$

$$MS(B) = SS(B) / DF(B) \quad (7.13)$$

$$MS(AB) = SS(AB) / DF(AB) \quad (7.14)$$

$$MS(error) = SS(error) / DF(error) \quad (7.15)$$

Degrees of freedom is used to determine the critical values where the acceptance or a rejection of a hypothesis can be defined. To calculate the two-way ANOVA, degrees of freedom consists of five types: degrees of freedom of factor A, degrees of freedom of factor B, degrees of freedom of both factors A and B, degrees of freedom of error value, and degrees of freedom of total values. These values are calculated using the formulas as

$$DF(A) = a - 1 \quad (7.16)$$

$$DF(B) = b - 1 \quad (7.17)$$

$$DF(AB) = (a - 1)(b - 1) \quad (7.18)$$

$$DF_{\text{error}} = (N - 1) - (a - 1) - (b - 1) - (a - 1)(b - 1) \quad (7.19)$$

$$DF_{\text{total}} = N - 1 \quad (7.20)$$

$F$ -values are used to test whether the values are statistically significant or not.  $F$ -values can be determined with the help of mean squares. These values can also be classified into three categories:  $F$ -value for factor A,  $F$ -value for factor B, and  $F$ -value for both factors A and B. The  $F$ -values can be calculated as

$$F(A) = MS(A)/MS(\text{error}) \quad (7.21)$$

$$F(B) = MS(B)/MS(\text{error}) \quad (7.22)$$

$$F(AB) = MS(AB)/MS(\text{error}) \quad (7.23)$$

From these components, the two-way ANOVA can be calculated by declaring the factors A and B as day and country, respectively. The result of the calculation is shown in [Table 7.1](#).

From [Table 7.1](#), the  $F$ -value is calculated as 14.273. So it is denoted as calculated  $F$ -value.

$$F - \text{value}_{\text{calculated}} = 14.273 \quad (7.24)$$

From the degrees of freedom in [Table 7.1](#), such as 29 and 25, the tabulated  $F$ -value is 1.85. So it is denoted as tabulated  $F$ -value.

$$F - \text{value}_{\text{table}} = 1.85 \quad (7.25)$$

When the calculated  $F$ -value is greater than the tabulated  $F$ -value, the null hypothesis can be rejected. Here the calculated  $F$ -value is 14.273 and the tabulated  $F$ -value is 1.85. Hence, the null hypothesis can be rejected and the alternative hypothesis can be accepted.

**Table 7.1** Analysis of variance result for trial 1 experimentation.

Source	Sum of squares	Degree of freedom	Mean squares	$F$ -value
Day	3,128,450,048	29	107,877,587.8	14.243
Country	5.302E+10	25	2,120,698,490	280.003
Day x Country	7.826E+10	724	108,098,823.6	14.273
Error	7,573,832.000	1	7,573,832.000	
Total	1.358E	780		

RESULT: From the trial 1 experimentation, analysis proves that there is significant relationship between the mean values of the two factors. By declaring the days and countries as two factors, the two-way analysis proves that the mean values of the number of confirmed cases of COVID-19 vary between each country and each day [40].

The spread of the disease was rapid. During the first 30 days of outbreak, it reached only 26 countries, but in the next 30 days, it traveled to almost 175 countries [47]. Trial 2 experiment is performed using the set 2 data that contains a total of 175 countries and the result is shown in Table 7.2.

From Table 7.2, the  $F$ -value is calculated as 117.22. So it is denoted as calculated  $F$ -value.

$$F - \text{value}_{\text{calculated}} = 117.22 \quad (7.26)$$

From the degrees of freedom in Table 7.2, such as 29 and 174, the tabulated  $F$ -value is 1.52. So it is denoted as tabulated  $F$ -value.

$$F - \text{value}_{\text{table}} = 1.52 \quad (7.27)$$

when the calculated  $F$ -value is greater than the tabulated  $F$ -value, the null hypothesis can be rejected. Here the calculated  $F$ -value is 117.22 and the tabulated  $F$ -value is 1.52. Hence, the null hypothesis can be rejected and the alternative hypothesis can be accepted.

RESULT: From the trial 2 experimentation, analysis proves that there is a significant relationship between the mean values of the two factors. By declaring the days and countries as two factors, the two-way analysis proves that the mean values of the number of confirmed cases of COVID-19 vary between each day and each country. When compared to the first month of the outbreak, the level of growth of confirmed cases is drastically high, with a growth percentage of 45%. The number of confirmed cases in the 60th day is eight times more than the number of cases at the 30th day [23]. It affects not only the developed countries such as China, America, and Italy but also small islands such as Cayman Islands and Virgin Islands. At the end of the 60th day, totally 154 countries, 15 territories, and 6 islands were affected by the severe attack of COVID-19 [22,54].

**Table 7.2** Analysis of variance result for trial 2 experimentation.

Source	Sum of squares	Degree of freedom	Mean squares	$F$ -value
Day	353,007,300.5	29	12,172,665.54	1,460,719,864
Country	4.836E+11	175	2,779,570,351	3.335E+11
Day x Country	4.871E+11	4986	9,768,931.382	117.2271766
Error	0.500	60	0.008	
Total	5.368E+11	5250		

## 4.2 Variation analysis: January 20, 2020 to March 19, 2020

The mean distribution and standard deviation [55] of the number of confirmed cases in the first 30 days are shown in Fig. 7.5. The percentage of confirmed cases in the first month is less than 2% in most of the countries including Spain, Italy, and India. While in countries such as Malaysia, Australia, and Germany the percentage of confirmed cases ranged from 8% to 10%. When comparing two countries such as Egypt and Malaysia, the growth level of confirmed cases was 80% higher [8].

Fig. 7.6 shows the distribution of variance level of mean values and standard deviation values of the top five most affected countries in 30 days [22]. At the end of the first month, the countries that are highly affected due to the virus were China, Singapore, Japan, Thailand, and Korea. Except China, the level of confirmed cases in the other four countries were equal [56]. A linear graph was observed for all the four countries and suddenly a peak was observed for China. The number of cases in China reached more than 50,000.

Figs. 7.7–7.9 represent the variation of mean values and standard deviation values of the least affected countries due to the virus in the next 30 days. During the first 30 days of infection, only 26 countries were affected, but in the next 30 days the infection has reached to almost 175 countries. Several places such as Montenegro, Fiji Islands, and Barbados saw a sudden growth in the number of confirmed cases [57]. The countries such as Guam, Guatemala, and Mauritius faced a linear growth of confirmed cases, with a growth rate of 0.2%.

