RESEARCH ARTICLE

REvised Simple statistical insights into the COVID-19 data of Saudi Arabia: figures prior to vaccination campaign [version 2; peer review: 1 approved, 2 approved with reservations]

Omar S. El-Masry

Clinical Laboratory Sciences Department, College of Applied Medical Sciences, Imam Abdulrahman Bin Faisal university, Dammam, Rakkah, 31441, Saudi Arabia

First published: 21 May 2021, 10:410
https://doi.org/10.12688/f1000research.52600.1
Latest published: 04 Aug 2021, 10:410
https://doi.org/10.12688/f1000research.52600.2

Abstract

Background: COVID-19, the disease caused by the newly emerging coronavirus, SARS-COV2, is still a major health burden worldwide as it continues to spread rapidly in many countries after being contained for a while. The aim of the study was to analyze the official current disease estimates in the Kingdom of Saudi Arabia to anticipate future risks and needs.

Methods: Publicly available COVID-19 data published by the Saudi Ministry of Health were analyzed to extract statistical estimates of the disease. These include monthly case fatality rates, death rates/1000, comparison of death figures and regression analysis.

Results: The number of confirmed, recovered and deaths surged in the middle of the outbreak (June and July). The case fatality rates reported later in September-November were the highest despite the decline in the number of confirmed cases. The death rates/1000 were higher during the middle of the outbreak, where the highest numbers of deaths were recorded. The number of recovered cases was the highest as well during this time. Regression analysis showed that the number of deaths was related to that of confirmed cases, especially during the peak time. On the other hand, the number of recovered cases was related to that of confirmed cases at the beginning of the outbreak.

Conclusion: Statistical estimates of COVID-19 fatalities provide simple figures to understand the disease progression pattern and the health care management success in disease containment. However, the absolute numbers should never be disregarded to reflect on the real situation.

Keywords
COVID-19, Case fatality rate, Death rate, outbreak

Open Peer Review

Approval Status ✔ ? ?

version 2 (revision)
04 Aug 2021 view
version 1 21 May 2021 view

1. Yahia Ali Kaabi, Jazan University, Jazan, Saudi Arabia
2. Ahmad Abubakar Suleiman, Kano University of Science and Technology, Wudil, Nigeria
Universityiti Teknologi PETRONAS, Perak, Malaysia
3. Mohammad Nayeem Hasan, Shahjalal University of Science & Technology, Sylhet, Bangladesh

Any reports and responses or comments on the article can be found at the end of the article.
Introduction
COVID-19, the disease caused by the newly merging coronavirus (SARS-COV2) is a major health concern worldwide currently (Zhang et al., 2020). The pandemic is unprecedented in recent history and a lot of restrictions were imposed across the globe in an attempt to contain the spread of the infection hoping to reduce the number of carriers and, consequently, the fatality rates. This, in turn will relieve the pressure on health authorities. The outbreak has spread rapidly from its start point in the mainland of China to more than 200 countries, according to World Health Organization figures (Zhou et al., 2020). The large number of patients who need medical care, especially those with an immediate requirement of intensive care and hospitalization resulted in overwhelming the financial, manpower and the managerial facets of healthcare centers and represented a huge burden on states (Zhang et al., 2020).

Data available so far indicates a remarkable quick spread of the virus, suggesting a high infectivity in comparison to SARS-COV of the same genus (Zhou et al., 2020). Regarding the virus spread, SARS-COV2, which causes COVID-19, can be transmitted via direct contact with infected surfaces or airborne droplets. Vertical transmission was not proven so far, therefore, transmission of the virus to babies from mothers is unlikely, according to the available evidence (Zhu et al., 2020). It was also reported that SARS-COV2 has a higher morbidity and mortality, and higher fatality rates than SARS-COV. At this end, COVID-19 patients suffer from severe suppression of cellular-mediated immune response to the virus as evident by the low count of CD3+, CD4+, and CD8+ cells (Chen et al., 2020). In this respect, SARS-COV, downregulates interferon (IFN) regulatory factor 3 (IRF-3), which ameliorates the antiviral activity mediated by IFN α and β axis; this scenario might also contribute to the suppression of antiviral activity in COVID-19 patients (Kuri and Weber, 2010). In a similar respect, early reports this year indicated that SARS-COV2 is detectable in clinical samples of non-survivors throughout the clinical course of the disease (Weiss and Murdoch, 2020). Moreover, an American study postulated that ethnicity and gender differences might have a potential influence on hospitalization and mortality of COVID-19 patients (Price-Haywood et al., 2020).

