



Informative Statistical Analysis on COVID Pandemic and Population as a Whole

Christian Osagie¹; Chukwuemeka Ezeliara Daniel²; Chinenye Adaobi Igwegbe³; Soumya Ghosh⁴; Shabnam Ahmadi^{5,6*}

¹Environmental and Natural Science, Brandenburg University of Technology, Cottbus-Senftenberg, Germany.

²Industrial and Production Engineering, Nnamdi Azikiwe University, Awka, Nigeria.

³Department of Chemical Engineering, Nnamdi Azikiwe University, Awka, Nigeria.

⁴Department of Genetics, Faculty of Natural and Agricultural Sciences, University of the Free State, Bloemfontein 9301, South Africa.

⁵Department of Environmental Health Engineering, Faculty of Medical Sciences, Tarbiat Modares University, Tehran, Iran.

⁶Faculty of Environmental Health, Zabol University of Medical Sciences, Zabol, Iran.

***Corresponding Author(s): Shabnam Ahmadi**

Faculty of Environmental Health, Zabol University of
Medical Sciences, Zabol, Iran.

Email: sh.ahmadi398@gmail.com

Abstract

COVID-19 is the latest pandemic that has triggered respiratory diseases worldwide at an alarmingly fast pace. The symptoms of this disease are acute respiratory illness, hyperthermia, cough, sore throat, breathlessness and diarrhea. The present study through a statistical analysis relates the COVID-19 incidence and mortality dynamics. The required data (i.e., total number of confirmed deaths reported globally in the year 2020, by 23rd July, the population in millions, and number of deaths per million) was collected. The COVID-19 confirmed deaths, the population in millions, and number of deaths per million as of July 23, 2020, by country were analyzed via the descriptive and inferential statistics using the SPSS version 21 and Minitab 17 to understand their relationship and the impact of COVID-19 on the world population. The data generated was normally distributed and suitable to portray the variables statistically. The Pearson correlation and Kendall's correlations revealed that the number of coronavirus death rates and death per million have the strongest relationship. However, there is no significant relationship at all with the population in millions and the death per million which has a significant value of 0.978. The results of Kendall's correlations show that there is a significant relationship between the numbers of coronavirus death rates with the population in millions with a significant value of 0.000. The current data with its statistical analysis helped to understand the impact of COVID-19 on the world population and the present global scenario.

Received: Nov 28, 2022

Accepted: Dec 13, 2022

Published Online: Dec 15, 2022

Journal: Annals of Epidemiology and Public Health

Publisher: MedDocs Publishers LLC

Online edition: <http://meddocsonline.org/>

Copyright: © Ahmadi S (2022). *This Article is distributed under the terms of Creative Commons Attribution 4.0 International License*

Keywords: Environment; COVID-19; Novel Corona virus; Exposure; Statistical data analysis; Descriptive statistics.



Specifications table

Subject area	COVID 19
More specific subject area	COVID 19 pandemic
Protocol name	The present study through a statistical analysis relates the COVID-19 incidence and mortality dynamics
Reagents/tools	NA
Experimental design	Data collections
Trial registration	NA
Name and reference of original method	NA
Resource availability	N/A

Method details

Background

Coronavirus 2019 or COVID 19 has become the most recent global pandemic to cause respiratory disease [1-6]. It was first recognized among twenty-seven individuals with an unknown etiology of Pneumonia on 31st December 2019, in Wuhan district, capital of Hebei province of China with a population of 11 million [7-9]. Since then, the dispersion of the virus has posed a global threat affecting 250 countries, ultimately leading the World Health Organization (WHO) to declare the upsurge as an epidemic on 31st January 2020, and eventually as a global pandemic on 11th March 2020 [10].

Structurally, the COVID-19 has shown similarity with the acute respiratory syndrome (SARS) which rooted two major pandemics: SARS [11] and Middle East Respiratory Syndromes (MERS) [12, 13] in the last two decades. Similar to SARS-CoV infection, COVID-19 infection varies in its course from asymptomatic to acute respiratory distress syndrome (ARDS) [14-16]. Although the commonly observed symptoms are pneumonia, cough, tiredness, complicated dyspnea, aches, and fever while the less common indications are headache, runny nose, hemoptysis, phlegm-producing cough, and diarrhea [15, 17]. Cases have been detected were people with minor symptoms take a week or more to recuperate. On the other hand, critical patients regress into respiratory failure, usually leading to death [8]. Precisely, middle aged/elderly people (>59 years) with comorbidities, such as cirrhosis, coronary heart disease, hypertension, Parkinson's disease, and cardiovascular diseases, generally get affected with serious illness, with high mortality and morbidity rate [17, 18], with fatality rates more incident in males than in females. However, the infants and children experience minor symptoms and improved prognosis, regardless of the sex