According to Fig. 7.8 the standard deviation attains a linear growth with less amount of variance while the mean values of countries are highly differentiable. Countries such as Venezuela, Jordan, and Turkey reached a sudden growth of confirmed cases, in which Armenia showed a high case percentage of 12%.

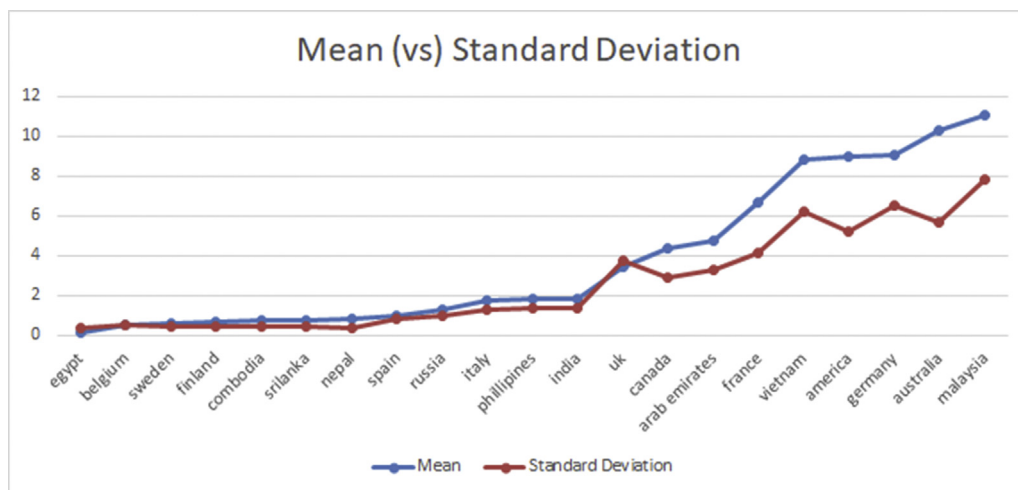


FIGURE 7.5 Mean versus standard deviation graph for the least 21 countries among 26.



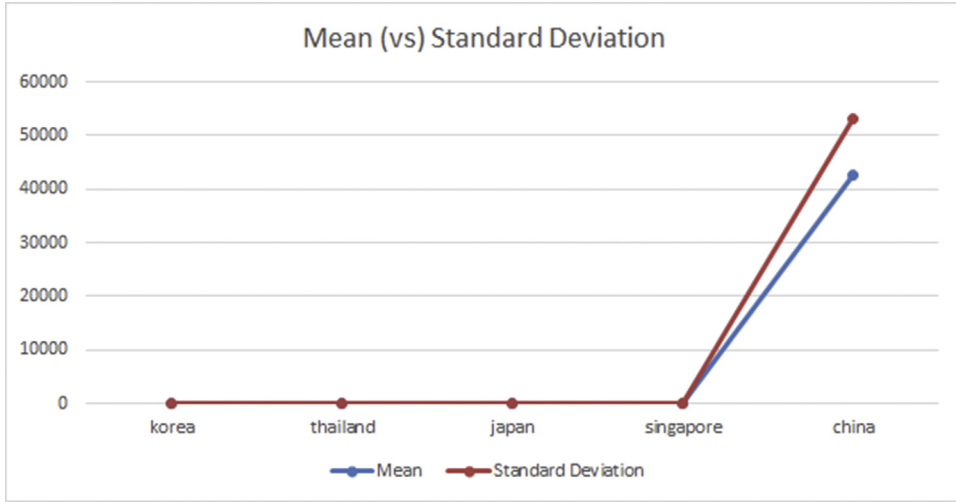


FIGURE 7.6 Mean versus standard deviation graph for the top 5 countries among 26.

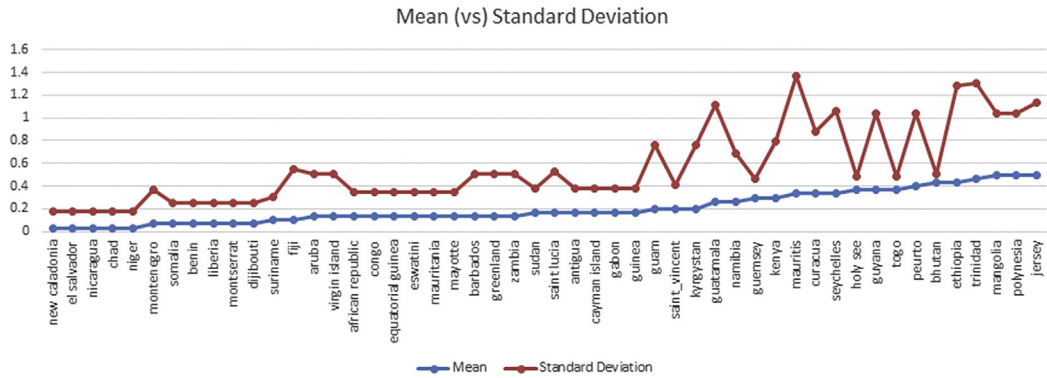


FIGURE 7.7 Mean versus standard deviation graph for the least 50 countries among 175.

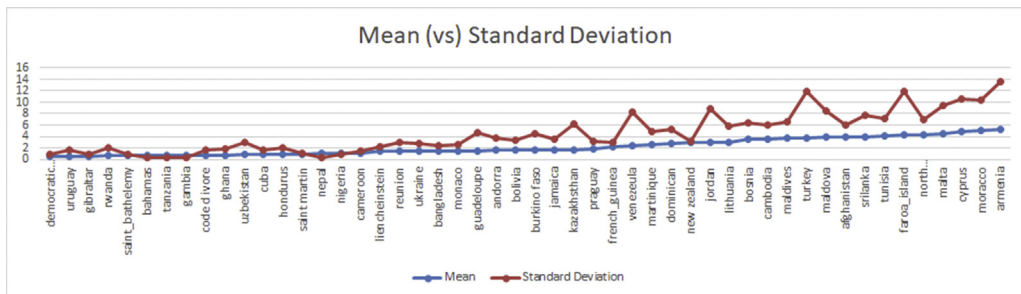


FIGURE 7.8 Mean versus standard deviation graph for the next least 50 countries among 175.



### 4.3 One-way analysis: January 20, 2020 to April 25, 2020

**Table 7.3** represents the summary of total confirmed cases versus total recovered cases. As the table shows, the virus has been spread over more than 175 countries. The total confirmed cases in 175 countries is 54,51,327. It is the pandemic disease that causes more number of infections [58].