The aim of the current study was to provide up to date simple statistical estimates and insights into the official numbers published by the Saudi Ministry of Health and establish an overview to understand the epidemiological data of COVID-19 in the Kingdom of Saudi Arabia. This might help to set future action plans to improve the overall outcomes of disease management.

Methods
This is a data-based study in which the official data on COVID-19 disease that were published by the Saudi Ministry of Health on COVID-19 dashboard were analyzed. Data are available from: https://covid19.moh.gov.sa/. The data published on the dashboard are presented as graphs showing the numbers of confirmed cases, recovered cases and deaths day by day. The numbers are updated on a daily basis.

The following data were extracted from the dashboard for this study: minimum and maximum numbers for confirmed cases of COVID-19, recovered cases of COVID-19, and deaths relating to COVID-19 reported for March-November 2020. The total number of people living in Saudi Arabia was determined according to the estimates of the United Nations that were published on the Worldometer website: https://www.worldometers.info/.

The total number of deaths in each month and case fatality rates were compared between the months in 2020, and death rates/1000 of the population were calculated. Regression analysis was also performed to underline the relationship between different figures reported for each month including the relationships between: (1) the number of confirmed cases (daily reported infections in each month) and the number of the recovered cases (daily reported recoveries in each month), and (2) the number of confirmed cases and the number of monthly reported COVID-19-related deaths (daily reported deaths in each month).

Statistical presentations and calculations were performed using Graphpad Prism (version 7). The data were tested for normality and the Kruskal Wallis test was employed for comparison between groups, followed by Dunn’s multiple comparison test. The difference was considered statistically significant at p ≤ 0.05.
Table 1. Numbers of confirmed and recovered cases, and deaths related to COVID-19 for 2020 (March-November) in the Kingdom of Saudi Arabia.

| Month | March | April | May   | June  | July  | August | September | October | November |
|-------|-------|-------|-------|-------|-------|--------|-----------|---------|----------|
| **Confirmed cases** |       |       |       |       |       |        |           |         |          |
| Minimum | 0     | 154   | 1362  | 1869  | 1573  | 898    | 403       | 323     | 217      |
| Maximum  | 205   | 1351  | 2840  | 4919  | 4207  | 1569   | 833       | 501     | 473      |
| Sum     | 1563  | 22543 | 63045 | 107083| 83253 | 39192  | 18337     | 12559   | 9967     |
| **Recovered cases** |       |       |       |       |       |        |           |         |          |
| Minimum | 0     | 19    | 210   | 806   | 1742  | 718    | 592       | 335     | 357      |
| Maximum  | 50    | 392   | 3559  | 9651  | 7718  | 5426   | 1454      | 626     | 495      |
| Sum     | 165   | 3390  | 60751 | 76457 | 104788| 54966  | 28232     | 14490   | 12940    |
| **Deaths** |       |       |       |       |       |        |           |         |          |
| Minimum | 0     | 3     | 7     | 24    | 20    | 27     | 24        | 14      | 11       |
| Maximum  | 4     | 9     | 23    | 50    | 58    | 42     | 37        | 29      | 20       |
| Sum     | 10    | 159   | 356   | 1173  | 1189  | 1042   | 865       | 626     | 487      |

Table 2. Summary of the Kruskal Wallis statistics comparing the number of deaths in each month (Dunn’s multiple comparison test).