differences [19, 20]. The mean growth time for COVID-19 has been detected as 5.2 days while the duration of the inception of symptoms until death is 14 days [21, 22]. Studies have also indicated that the geographical distribution of COVID-19 is affected by warm climates as it reduces its infectivity rates [23], along with factors such as air pollution leading to viral spread [24].

COVID-19 is contagious [1] and the dissemination primarily occurs through coughs or sneezes, saliva droplets, or nose discharge from the infected patients. Initially, it had been described [18, 25, 26] that the infections were zoonotic in nature, i.e., animal-human transmission. A subsequent study showed that the infection spread from human to human [27]. By mid-2020, the world had witnessed over 4,000,000 cases of COVID-19 and several 100 thousands deaths, with the infection spreading rapidly across borders infecting countries mostly by tourists [28], or within a community's day to day life through human contact such as family and hospital settings [14, 29]. As of September 2020, there were 26,661,644 confirmed cases and 876,191 deaths worldwide, out of which 18,786,138 have been reported to be recovered [30].

The present study through a statistical analysis relates the COVID-19 incidence and mortality dynamics. The analysis has specifically focused on the deaths caused worldwide per one million population in a country. The current data may aid to apprehend the impact of COVID-19 on the world population and the present global scenario.

Experimental design, materials, and methods

Sample collection, analytical procedures, and data analysis

The required data (the total number of patients and deaths reported in the world in the year 2020, 23 July) were collected. The relationship between COVID-19 exposure cases, population, and deaths worldwide per million as of July 23, 2020, for each country were analyzed via the descriptive, and inferential statistics using the SPSS version 21 and Minitab 17 to understand their relationship and the impact of COVID-19 on the world population.

Descriptive statistical results

Table 1 shows the descriptive statistics on the confirmed cases, population, and deaths per million. User-defined missing values are treated as missing. All non-missing data are used. The descriptive variables which are the number of confirmed death, the population of the countries, and the number of death per million populations were statistically described to understand the data description. Data has been retrieved from 148 countries.

Table 1: Descriptive statistics on the confirmed cases, population, and deaths per million.

Indicators' name	Data (148 countries)	Mean		Std. Deviation	Variance	Skewness	
	Statistic	Statistic	Std. Error	Statistic	Statistic	Statistic	Std. Error
Confirmed death (absolute)	148	4208.0946	1242.39622	15114.40236	228445158.630	6.696	.199
Population (in millions)	148	49.2847	13.43113	163.39680	26698.515	7.396	.199
Death per million	148	84.2840	12.28667	149.47378	22342.412	2.738	.199
Valid N (listwise)	148	-	-	-	-	-	-

Probability plot on the variables

Figure 1 shows the probability plot for each of the analyzed variables (confirmed cases, population, and deaths per million).

The variables show that the data generated are normally distributed and suitable to portray the variables statistically. These variables are also significant with significant values that are less than 0.005 each.

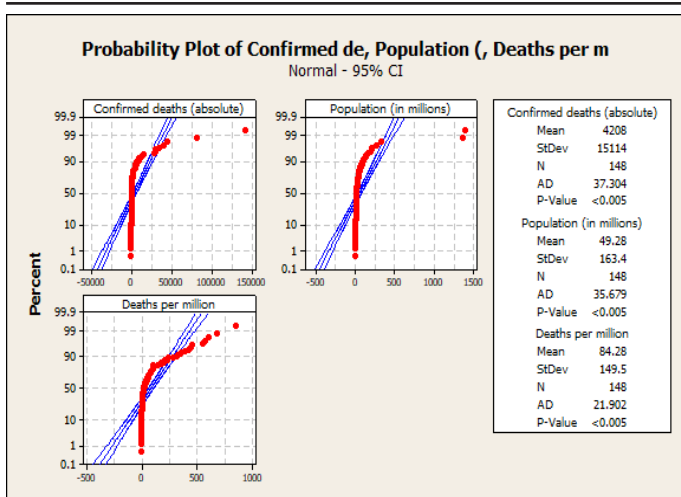


Figure 1: The probability plot of confirmed cases, population, and deaths per million.