**Table 7.4** shows the comparison variation of the number of recovered cases in all countries. With the degrees of freedom 1 and 348, the  $F$ -value calculated is 4.340,913,847. It has been conducted by fixing the categories such as between groups and within groups. From these components, ANOVA of the number of recovered cases has been calculated [59].

**Table 7.5** represents the summary of total confirmed cases versus total deaths. While comparing the total cases, 5,451,327 to the total deaths, 382,173, the level of death rate is only 5%. It shows that this disease causes less number of deaths cases when compared with other previous diseases [60].

**Table 7.6** represents the one-way ANOVA of total confirmed cases and total death cases. With the degrees of freedom 1 and 348, the  $F$ -value calculated was 7.585703874. It has a big variation when compared with the number of recovered cases. This variance shows that the death rates in different countries are highly varied among each other.

From **Tables 7.4 and 7.6**, there is a conclusion that the percentage of variance in total recovered cases is more than the percentage of variance in total death cases. Hence, this calculation proves that many people are getting recovered from the disease and the count of people died due to the disease was comparatively low.

**Table 7.3** Summary of total cases/total recovered.

Groups	Count	Sum	Average	Variance
Total cases	175	5,451,327	31,150.44	19,241,348,449
Total recovered	175	149,825	8561.468571	1,329,385,833

**Table 7.4** Variation comparison of total recovered cases.

Source of variation	SS	df	MS	F	P-value	F crit
Between groups	44,647,892,643	1	44,647,892,643	4.340913847	0.037936	3.868319
Within groups	3.57931E+12	348	10,285,367,141			
Total	3.62396E+12	349				

*df*, degree of freedom; *MS*, mean squares; *SS*, sum of squares.

**Table 7.5** Summary of total cases/total deaths.

Groups	Count	Sum	Average	Variance
Total cases	175	5,451,327	31,150.44	19,241,348,449
Total deaths	175	382,173	2183.845714	115,606,386.5

**Table 7.6** Variation comparison of total deaths.

Source of variation	SS	df	MS	F	P-value	F crit
Between groups	73,418,063,645	1	73,418,063,645	7.585703874	0.006192	3.868319
Within groups	3.36811E+12	348	9,678,477,417			
Total	3.44153E+12	349				

*df*, degree of freedom; *MS*, mean squares; *SS*, sum of squares.

## 5. Outbreak of COVID-19, as of March 31, 2020

In the previous section, statistical analysis is performed till March 20, 2020. The next section explains about the outbreak of COVID-19 and graphs were plotted for data as of March 31, 2020 [37,47]. At the end of 2 months of the outbreak the total number of confirmed cases, active cases, and deaths among various countries has a significant difference. For the purpose of analysis, the top 14 countries based on their total number of confirmed cases were taken into consideration. At the end of the analysis, three graphs were plotted [54].

Fig. 7.11 shows the comparison between the total number of confirmed cases and active cases in each 14 countries. From the graph it is clear that China, the place of origin of the virus, has low number of active cases compared with other countries. The United States tops the list with 67,063 active cases out of 68,489 total confirmed cases. Next comes Italy with a total of 57,521 active cases followed by Spain with 40,501 active cases; Germany, France, Iran, and other countries follow respectively.

Fig. 7.12 shows the comparison between the total number of confirmed cases and recovered cases. China tops this list with 74,051 recoveries out of its 81,285 confirmed cases. Except China, no other country has a good amount of recovered cases; the United States has just 394 recovered cases out of its 68,489 confirmed cases. Italy and Spain also have a relatively low number of recovered cases compared with their total number of confirmed cases.

Fig. 7.13 explains the comparison between the total number of confirmed cases and deaths. Italy has a high death rate of 7503 people, which is twice that of China. Spain has a death toll of 3647, which is higher than China's death toll of 3287. Iran, France, and the United States come next in the list. The United States with the highest number of active cases has a death toll of 1032. Comparing all the three graphs, it is clear that China has

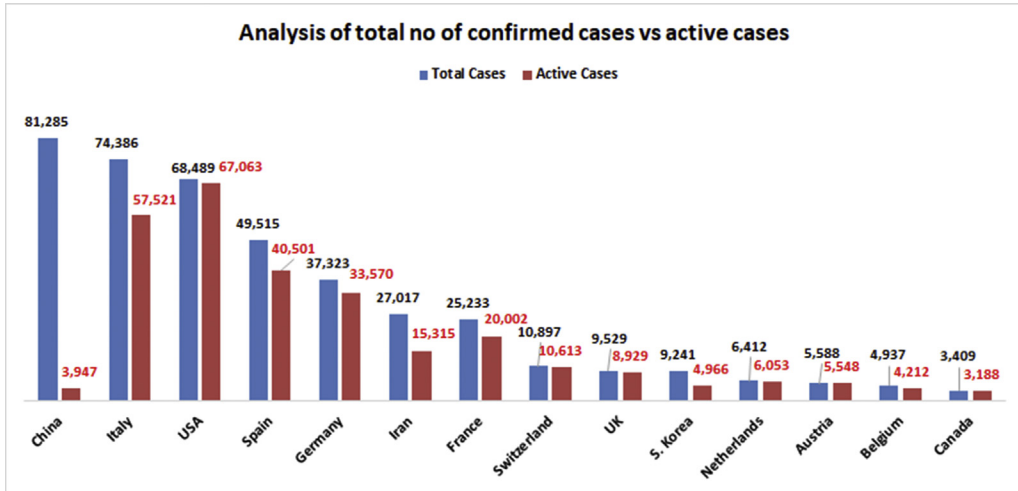


FIGURE 7.11 The total number of confirmed cases versus active cases, as of March 26, 2020.