| Month | March | April | May  | June | July | August | September | October | November |
|-------|-------|-------|------|------|------|--------|-----------|---------|----------|
| March | NS    | *     | ***** | **** | **** | ****    | ****      | ****    | ***      |
| April | NS    | NS    | ***   | **** | **** | ****    | ****      | **      | NS       |
| May   | *     | NS    | ****  | **** | **** | ****    | ****      | NS      | NS       |
| June  | ****  | ****  | ****  | NS   | NS   | NS      | NS        | ****    | ****     |
| July  | ****  | ****  | ****  | NS   | NS   | NS      | NS        | ***     | ****     |
| August| ****  | ****  | ****  | NS   | NS   | NS      | NS        | **      | ****     |
| September| ****  | ****  | ****  | NS  | NS  | NS      | NS        | NS      | *        |
| October| ****  | **    | NS    | **** | **   | NS      | NS        |         | NS       |
| November| ***  | NS    | NS    | **** | **** | ****    | ****      | *       | NS       |

NS: Not significant (p value > 0.05). *: Significant difference (p value < 0.05). **: Significant difference (p value < 0.01). ***: Significant difference (p value < 0.001). ****: Significant difference (p value < 0.0001)
Results
Table 1 shows the minimum and maximum numbers recorded for COVID-19 confirmed cases, recovered cases and deaths for each month of 2020. The maximum numbers of confirmed cases were recorded during June, while the maximum recovered cases and deaths were recorded in July. There was an increase in the number of confirmed cases from March-June, and then the numbers started to decrease during July-November. The low number recorded in the first one or two months might be associated with the total number of performed tests, where the capacity of health authorities worldwide may not have been fully prepared for such a surprising pandemic.

Comparing death figures
Table 2 shows the statistical comparison between the numbers of daily deaths recorded in each month (March-November). As the data were not following the Gaussian distribution, the Kruskal Wallis test was performed, followed by Dunn’s multiple comparison test, which indicated that the numbers of deaths recorded in May-November were significantly higher than the numbers recorded in March (p < 0.0001; except for May, p < 0.05). The difference between the numbers of deaths recorded in April and May was not significant, whilst the numbers recorded in June-October were significantly higher than April and May figures (p < 0.0001), except for the comparison between May and October, the difference was not significant, perhaps due to the reduction in number of deaths in October and November. Also, the numbers of deaths in October and November were significantly lower than that reported in June-August indicating the gradual reduction in the total number of deaths. In addition, there was no significant difference between, June, July, August and September figures, which might explain that the average reduction in the death numbers in September was not the desirable one as compared to the reduction observed in October and November. In other words, despite the decline in the number of confirmed cases in September in comparison to June-August period, the decline in the reported deaths in this month was not significantly different from death figures in June-August.

Case fatality rates and death rates/1000
The case fatality rate is defined as the total number of deaths from a particular cause (COVID-19, in this case)/total number of cases x 100. The case fatality rates in each month are presented in Table 3. The lowest numbers of confirmed cases were reported throughout March and case fatality rate was higher than the rate reported for May (0.64 vs. 0.56%), perhaps due to the surprising outbreak of the pandemic and lack of the necessary preparations to confront such situations. Likewise, the rate in April was higher than that of May (0.7). Also, despite the declining figures of confirmed disease and death cases until November following the surge in June and July, the case fatality rates in September-November were the highest (4.71%, 4.98%, and 4.88%, respectively), higher than June and July, wherein the highest numbers of confirmed cases and deaths were reported. This might suggest the poor prognosis in the patient group of non-survivors, or the increased virulence of the virus, but we have to bear in mind the large number of cases in June and July in comparison to the following months. In the same context, the death rates/1000 of the population (Table 3) were also calculated. The results suggested that the highest death rates/1000 were recorded in June, July and August (0.03). Obviously, the lowest figures of this parameter were seen in March and April. In addition, the figures of September-November were less than that of June and July, despite the highest case fatality rates observed in these months.