Pearson correlation on the collected data

User-defined missing values are treated as missing. Statistics for each pair of variables are based on all the cases with valid data for that pair. Correlation is significant at the 0.01 level. Pearson is a parametric correlation. The results show that there is a significant relationship between the numbers of coronavirus death rates with the population in millions with a significant value of 0.000. Also, there is a significant relationship between the numbers of coronavirus death rates with death per millions with a significant value of 0.000. The Pearson correlation revealed that the number of coronavirus death rates and death per million have the strongest relationship. However, there is no significant relationship at all with the population in millions and the death per million which has a significant value of 0.978.

Table 2: Pearson Correlations on the collected data.

Indicators' name		Confirmed death (absolute)	Population (in millions)	Death per million
Confirmed death (absolute)	Pearson Correlation	1	.289**	.539**
	Sig. (2-tailed)		.000	.000
	N	148	148	148
Population (in millions)	Pearson Correlation	.289**	1	.002
	Sig. (2-tailed)	.000		.978
	N	148	148	148
Death per million	Pearson Correlation	.539**	.002	1
	Sig. (2-tailed)	.000	.978	
	N	148	148	148

Table 3: Nonparametric Correlations on the data.

Correlations		Confirmed death (absolute)	Population (in millions)	Death per million	
Kendall's tau_b	Confirmed death (absolute)	Correlation Coefficient	1.000	.388**	.611**
		Sig. (2-tailed)	.	.000	.000
		N	148	148	148
	Population (in millions)	Correlation Coefficient	.388**	1.000	-.002
		Sig. (2-tailed)	.000	.	.975
		N	148	148	148
	Death per million	Correlation Coefficient	.611**	-.002	1.000
		Sig. (2-tailed)	.000	.975	.
		N	148	148	148
Spearman's rho	Confirmed Death (absolute)	Correlation Coefficient	1.000	.554**	.802**
		Sig. (2-tailed)	.	.000	.000
		N	148	148	148
	Population (in millions)	Correlation Coefficient	.554**	1.000	.002
		Sig. (2-tailed)	.000	.	.981
		N	148	148	148
	Death per million	Correlation Coefficient	.802**	.002	1.000
		Sig. (2-tailed)	.000	.981	.
		N	148	148	148

Kendall’s correlations and the Spearman’s correlations on the data

User-defined missing values are treated as missing. Statistics for each pair of variables are based on all the cases with valid data for that pair. Correlation is significant at the 0.01 level. The correlation analysis shows the results of the Kendall’s correlations and the Spearman’s correlations of nonparametric correlations. The results of Kendall’s correlations show that there is a significant relationship between the numbers of coronavirus death rates with the population in millions with a significant value of 0.000. Also, there is a significant relationship between the numbers of coronavirus death rates with death per millions with a significant value of 0.000. The Kendall’s correlations reveal that the numbers of coronavirus death rate and death per million have the strongest relationship with a correlation coefficient of 61.1%. However, there is no significant relationship at all with the population in millions and the death per million which has a significant value of 0.975. Furthermore, the results of Spearman’s correlations show that there is a significant relationship between the numbers of coronavirus death rates with the population in millions with a significant value of 0.000. Also, there is a significant relationship between the numbers of coronavirus death rates with death per millions with a sig-

nificant value of 0.000. The Spearman’s correlations reveal that the numbers of coronavirus death rate and death per million have the strongest relationship with a correlation coefficient of 80.2%. However, there is no significant relationship at all with the population in millions and the death per million which has a significant value of 0.981.

Monte Carlo statistics on the data

Table 4 shows a nonparametric statistical test for chi-square where user-defined missing values are treated as missing. Statistics for each test are based on all cases with valid data for the variable(s) used in that test. Monte Carlo statistics in chi-square was used as a tool for this analysis.

Optimal binning analysis of the variables

Table 5 shows the descriptive statistics for optimal binning analysis of the variables where the user-defined missing values are treated as missing. The equal frequency method operates on each guide and binning input variable pair only cases with valid data for both variables are used. Table 6 shows the entropy model for the optimal binning analysis of the variables where the user-defined missing values are treated as missing. Smaller model entropy indicates higher predictive accuracy of the binned variable on guide variable death per million.

Table 4: Nonparametric statistical test of frequencies for Chi-square.