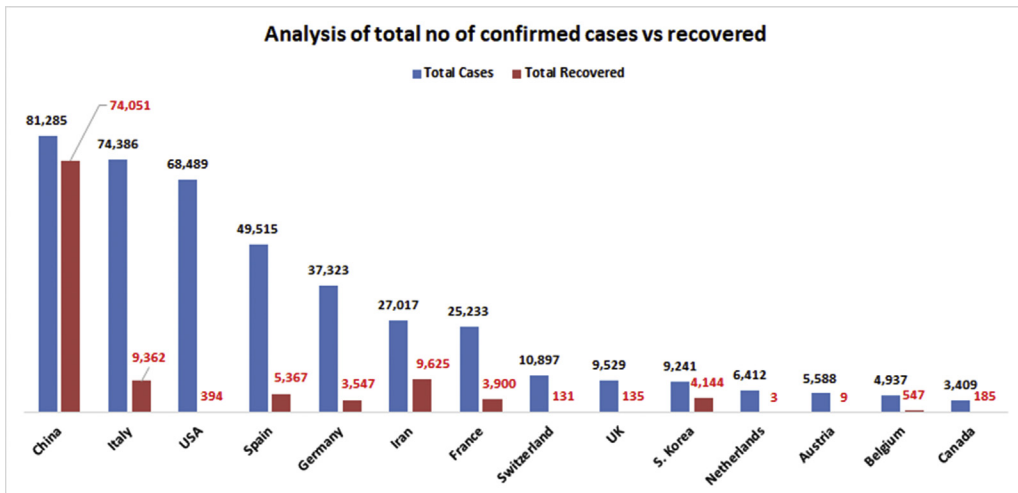


FIGURE 7.12 The total number of confirmed cases versus recovered cases, as of March 26, 2020.

almost recovered from the effect of COVID-19. It has comparatively less number of active cases and most of the affected people have already recovered. But the rest of the world is still suffering, especially Italy, Spain, and the United States. Italy has become the epicenter of the pandemic, with the death toll almost double that of China. The number of cases is increasing rapidly in Spain, the United States, and France.

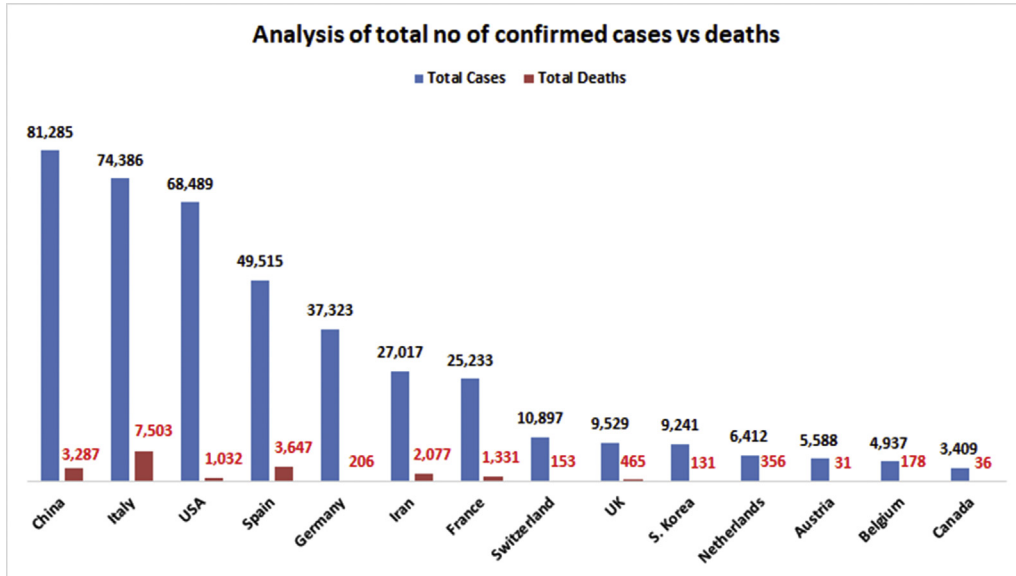


FIGURE 7.13 The total number of confirmed cases versus deaths, as of March 26, 2020.

## 6. Outbreak of COVID-19, as of April 25, 2020

As the spreading of COVID-19 is growing rapidly. Many people are affected by the disease in many countries. Here, the total analysis of people who are affected due to the coronavirus infection is represented. At the end of the 100th day of novel coronavirus outbreak, it has spread into more than 200 countries and territories. Fig. 7.14 represents the active cases of COVID-19 worldwide. This graph consists of three parts: number of cases more than 1 lakh, between 50,000 and 1 lakh, and less than 50,000. Here, countries such as the United States, Spain, Italy, France, Germany, and the United Kingdom had the highest number of active cases. There are more than 7 lakh people affected due to the novel coronavirus in America. Even though China was the first one to be affected by the disease, America was country with the topmost number of confirmed COVID-19 cases. The countries that had the number of infected people of more than 100,000 were America, Spain, Italy, France, Germany, and the United Kingdom. Austria and South Korea are the countries that had less than 5000 cases. The number of confirmed cases in South Korea, Chile, and Saudi Arabia is comparatively equal. This graph clearly explains that the number of active cases in all countries is somewhat low when compared with the total number of confirmed cases.

The percentage of increase in the pandemic disease spread is too fast, and when compared with the number of death cases, the number of recovered cases is more. Fig. 7.15 represents the total recovered cases of COVID-19 worldwide. When compared with other countries, China attained the highest number of recovered cases. Mostly

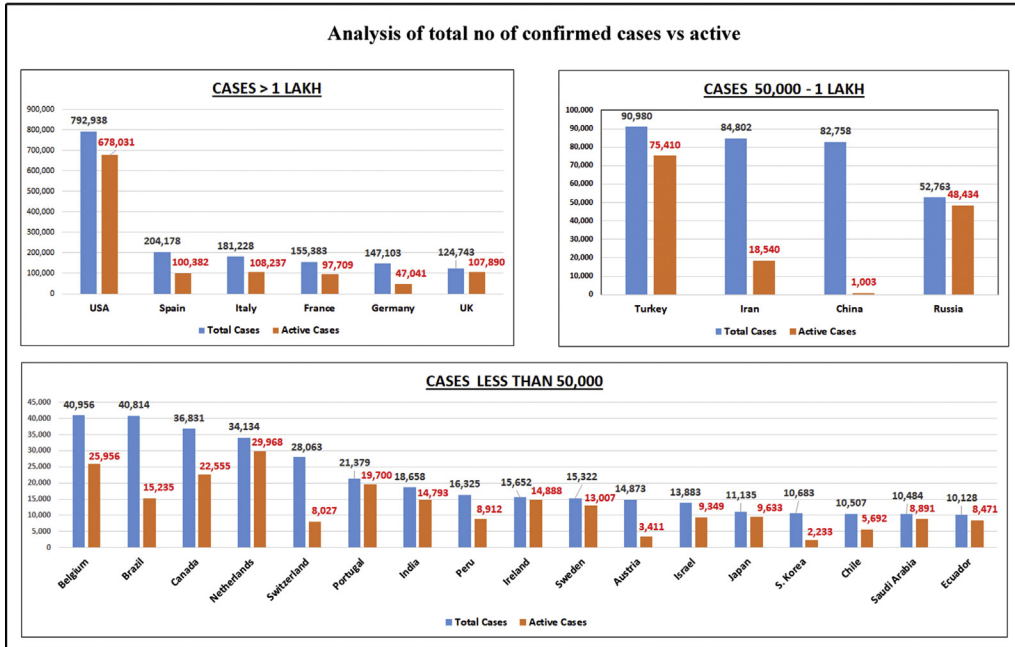


FIGURE 7.14 The total number of confirmed cases versus active cases, as of April 25, 2020.