| Month    | Total number of deaths | Case fatality rate (%) | Death rate/1000 |
|----------|------------------------|------------------------|-----------------|
| March    | 10                     | 0.64                   | 0.00028         |
| April    | 159                    | 0.7                    | 0.0045          |
| May      | 356                    | 0.56                   | 0.0102          |
| June     | 1173                   | 1.09                   | 0.0337          |
| July     | 1189                   | 1.43                   | 0.034           |
| August   | 1042                   | 2.66                   | 0.03            |
| September| 865                    | 4.71                   | 0.025           |
| October  | 626                    | 4.98                   | 0.018           |
| November | 487                    | 4.88                   | 0.014           |

*Case fatality rate = Total number of deaths from Covid-19/total number of confirmed cases*100. Death rate/1000 = Total number of recorded deaths from Covid-19/the number of Saudi population*1000.

*The Saudi population number was estimated at 34813871 at mid-year according to the UN estimates.*
Relationship between the number of confirmed cases and death figures

The relationship between the number of confirmed cases and the number of deaths in each month was investigated by regression analysis (Figure 1). Using this type of analysis we can understand how the change in the number of confirmed cases of COVID-19 is related to the number of deaths. The analysis showed that the slope was significantly non-zero (this means that the relationship between the tested variables was represented as a line with an apparent steepness, where the number of deaths (y) changes according to the change in the number of confirmed cases (x); simply the line is not horizontal) for March, though the association relationship was weak (R = 0.218; p value = 0.009). The slope was close to zero for April, May, and September (R = 0.12, 0.01, and 0.01; p value = 0.06, 0.51, and 0.49), which means that the change in the number of confirmed cases in these months was not associated with the change in the number of deaths. Regarding the figures of regression analysis in June and July, the best fit graph showed a stronger significant association between the number of confirmed cases and deaths (R = 0.5, and 0.68; p value = < 0.0001) as compared to the figures in the other months. August results showed a weaker positive association and a non-zero slope of the best fit line (R = 0.21 and p value = 0.008). The slope was close to zero for October indicating a lack of the relationship, while there was a weak association for November figures (R = 0.206 and p value = 0.012).

![Figure 1. Regression analysis of the number of confirmed cases and deaths.](image-url)

Regression analysis showed that the number of deaths could be weakly related to the number of confirmed cases in March (A), August (F), and November (I). The relationship was significantly strong in June (D) and July (E).
Relationship between the number of confirmed cases and recovery figures

Figure 2 shows the regression analysis results in March-November for the assessment of the relationship between the number of the confirmed cases and recovered cases. The results showed that the number of the recovered cases of COVID-19 patients in March, April, May and August could be explained by the number of confirmed cases (R = 0.22, 0.54, 0.2, and 0.2; p values < 0.008, 0.0001, 0.01, and 0.01). The strongest relationship between the number of confirmed cases and the number of the recovered cases was then observed during April, where 50% of the recovery figures could be explained by the confirmed cases figures. The slope was close to zero for June, July and September analysis, which means that the number of the recovered cases are not related to the numbers of the confirmed cases, although June and July witnessed the highest number of the recovered cases (R = 0.009, 0.014, and 0.08; p value = 0.52 and 0.12). However, these results are consistent with the high death figures observed in June, July and September. There was a significant relationship between the numbers of the recovered and the confirmed cases in October (R = 0.224 and p value = 0.007). The figures for November showed that the relationship was not significant.

Figure 2. Regression analysis of the number of confirmed cases and recovered cases. Regression analysis showed that the number of recovered cases could be significantly related to the number of confirmed cases in March (A), April (B) and May (C). A weaker association was reported for August (F) and October (H).
Discussion