		Confirmed Death (absolute)	Population (in millions)	Death per million
Chi-Square		18.405	.986	1.946
Df		130	146	145
Asymp. Sig.		1.000	1.000	1.000
Monte Carlo Sig.	Sig.		1.000	1.000
	99% Confidence Interval	Lower Bound	1.000	1.000
		Upper Bound	1.000	1.000

Table 5: Descriptive statistics for optimal binning.

Indicators' name	N	Minimum	Maximum	Number of Distinct Values	Number of Bins
Confirmed Death (absolute)	148	1.00	142991.00	131	59
Population (in millions)	148	1.14	1392.73	147	71

Table 6: Model entropy for optimal binning.

Indicators' name	Model Entropy
Confirmed Death (absolute)	1.799
Population (in millions)	1.433

Main effect and interaction of the variables

Figures 2-4 show the main effect in the variables of the population in millions, the death per million, and the numbers of coronavirus death rate. The result shows that the numbers of coronavirus deaths have an effect on the population in millions and the death per million and vice versa. Figure 5 shows that there is an interaction between the variables of the population in millions, the death per million, and the numbers of coronavirus death rate. The result shows that the numbers of coronavirus death rate have more impact and interact with few populations in millions and with few deaths per million and vice versa.

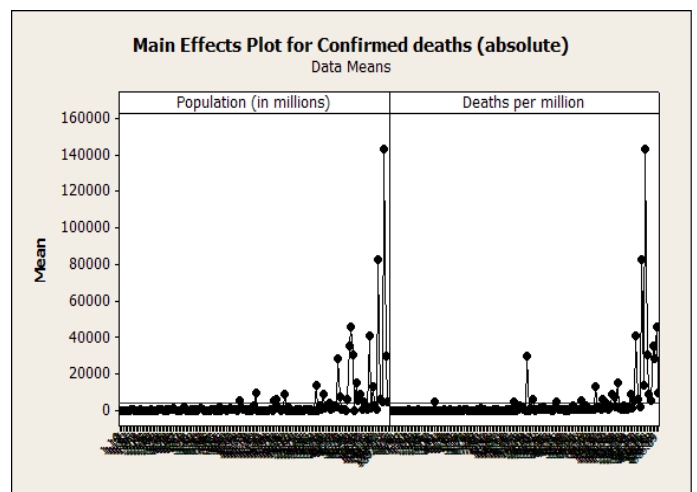


Figure 2: Main effects plot for confirmed deaths (absolute).

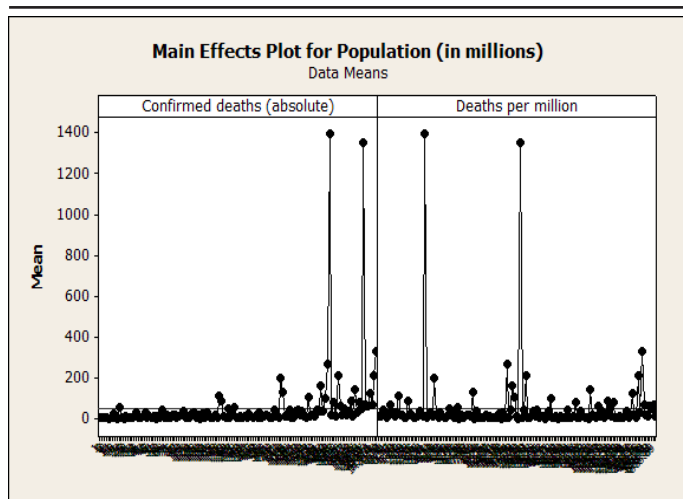


Figure 3: Main effects plot for the population (in millions).

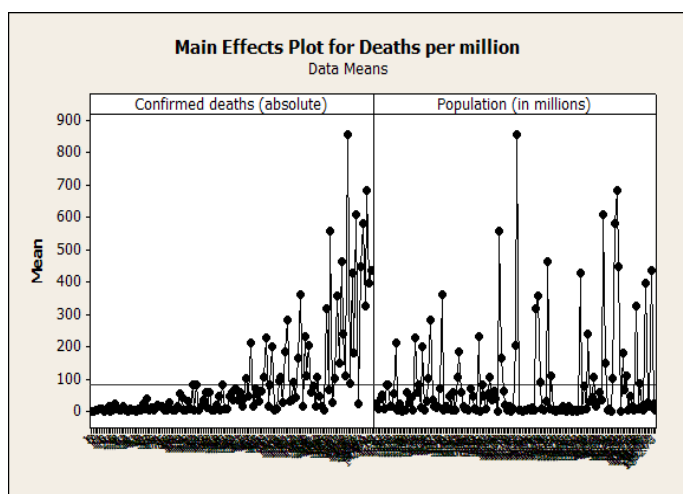


Figure 4: Main effects plot for deaths per million.