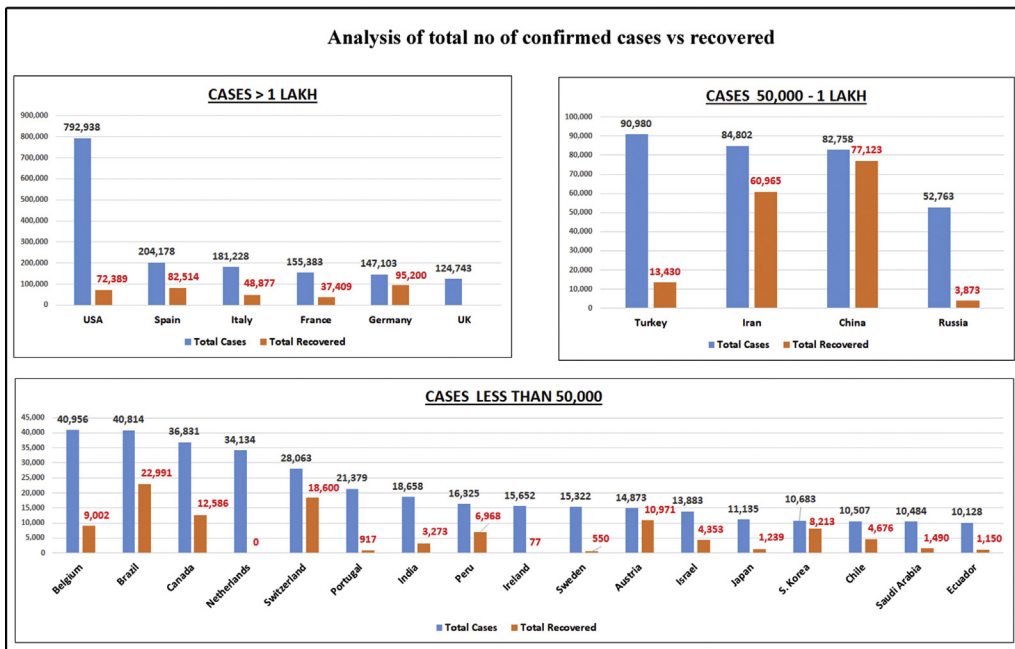
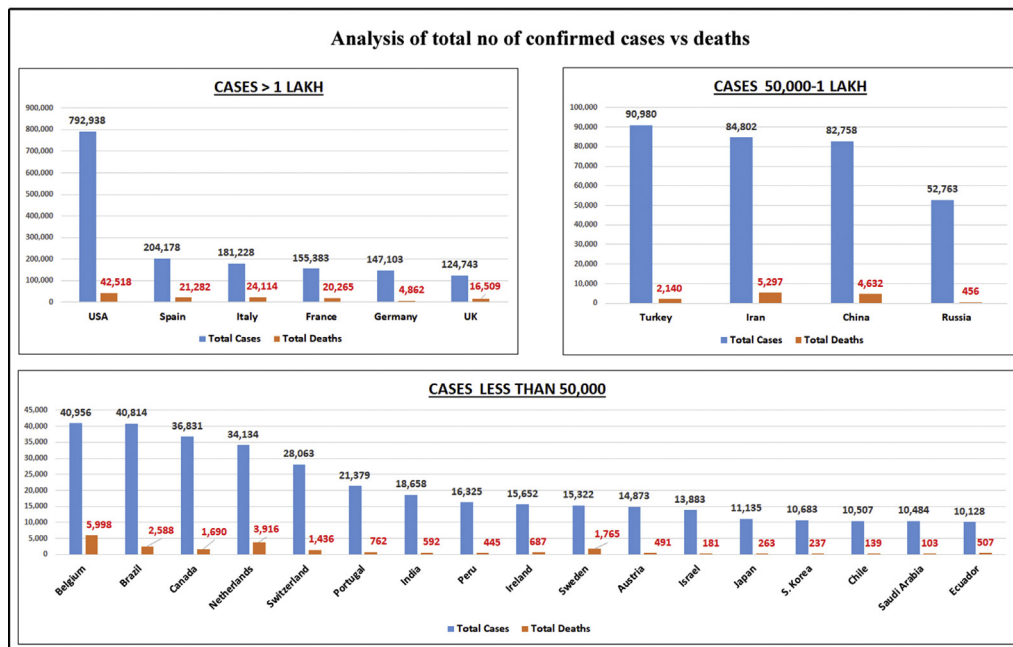


FIGURE 7.15 The total number of confirmed cases versus recovered cases, as of April 25, 2020.



two-thirds of people in China have recovered from the disease. The graph clearly describes that the number of recovered cases in the United States is very low and denotes that America was in a dangerous situation. Countries such as Netherlands, Sweden, Portugal, the United Kingdom, and Ireland have very low numbers recovered cases. In China and Iran, the rate of recovery is very high. It was because of their doctors and the rules followed throughout the country. The countries with zero rate of recovery are Netherlands and the United Kingdom. It shows that these two countries are in a very dangerous situation.

Many viral diseases have affected people for many years. When compared with the Spanish flu, the death rate of COVID-19 is very low. While analyzing the countries affected due to coronavirus outbreak, as per Fig. 7.16, Italy has the highest mortality rate, where more than 1000 people have died in a single day. In a survey, most of the people died due to the coronavirus infection are from the United States. More than 42,000 people died due to the disease in America. One relief in this COVID-19 outbreak is it does not cause more number of deaths when compared with the other previous diseases. The total mortality rate of the novel coronavirus is 3.4. It shows that only more number of people are affected but the death rate is very low. So there is more chance of people to recover from the disease. Belgium has reported more number of death cases. Its death rate is four times higher than that of America. Even though some countries have high mortality rates, many of the countries have low mortality rates. It was one of the positive things about COVID-19.