As the recently emerging pandemic of COVID-19 continues to attract attention all over the world with uncertainty about when the disease will vanish and cease new waves, the need continues to grow to assess the capacity of health care authorities to cope with the current challenge and be ready for all possibilities. In this respect, the assessment of the past and present situation data is important for accurate need assessment for future preparation. Therefore, the main aim of the current study was to analyze the data to conclude some statistical figures that could help to understand how effective the management of the situation was, and what should we do for better clinical management outcomes of the COVID-19 crisis. The results showed that the numbers of confirmed deaths reached the peak during June and July, with the highest death rates/1000 of the population. Also, despite the decrease in the number of the confirmed deaths in September-November, the case fatality rates were the highest throughout the period from March-November, which also showed a small number of confirmed cases that might have resulted in the high case fatality rate. The world is still overwhelmed with the pandemic; however, the situation seems to be under control in Saudi Arabia, despite the high case fatality rates recorded in September-November. Case fatality could be affected with multiple factors, such as patients’ age, comorbidities or virus virulence and immune response efficacy in patients. To this end, other data, such as patients’ demographics, comorbidities, medical history, autopsy reports and detailed lab records and clinical follow up data should be available for a more comprehensive figure. In this regard, it was reported that up to 81% of COVID-19 patients might suffer only from a mild disease, which does not require hospitalization (Price-Haywood et al., 2020). Having said that, the other 19%, which is a proportion large enough to overwhelm the health care sector, might suffer from severe symptoms that impose hospitalization and here is where mortality rates could increase as a result of poor clinical management.

It was also recently reported that the average case fatality rate of COVID-19 is less than 5% (Sun et al., 2020). This estimate suggests that the case fatality rates reported in the current study are within the global figures. The surprising hit of the pandemic and/or the virulence of the virus strain and its quick spread needed some time for the governments to confine the situation. In this regard, many comorbidities were linked to the increased severity and mortality of COVID-19 patients, including diabetes mellitus (Kumar et al., 2020). It is of importance to denote that diabetes mellitus is prevalent in Saudi Arabia (Alotaibi et al., 2017). However, more data should be available on COVID-19 casualties before assuming any link between the disease and COVID-19 mortality figures presented herein. It might be also of importance to mention that the case fatality rates of COVID-19 are lower than that of Severe Acute Respiratory Syndrome (SARS) and Middle East Respiratory Syndrome (MERS) (Sun et al., 2020), but we have to bear in mind that the number of cases of COVID-19 patients is much higher than that recorded for the other two counterparts. Similar case fatality rates were reported for both China and Italy (~2.3) at the beginning of the pandemic, with comorbidities that might have contributed significantly to mortality in elderly patients (Porcheddu et al., 2020). These figures are almost similar to the average case fatalities reported in the period (March-September) in Saudi Arabia. In line with the current study, Boretti compared some figures of the outbreak between Saudi Arabia and the United Kingdom in the middle of the outbreak; Boretti reported that the number of cases per million, the number of deaths per million and the number of newly daily deaths per million were much less in Saudi Arabia than those in the United Kingdom (Boretti, 2020).

In order to give a brief account on the figures post vaccination campaign, the case fatality rates were calculated from the officially published figures in the first 6 months in 2021 (January-June). The number of newly recorded cases escalated from ~5000 cases in January to ~37000 in June. This means that following flattening of the curve at the end of 2020, there was an ascending increase in the number of cases again after the start of the vaccination campaign. However, the number of cases still much less than the numbers recorded in the same months in 2020 before vaccination. The case fatality rate in January 2021 was 0.03 (3%), whilst that calculated for February-June was 0.01 (1%), which is similar to that recorded for June 2020; however, there was a clear reduction in the number of confirmed cases and deaths in 2021. Whether the reduction in confirmed cases and deaths numbers is due to vaccination or other factors, this is not confirmed due to lack of information about the number of vaccinated people who contracted the infection nor those who died after being vaccinated. Therefore, we have to wait the release of all official figures and information before we make a final conclusion about the efficacy of vaccination.

In conclusion, the results presented in the current study suggested that the fatality rates of COVID-19 mortality and their association with the number of confirmed cases and the recovery figures should be always seen in the same frame with the absolute figures to reflect the real situation to the policy makers for better needs assessment, future planning and prompt response in crisis time. Amongst the limitations of the current study, and maybe other related published articles, are the lack of information on cases and deaths age groups, comorbidities, ethnicity, and medical history that might have affected disease progression and mortality rates. The availability of such information could help to clarify risk factors and those predisposed for a severe disease clinical course. Also, we have to bear in mind that the number of COVID-19 cases are much higher than that of the disease counterparts SARS and MERS when we look at the fatality rates.
Data availability
All data presented herein were extracted by analyzing the publicly available data on COVID-19 that were uploaded to the COVID-19 data portal established by the Saudi Ministry of Health: https://covid19.moh.gov.sa/.