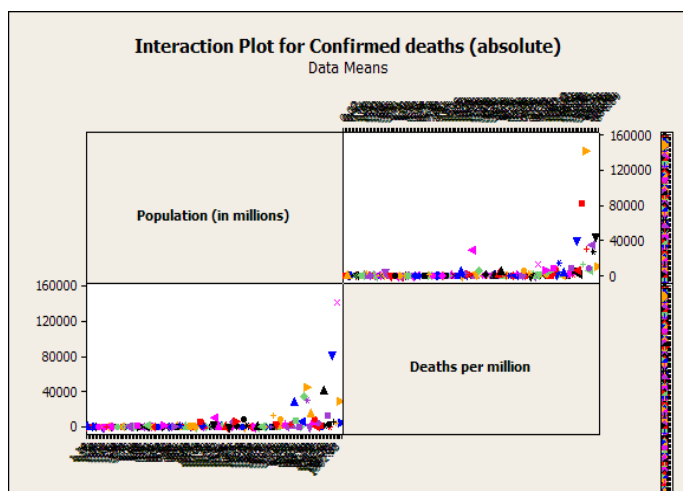


Figure 5: Interaction plot for confirmed deaths (absolute).

Conclusion

The association between COVID-19 reported deaths, the population in millions and the number of deaths per million as of 23 July 2020 by nation was evaluated by detailed and inferential statistics using the SPSS version 21 and Minitab 17 to understand their relationship and the impact of COVID-19 on the world population. The data produced is normally distributed and suitable for statistical representation of the variables. The Pearson correlation and Kendall correlations have dem-

onstrated that the number of coronavirus deaths and deaths per million have the greatest relationship. There is, however, no substantial association with the population in millions and death per million. The findings of Kendall's analyses indicate that there is a substantial association between the numbers of coronavirus death rates with the population in millions with a specific significant value of 0.000. The present data and analysis helped to explain the effect of COVID-19 on the global population and the current global scenario.

Ethical issues

The authors declare that there are no ethical issues involve in the study

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgements

We are thankful to Dr. Swagata Ghosh, Assistant Professor of English, Symbiosis Institute of Technology, Symbiosis International University, Pune, India for editing this manuscript.

References

1. Fauci AS, Lane HC, Redfield RR. Covid-19-Navigating the uncharted, *New Engl J Med.* 2020; 382: 1268-1269.
2. Ghosh S, Bornman C, Zafer MM. Antimicrobial Resistance Threats in the emerging COVID-19 pandemic: Where do we stand? *Journal of Infection and Public Health.* 2021; 14: 555-560.
3. Meskini M, Rezghi Rami M, Maroofi P, Ghosh S, et al. An overview on the epidemiology and immunology of COVID-19, *Journal of Infection and Public Health.* 2021.
4. Alhourri AL, Salloum A, Harfouch RM, Ghosh S. Possible Side Effects of using Detergents during the Covid19 Pandemic in Syria. *Annals of Clinical Cases.* 2020; 1: 071-074.
5. Ghosh S, Al-Sharify ZT, Maleka MT, Onyeaka H, Maleke M, et al. Propolis efficacy on SARS-COV viruses: a review on antimicrobial activities and molecular simulations, *Environ Sci Pollut Res.* 2022; 29: 58628-58647.
6. Alkafaas SS, Abdallah AM, Ghosh S, Loutfy SA, Elkaffas SS, et al. Hessian, Insight into the role of clathrin-mediated endocytosis inhibitors in SARS-CoV-2 infection, *Rev Med Virol.* 2022.
7. Alavi-Moghaddam M. A Novel Coronavirus Outbreak from Wuhan City in China, Rapid Need for Emergency Departments Preparedness and Response; a Letter to Editor. *Archives of Academic Emergency Medicine.* 2020; 8.
8. Li T, Wei C, Li W, Hongwei F, Shi J. Beijing Union Medical College Hospital on "pneumonia of novel coronavirus infection" diagnosis and treatment proposal, *Medical Journal Peking Union Medical College Hospital.* 2020.
9. Sokouti M, Sadeghi R, Pashazadeh S, Eslami S, Sokouti M, et al. Comparative Global Epidemiological Investigation of SARS-CoV-2 and SARS-CoV Diseases Using Meta-MUMS Tool Through Incidence, Mortality, and Recovery Rates, *Archives of Medical Research.* 2020; 51: 458-463.
10. Director W. WHO director general's opening remarks at the media briefing on COVID-19. 2020.
11. Poon LLM, Leung CSW, Tashiro M, Chan KH, Wong BWY, et al. Peiris, Rapid detection of the severe acute respiratory syndrome (SARS) coronavirus by a loop-mediated isothermal amplification