**FIGURE 7.16** The total number of confirmed cases versus deaths, as of April 25, 2020.

## 7. Comparison of COVID-19 in March and April

Since December 31, 2019, thousands of people are affected due to the novel coronavirus infection. Its attack not even leaves a small island or territories. More than 200,000 people are affected due to COVID-19 in 2020. Even though the infection has been ruining many lives, there are a lot of differences in the caution of infection on March and April. The major differences and effectiveness of the disease between the months of March and April are listed in the following:

**1. Topmost COVID-19-affected countries:** In the month of March, the most affected countries are China, Italy, Korea, Iran, Spain, France, Germany, International Conveyance (Diamond princess), America, and Japan. The graph of the confirmed cases in the topmost countries is shown in [Fig. 7.10](#). From this graph, it is clear that China is the most affected country due to the disease. As it has the highest population and is crowded, the level of spreading of the virus is drastically high. In the month of April, the most affected countries are America, Spain, Italy, France, Germany, the United Kingdom, Turkey, Iran, China, and Russia. The graph of the active cases in the topmost affected countries is shown in [Fig. 7.14](#). From this graph, it has been declared that America leads the count of confirmed cases than China. America reached more than 200,000 people who were affected by the disease.

**2. Count of recovered cases:** In the month of March, people had no awareness about the coronavirus infection. So they did not mind a mild cold or fever. This led to the growth of confirmed cases. Doctors struggled to treat people who were affected by the disease. China was the best example of recovery. In the beginning, China was the one that had a high number of infected cases. But later it slowly recovered from the disease.

In the month of April, the level of recovery is much higher than the previous month. In spite of getting more number of confirmed cases, the count of recovered cases was also good. Day by day, more amount of people recovered from the disease and were discharged safely to their homes. Many people have awareness about COVID-19 in the recent days.

**3. Spread of the virus:** In the month of March, many people were affected by the disease. This was mainly because of the lack of awareness of the people. America got the first place in the list of affected countries because of the increased spreading of the virus. The vulnerability of the disease is very high. People of all ages are infected by the disease, and elders do not survive the disease. Then the spread of virus slowly reduced. In the month of April, there was a less level of spread than the previous month. This was possible only because of the lockdown ordered by all the countries' governments. In April, most of the countries declared a whole lockdown for many days to prevent the spread of the virus. It was the best decision took by governments because only with the help of quarantine, most lives were saved.

**4. Mortality rate:** In the month of March, the mortality rate of COVID-19 worldwide was 2.5%. People aged more than 50 years died due to the virus in the beginning. But as the virus started to spread more and more, there was an issue where teenaged people also died due to the disease. It was considered as the big problem in all countries. Nonidentification of the disease in the starting stage was an important reason for the high mortality rate.

In the month of April, the mortality rate of COVID-19 in all countries as 3.4%. As the count of confirmed cases increases, the number of death cases also increases. The major reason for the growth of high mortality rate is there is no perfect treatment for the people who have reached the critical stage of the disease. But people who are in the starting stage of the disease can be easily rectified.

## 8. Conclusion

At present, there is no therapy or vaccines available to treat COVID-19. The national governments and the WHO and its partners are working urgently to coordinate the rapid development of medical countermeasures. Ongoing, comprehensive, and verified global surveillance data about COVID-19 is crucial for responding at the global, national, and local levels. Epidemiologic surveillance information is collected from all countries, territories, and local areas and the collected information is made accessible through multiple channels, including a dynamic dashboard and a daily situation report, and also displayed in various government websites.

The outbreak of COVID-19 has drastically affected various countries. Countries that had less number of confirmed cases in the first 30 days had seen a rapid increase in the next 30 days. The dataset was prepared for 175 affected countries and territories as of March 19, 2020 [56]. The spread of the virus was so rapid that on March 26, 2020, it affected around 186 countries and territories and on March 27, 2020, it reached 197 countries and territories. Another thing to be noted is that China, where the outbreak started, has almost recovered from the effects of COVID-19 [13]. But the rest of the world is still struggling to control the spread of this novel virus, especially Italy, Spain, and the United States are the worst affected countries due to the pandemic. As of now, there is no vaccine or drug for this novel virus, so the only way to protect from this disease is to avoid exposure to the virus.

Self-isolation and social distancing can only be the possible ways to reduce the number of cases in each country. With the world facing an unprecedented threat, there is an opportunity to emerge with stronger health systems and improved global collaboration to face the next health threat. As we focus on the immediate response to the COVID-19 crisis, it is important to keep in mind the depth of the consequences that are already being felt across the globe. The pandemic has left a great impact on us, so it is our duty now to be responsible and make the world a safer place to live in the future.

## References

- [1] T. Chen, R.K.-K. Leung, Z. Zhou, R. Liu, X. Zhang, L. Zhang, Investigation of key interventions for shigellosis outbreak control in China, *PLoS One* 9 (4) (2014).
- [2] T. Chen, R.K.-K. Leung, R. Liu, F. Chen, X. Zhang, J. Zhao, S. Chen, Risk of imported ebola virus disease in China, *Trav. Med. Infect. Dis.* 12 (6) (2014) 650–658.
- [3] G. Zhou, G. Yan, Severe Acute Respiratory Syndrome Epidemic in Asia.