References

Alotaibi A, Perry L, Gholizadeh L, et al.: Incidence and prevalence rates of diabetes mellitus in Saudi Arabia: An overview. J Epidemiol Glob Health. 2017; 7: 211-218. PubMed Abstract | Publisher Full Text | Free Full Text

Boretti A: COVID-19 fatality rate for Saudi Arabia, updated 3 June 2020. J Glob Antimicrob Resist. 2020; 22: 845-846. PubMed Abstract | Publisher Full Text | Free Full Text

Chen F, Liu ZS, Zhang FR, et al.: [First case of severe childhood novel coronavirus pneumonia in China]. Zhonghua Er Ke Za Zhi. 2020; 58: E005. PubMed Abstract | Publisher Full Text | Free Full Text

Kumar A, Arora A, Sharma P, et al.: Is diabetes mellitus associated with mortality and severity of COVID-19? A meta-analysis. Diabetes Metab Syndr. 2020; 14: 535-545. PubMed Abstract | Publisher Full Text | Free Full Text

Kuri T, Weber F: Interferon interplay helps tissue cells to cope with SARS-coronavirus infection. Virulence. 2010; 1: 273-275. PubMed Abstract | Publisher Full Text | Free Full Text

Porcheddu R, Serra C, Kelvin D, et al.: Similarity in Case Fatality Rates (CFR) of COVID-19/SARS-COV-2 in Italy and China. J Infect Dev Ctries. 2020; 14: 125-128. PubMed Abstract | Publisher Full Text

Price-Haywood EG, Burton J, Fort D, et al.: Hospitalization and Mortality among Black Patients and White Patients with Covid-19. N Engl J Med. 2020; 382: 2534-2543. PubMed Abstract | Publisher Full Text | Free Full Text

Sun P, Qie S, Liu Z, et al.: Clinical characteristics of hospitalized patients with SARS-CoV-2 infection: A single arm meta-analysis. J Med Virol. 2020; 92: 612-617. PubMed Abstract | Publisher Full Text | Free Full Text

Weiss P, Murdoch DR: Clinical course and mortality risk of severe COVID-19. Lancet. 2020; 395: 1014-1015. PubMed Abstract | Publisher Full Text | Free Full Text

Zhang L, Yan X, Fan Q, et al.: D-dimer levels on admission to predict in-hospital mortality in patients with Covid-19. J Thromb Haemost. 2020; 18: 1324-1329. PubMed Abstract | Publisher Full Text | Free Full Text

Zhou MY, Xie XL, Peng YG, et al.: From SARS to COVID-19: What we have learned about children infected with COVID-19. Int J Infect Dis. 2020; 96: 710-714. PubMed Abstract | Publisher Full Text | Free Full Text

Zhu H, Wang L, Fang C, et al.: Clinical analysis of 10 neonates born to mothers with 2019-nCoV pneumonia. Transl Pediatr. 2020; 9: 51-60. PubMed Abstract | Publisher Full Text | Free Full Text
Mohammad Nayeem Hasan

Department of Statistics, Shahjalal University of Science & Technology, Sylhet, Bangladesh

The study entitled “Simple statistical insights into the COVID-19 data of Saudi Arabia: figures prior to vaccination campaign” tries to evaluate the official current disease estimates in the Kingdom of Saudi Arabia to anticipate future risks and needs. The manuscript is well organized and the methodology is considerable for acceptance. However, I suggest some improvements below:

1. Provide some research gaps in the introduction.
2. Please explain all variables and how you categorized them. If needed, provide references.
3. Please follow some recently published papers related to your study from high-impact journals.
4. Please explain, how you reached the final model, do you use any model evaluation techniques like AUROC, multicollinearity, feature selection, P-value, etc.?
5. Please, explain the discussion based on your findings in brief. You can follow them: [Ref 1,2,3 and 4]

References
1. Sangkham S, Islam M, Sarndhong K, Vongruang P, et al.: Effects of fine particulate matter (PM2.5) and meteorological factors on the daily confirmed cases of COVID-19 in Bangkok during 2020–2021, Thailand. Case Studies in Chemical and Environmental Engineering. 2023; 8. Publisher Full Text
2. Haider N, Hasan MN, Guitian J, Khan RA, et al.: The disproportionate case-fatality ratio of COVID-19 between countries with the highest vaccination rates and the rest of the world. IJID Reg. 2023; 6: 159-166 PubMed Abstract | Publisher Full Text
3. Hasan M, Khalil I, Chowdhury M, Rahman M, et al.: Two Decades of Endemic Dengue in Bangladesh (2000-2022): Trends, Seasonality, and impact of Temperature and Rainfall Patterns on transmission dynamics. medRxiv. 2023. Publisher Full Text
4. Haider N, Hasan MN, Guitian J, Khan RA, et al.: The disproportionate case-fatality ratio of COVID-19 between countries with the highest vaccination rates and the rest of the world. IJID Reg. 2023; 6: 159-166 PubMed Abstract | Publisher Full Text

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Public Health

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.
Kingdom of Saudi Arabia prior to the vaccination campaign.

4. The authors have clearly mentioned the source of the datasets used in this investigation. Please add the access dates (format: Date Month Year), e.g., accessed on 1 January 2020.

5. Mathematical equations for regression analysis must be provided:
   1. The number of confirmed cases (x) and the number of recovered cases (y)
   2. The number of confirmed cases (x) and the number of monthly reported COVID-19-related deaths (y). Add this reference (https://www.sciencedirect.com/science/article/pii/S2590053620301087)

The same thing must be done to obtain Kruskal-Wallis and Dunn's tests. Please write the case fatality rate formula correctly as an equation and provide an equation number in all cases. Put all equations in the Methods section.

1. Simple models seem to have predicted the trend well in the near future; however, the data changed with the date. So, the regression model adjusted for the date should fit the data more robustly in the real situation. Suggesting the conclusion and recommendations using polynomial model-based results were drawn.

2. You may cite this reference to support your Results Section (https://www.mdpi.com/2071-1050/15/18/13647).

3. The conclusion section must be more precise. Be specific while writing your main findings.

4. References must be updated.

References
1. Suleiman AA, Suleiman A, Abdullahi UA, Suleiman SA: Estimation of the case fatality rate of COVID-19 epidemiological data in Nigeria using statistical regression analysis. *Biosaf Health*. 2021; 3 (1): 4-7 PubMed Abstract | Publisher Full Text
2. Salleh S, Suleiman A, Daud H, Othman M, et al.: Tropically Adapted Passive Building: A Descriptive-Analytical Approach Using Multiple Linear Regression and Probability Models to Predict Indoor Temperature. *Sustainability*. 2023; 15 (18). Publisher Full Text
3. Suleiman A, Othman M, Daud H, Abdullah M, et al.: Forecasting the Volatility of Real Residential Property Prices in Malaysia: A Comparison of Garch Models. *Real Estate Management and Valuation*. 2023; 31 (3): 20-31 Publisher Full Text

Is the work clearly and accurately presented and does it cite the current literature?
Partly

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Partly

If applicable, is the statistical analysis and its interpretation appropriate?
Partly

Are all the source data underlying the results available to ensure full reproducibility?
Yes
Are the conclusions drawn adequately supported by the results?
Yes

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Statistics

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 04 August 2021

https://doi.org/10.5256/f1000research.59066.r91211

© 2021 Kaabi Y. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

- Yahia Ali Kaabi
  Department of Medical Laboratory Technology, Faculty of Applied Medical Sciences, Jazan University, Jazan, Saudi Arabia
  The author provided all needed amendments successfully.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

**Competing Interests:** No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of
Yahia Ali Kaabi
Department of Medical Laboratory Technology, Faculty of Applied Medical Sciences, Jazan University, Jazan, Saudi Arabia

The manuscript titled “Simple statistical insights into the COVID-19 data of Saudi Arabia: figures prior to vaccination campaign” is a good attempt to brief the disease figures in Saudi Arabia by providing a statistical summary of death and recovery rates, as well as the number of confirmed cases throughout the past year before starting vaccination.