- assay, *Clinical Chemistry*. 2004; 50: 1050-1052.
12. Zaki AM, van Boheemen S, Bestebroer TM, Osterhaus ADME, Fouchier RAM. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia, *New Engl J Med*. 2012; 367: 1814-1820.
 13. Zhou P, Yang XL, Wang XG, Hu B, Zhang L, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*. 2020; 579: 270-273.
 14. Chan JFW, Yuan SF, Kok KH, To KKW, Chu H, et al. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *Lancet*. 2020; 395: 514-523.
 15. Huang Y, Wang X, Li L, Ren J, Zhao Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China, *Lancet (London, England)* 2020; 395: 497-506.
 16. Wang D, Hu B, Hu C, Zhu F, Liu X, et al. Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China. *JAMA*. 2020; 323: 1061.
 17. Fan JC, Liu XD, Pan WM, Douglas MW, Bao SS. Epidemiology of coronavirus disease in Gansu Province, China, 2020. *Emerg Infect Dis*. 2020; 26: 1257-1265.
 18. Li Q, Guan XH, Wu P, Wang XY, Zhou L, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia, *New Engl J Med*. 2020; 382: 1199-1207.
 19. Nicholls JM, Poon LLM, Lee KC, Ng WF, et al. Lung pathology of fatal severe acute respiratory syndrome. *Lancet*. 2003; 361: 1773-1778.
 20. Chen J, Subbarao K. The immunobiology of SARS, *Annual Review of Immunology*. 2007; 25: 443-472.
 21. Wang WE, Tang JM, Wei FQ. Updated understanding of the outbreak of 2019 novel coronavirus (2019-nCoV) in Wuhan, China. *Journal of Medical Virology*. 2020; 92: 441-447.
 22. Rothe C, Schunk M, Sothmann P, Bretzel G, Froeschl G, et al. Transmission of 2019-nCoV Infection from an Asymptomatic Contact in Germany. *New Engl J Med*. 2020; 382: 970-971.
 23. Benedetti F, Pachetti M, Marini B, Ippodrino R, Gallo RC, et al. Inverse correlation between average monthly high temperatures and COVID-19-related death rates in different geographical areas, *Journal of Translational Medicine*. 2020; 18.
 24. Wu X, Nethery RC, Sabath BM, Braun D, Dominici F. Exposure to air pollution and COVID-19 mortality in the United States: A nationwide cross-sectional study, *medRxiv*. 2020.
 25. Backer JA, Klinkenberg D, Wallinga J. Incubation period of 2019 novel coronavirus (2019-nCoV) infections among travellers from Wuhan, China. *Eurosurveillance*. 2020; 25: 10-15.
 26. Peeri NC, Shrestha N, Rahman MS, Zaki R, Tan Z, et al. The SARS, MERS and novel coronavirus (COVID-19) epidemics, the newest and biggest global health threats: what lessons have we learned?, *Int J Epidemiol*. 2020; 49: 717-726.
 27. Adeniyi MO, Ekum MI, Ogunsanya AS, Akinyemi JA, Oke SI, et al. Dynamic model of COVID-19 disease with exploratory data analysis. *Scientific African*. 2020; 1-21.
 28. Duffey RB, Zio E. Prediction of CoVid-19 infection, transmission and recovery rates: A new analysis and global societal comparisons, *Safety Science*. 2020; 129.
 29. Guan W, Ni Z, Hu Y, Liang W, Ou C, et al. Grp, Clinical Characteristics of Coronavirus Disease 2019 in China. *New Engl J Med*. 2020; 382: 1708-1720.
 30. Worldometer. COVID-19 CORONAVIRUS PANDEMIC. 2020.