- [4] H. Nishiura, S.-M. Jung, N.M. Linton, R. Kinoshita, Y. Yang, K. Hayashi, T. Kobayashi, B. Yuan, A.R. Akhmetzhanov, The Extent of Transmission of Novel Coronavirus in Wuhan, China, 2020, 2020.
- [5] C. Wang, X. Wang, Prevalence, nosocomial infection and psychological prevention of novel coronavirus infection, *Chin. General Pract. Nurs.* 18 (2020) 2–3.
- [6] B. Yi, Y. Chen, X. Ma, J. Rui, J.-A. Cui, H. Wang, J. Li, S.-F. Chan, R. Wang, K. Ding, et al., Incidence dynamics and investigation of key interventions in a dengue outbreak in Ningbo city, China, *PLoS Negl. Trop. Dis.* 13 (8) (2019) e0007659.
- [7] I.M. Longini, A. Nizam, S. Xu, K. Ungchusak, W. Hanshaoworakul, D.A. Cummings, M.E. Halloran, Containing pandemic influenza at the source, *Science* 309 (5737) (2005) 1083–1087.
- [8] T. Chen, Q. Chen, R. Liu, A. Szot, S. Chen, J. Zhao, S. Zhou, The transmissibility estimation of influenza with early stage data of small-scale outbreaks in Changsha, China, 2005–2013, *Epidemiol. Infect.* 145 (3) (2017) 424–433.
- [9] N.E.M. Khalifa, M.H.N. Taha, A.E. Hassanien, S. Elghamrawy, Detection of Coronavirus (COVID-19) Associated Pneumonia Based on Generative Adversarial Networks and a Fine-Tuned Deep Transfer Learning Model Using Chest X-Ray Dataset, arXiv, preprint arXiv:2004.01184.
- [10] M. Xu, Y. Zhang, Investigation on the psychological status of the first batch of clinical first-line support nurses to fight against pneumonia caused by novel coronavirus, *Chin. Nurs. Res.* 34 (2020) 1–3.
- [11] T. Que, V. Wong, K. Yuen, Treatment of severe acute respiratory syndrome with lopinavir/ritonavir: a multicentre retrospective matched cohort study, *Hong Kong Med. J.* 9 (6) (2003) 399–406.
- [12] Z. Zhao, F. Zhang, M. Xu, Clinical analysis of 190 cases of outbreak with atypical pneumonia in Guangzhou in spring, 2003, *Guangdong Med. J.* 24 (S1) (2003) 73–76.
- [13] Z. Tang, W. Zhao, X. Xie, Z. Zhong, F. Shi, J. Liu, D. Shen, Severity Assessment of Coronavirus Disease 2019 (COVID-19) Using Quantitative Features From Chest Ct Images, arXiv, preprint arXiv:2003.11988.
- [14] Z.-Q. Xia, J. Zhang, Y.-K. Xue, G.-Q. Sun, Z. Jin, Modeling the transmission of middle east respirator syndrome corona virus in the Republic of Korea, *PLoS One* 10 (12) (2015).
- [15] T.G. Ksiazek, D. Erdman, C.S. Goldsmith, S.R. Zaki, T. Peret, S. Emery, S. Tong, C. Urbani, J.A. Comer, W. Lim, et al., A novel coronavirus associated with severe acute respiratory syndrome, *N. Engl. J. Med.* 348 (20) (2003) 1953–1966.
- [16] C. Chu, V. Cheng, I. Hung, M. Wong, K. Chan, K. Chan, R. Kao, L. Poon, C. Wong, Y. Guan, et al., Role of lopinavir/ritonavir in the treatment of SARS: initial virological and clinical findings, *Thorax* 59 (3) (2004) 252–256.
- [17] A. Banerjee, D. Santra, S. Maiti, Energetics Based Epitope Screening in SARS CoV-2 (COVID 19) Spike Glycoprotein by Immuno-Informatic Analysis Aiming to a Suitable Vaccine Development, bioRxiv.
- [18] C.-C. Lai, T.-P. Shih, W.-C. Ko, H.-J. Tang, P.-R. Hsueh, Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and corona virus disease-2019 (COVID-19): the epidemic and the challenges, *Int. J. Antimicrob. Agents* (2020) 105924.
- [19] S. Riley, C. Fraser, C.A. Donnelly, A.C. Ghani, L.J. Abu-Raddad, A.J. Hedley, G.M. Leung, L.-M. Ho, T.-H. Lam, T.Q. Thach, et al., Transmission dynamics of the etiological agent of sars in Hong Kong: impact of public health interventions, *Science* 300 (5627) (2003) 1961–1966.
- [20] C. Sohrabi, Z. Alsafi, N. O'Neill, M. Khan, A. Kerwan, A. Al-Jabir, C. Iosifidis, R. Agha, World Health Organization declares global emergency: a review of the 2019 novel coronavirus (COVID-19), *Int. J. Surg.* (2020).
- [21] Q.-V. Pham, D. Nguyen, W.-J. Hwang, P. N. Pathirana, et al., Artificial Intelligence (AI) and Big Data for Coronavirus (COVID-19) Pandemic: A Survey on the State-of-the-Arts.

- [22] S. Cauchemez, C. Fraser, M.D. Van Kerkhove, C.A. Donnelly, S. Riley, A. Rambaut, V. Enouf, S. van der Werf, N.M. Ferguson, Middle east respiratory syndrome coronavirus: quantification of the extent of the epidemic, surveillance biases, and transmissibility, *Lancet Infect. Dis.* 14 (1) (2014) 50–56.
- [23] R. Dandekar, G. Barbastathis, Neural Network Aided Quarantine Control Model Estimation of Covid Spread in Wuhan, China, arXiv, preprint arXiv:2003.09403.
- [24] N. Chen, M. Zhou, X. Dong, J. Qu, F. Gong, Y. Han, Y. Qiu, J. Wang, Y. Liu, Y. Wei, et al., Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study, *Lancet* 395 (10223) (2020) 507–513.
- [25] P. Zhou, X.-L. Yang, X.-G. Wang, B. Hu, L. Zhang, W. Zhang, H.-R. Si, Y. Zhu, B. Li, C.-L. Huang, et al., A pneumonia outbreak associated with a new coronavirus of probable bat origin, *Nature* 579 (7798) (2020) 270–273.
- [26] V.M. Corman, O. Landt, M. Kaiser, R. Molenkamp, A. Meijer, D.K. Chu, T. Bleicker, S. Brünink, J. Schneider, M.L. Schmidt, et al., Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR, *Euro Surveill.* 25 (3) (2020) 2000045.
- [27] J.Y. Kim, J.Y. Song, Y.K. Yoon, S.-H. Choi, Y.G. Song, S.-R. Kim, H.-J. Son, S.-Y. Jeong, J.-H. Choi, K. M. Kim, et al., Middle east respiratory syndrome infection control and prevention guideline for healthcare facilities, *Inf. Chemother.* 47 (4) (2015) 278–302.
- [28] A.S. Omrani, M.M. Saad, K. Baig, A. Bahloul, M. Abdul-Matin, A.Y. Alaidaroos, G.A. Almakhlaifi, M. M. Albarrak, Z.A. Memish, A.M. Albarrak, Ribavirin and interferon alfa-2a for severe middle east respiratory syndrome coronavirus infection: a retrospective cohort study, *Lancet Infect. Dis.* 14 (11) (2014) 1090–1095.
- [29] B.R. Beck, B. Shin, Y. Choi, S. Park, K. Kang, Predicting Commercially Available Antiviral Drugs That May Act on the Novel Coronavirus (2019-nCoV), Wuhan, China Through a Drug-Target Interaction Deep Learning Model, bioRxiv.
- [30] E. Ong, M.U. Wong, A. Huffman, Y. He, COVID-19 Coronavirus Vaccine Design Using Reverse Vaccinology and Machine Learning, bioRxiv.
- [31] S.P. Adhikari, S. Meng, Y.-J. Wu, Y.-P. Mao, R.-X. Ye, Q.-Z. Wang, C. Sun, S. Sylvia, S. Rozelle, H. Raat, et al., Epidemiology, causes, clinical manifestation and diagnosis, prevention and control of coronavirus disease (COVID-19) during the early outbreak period: a scoping review, *Infect. Dis. Poverty* 9 (1) (2020) 1–12.
- [32] B. Sarkar, M. A. Ullah, F. T. Johora, M. A. Taniya, Y. Araf, The Essential Facts of Wuhan Novel Coronavirus Outbreak in China and Epitope-Based Vaccine Designing Against 2019-nCoV, bioRxiv.
- [33] Y.-H. Jin, L. Cai, Z.-S. Cheng, H. Cheng, T. Deng, Y.-P. Fan, C. Fang, D. Huang, L.-Q. Huang, Q. Huang, et al., A rapid advice guideline for the diagnosis and treatment of 2019 novel coronavirus (2019-nCoV) infected pneumonia (standard version), *Mil. Med. Res.* 7 (1) (2020) 4.
- [34] S. Zhang, Q. Hu, Z. Deng, S. Hu, F. Liu, S. Yu, R. Liu, C. He, H. Li, L. Gao, et al., Transmissibility of acute haemorrhagic conjunctivitis in small-scale outbreaks in Hunan province, China, *Sci. Rep.* 10 (1) (2020) 1–9.
- [35] T.-M. Chen, J. Rui, Q.-P. Wang, Z.-Y. Zhao, J.-A. Cui, L. Yin, A mathematical model for simulating the phase-based transmissibility of a novel coronavirus, *Infect. Dis. Poverty* 9 (1) (2020) 1–8.
- [36] W. Wang, J. Tang, F. Wei, Updated understanding of the outbreak of 2019 novel coronavirus (2019-nCoV) in Wuhan, China, *J. Med. Virol.* 92 (4) (2020) 441–447.
- [37] Infection Prevention and Control During Health Care When Novel Coronavirus (nCoV) Infection Is Suspected, 2020 (Accessed 26 January 2020), [www.who.int/publications-detail/infection-prevention-and-control-duringhealth-care-when-novel-coronavirus-\(ncov\)-infection-is-suspected-20200125](http://www.who.int/publications-detail/infection-prevention-and-control-duringhealth-care-when-novel-coronavirus-(ncov)-infection-is-suspected-20200125).