The article rationale is good and the author summarized the official figures of COVID-19 in Saudi Arabia that were published by the Ministry of Health in a simple way for the readership. This simple presentation of COVID-19 data is needed to understand the graph of the outbreak and what health authorities did to contain the outbreak. Availability of such figures is also needed to help health authorities anticipate their future needs and put plans for future actions and precautionary measures.

The introduction provided a sufficient background on Coronavirus and the study aim was clearly mentioned at the end of this section. However, I would suggest adding one paragraph to provide an account on the current situation worldwide and in Saudi Arabia.

The methods section was concise, but clear and data source and all statistical analysis were stated and justified by the author.

The results were also clearly presented; in particular the detailed comparison of death figures in each month was important when correlating to the number of new incidents/monthly. In addition, comparison between local figures and international figures highlighted the similarity between Saudi Arabia and many other international figures, which was important to show how far or close are we from international statistics. It would be good if the author could also explain how death rate/1000 is calculated below table 3, as done for case fatality rate. In Figure 2, the legend should be corrected to describe the recovered cases not deaths.

I would also suggest adding one paragraph to the discussion section to report on the figures after vaccination campaign till now, just to provide a simple comparison between the situations before
and after vaccination to assess the success of vaccination in containing the mortality effects of COVID-19 in the country.

Is the work clearly and accurately presented and does it cite the current literature?  
Yes

Is the study design appropriate and is the work technically sound?  
Yes

Are sufficient details of methods and analysis provided to allow replication by others?  
Yes

If applicable, is the statistical analysis and its interpretation appropriate?  
Yes

Are all the source data underlying the results available to ensure full reproducibility?  
Yes

Are the conclusions drawn adequately supported by the results?  
Partly

**Competing Interests:** No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

---

Author Response 27 Jul 2021  
**Omar S. El-Masry**

In response to the valuable comment on the introduction section, I think adding more information on the current situation in the introduction could also require a change in the paper's title, which mainly focuses on the figures prior to the vaccination campaign in the Kingdom. However, the current figures will be briefly discussed in the discussion section.

Thank you for your valuable comment on the results section, the equation for death rate/1000 was added below table 3 and Figure 2 legend was corrected as indicated, thank you.

The following paragraph was added to the discussion section as recommended; I focused on the case fatality rate, which reflects directly on the number of confirmed cases and the number of deaths, death rate /1000 was not calculated as the number of Saudi population in 2021 is different than that in 2020 therefore comparison is not significant.

In order to give a brief account on the figures post-vaccination campaign, the case fatality rates were calculated from the officially published figures in the first 6 months in 2021.
(January-June). The number of newly recorded cases escalated from ~5000 cases in January to ~37000 in June. This means that following flattening of the curve at the end of 2020, there was an ascending increase in the number of cases again after the start of the vaccination campaign. However, the number of cases is still much less than the numbers recorded in the same months in 2020 before vaccination. The case fatality rate in January 2021 was 0.03 (3%), whilst that calculated for February-June was 0.01 (1%), which is similar to that recorded for June 2020; however, there was a clear reduction in the number of confirmed cases and deaths in 2021. Whether the reduction in confirmed cases and deaths numbers is due to vaccination or other factors, this is not confirmed due to lack of information about the number of vaccinated people who contracted the infection nor those who died after being vaccinated. Therefore, we have to wait for the release of all official figures and information before we make a final conclusion about the efficacy of vaccination.

An updated version was submitted for your kind review.

Competing Interests: No competing interests were disclosed.