- [38] L. Li, M. Ren, Y. Zhang, W. Li, H. Zhao, L. Liang, et al., Lung CT image of a confirmed case of the 2019 novel coronavirus (2019-nCoV) infected pneumonia (with differential diagnosis of the SARS), *Yixue Xinzhi* 30 (1) (2020) 4–6.
- [39] Z. Hou, F. Du, H. Jiang, X. Zhou, L. Lin, Assessment of Public Attention, Risk Perception, Emotional and Behavioural Responses to the Covid-19 Outbreak: Social Media Surveillance in China, 2020.
- [40] O. Gozes, M. Frid-Adar, N. Sagie, H. Zhang, W. Ji, H. Greenspan, Coronavirus Detection and Analysis on Chest CT With Deep Learning, arXiv, preprint arXiv:2004.02640.
- [41] A review of coronavirus disease-2019 (Covid-19), *Indian J. Pediatr.* (2020) published online.
- [42] G. Guyatt, A. Oxman, G. Vist, G. working group, et al., An emerging consensus on rating quality of evidence and strength of recommendations, *Br. Med. J.* (2008).
- [43] S. Su, G. Wong, W. Shi, J. Liu, A.C. Lai, J. Zhou, W. Liu, Y. Bi, G.F. Gao, Epidemiology, genetic recombination, and pathogenesis of coronaviruses, *Trends Microbiol.* 24 (6) (2016) 490–502.
- [44] G. Carpenter, S. Cohen, Epidermal growth factor, *J. Biol. Chem.* 265 (14) (1990) 7709–7712.
- [45] A. Maharatna, Fertility, mortality and gender bias among tribal population: an indian perspective, *Soc. Sci. Med.* 50 (10) (2000) 1333–1351.
- [46] A.V. D’Amico, J. Moul, P.R. Carroll, L. Sun, D. Lubeck, M.-H. Chen, Prostate specific antigen doubling time as a surrogate end point for prostate cancer specific mortality following radical prostatectomy or radiation therapy, *J. Urol.* 172 (5) (2004) S42–S47.
- [47] Coronavirus, World Health Organization, cited 19 January 2020, Available from: <https://www.who.int/health-topics/coronavirus>.
- [48] Available from: [www.worldometers.info/coronavirus/](http://www.worldometers.info/coronavirus/), *Worldometers.info*, cited 30 April 2020.
- [49] B. Ghoshal, A. Tucker, Estimating Uncertainty and Interpretability in Deep Learning for Coronavirus (COVID-19) Detection, arXiv, preprint arXiv:2003.10769.
- [50] B. Udugama, P. Kadhiresan, H.N. Kozlowski, A. Malekjahani, M. Osborne, V.Y. Li, H. Chen, S. Mubareka, J. Gubbay, W.C. Chan, Diagnosing COVID-19: the disease and tools for detection, *ACS Nano* (2010).
- [51] J. Rydén, S.E. Alm, The effect of interaction and rounding error in two-way ANOVA: example of impact on testing for normality, *J. Appl. Stat.* 37 (10) (2010) 1695–1701.
- [52] T. Niedoba, P. Piéta, Applications of ANOVA in mineral processing, *Min. Sci.* 23 (2016).
- [53] M.T. Sow, Using ANOVA to examine the relationship between safety & security and human development, *J. Int. Bus. Econom.* 2 (4) (2014) 101–106.
- [54] H. Li, S.-M. Liu, X.-H. Yu, S.-L. Tang, C.-K. Tang, Coronavirus disease 2019 (COVID-19): current status and future perspective, *Int. J. Antimicrob. Agents* (2020) 105951.
- [55] Q. Li, X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, R. Ren, K.S. Leung, E.H. Lau, J.Y. Wong, et al., Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia, *N. Engl. J. Med.* (2020).
- [56] A.R. Fehr, R. Channappanavar, S. Perlman, Middle east respiratory syndrome: emergence of a pathogenic human coronavirus, *Annu. Rev. Med.* 68 (2017) 387–399.
- [57] L.E. Gralinski, V.D. Menachery, Return of the coronavirus: 2019-nCoV, *Viruses* 12 (2) (2020) 135.
- [58] M.R. Stoline, The status of multiple comparisons: simultaneous estimation of all pairwise comparisons in one-way ANOVA designs, *Am. Stat.* 35 (3) (1981) 134–141.
- [59] A.C. Tamhane, Multiple comparisons in model i one-way ANOVA with unequal variances, *Commun. Stat. Theor. Methods* 6 (1) (1977) 15–32.
- [60] H.M. Park, Comparing Group Means: T-Tests and One-Way ANOVA Using Stata, SAS, R, and SPSS.