

## PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://bmjopen.bmj.com/site/about/resources/checklist.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

### ARTICLE DETAILS

<b>TITLE (PROVISIONAL)</b>	Factors associated with COVID-19 infections and mortality in Africa: A cross-sectional study using publicly available data
<b>AUTHORS</b>	Okeahalam, Charles; Williams, Victor; Otwombe, Kennedy

### VERSION 1 – REVIEW

<b>REVIEWER</b>	Elisa M Maffioli University of Michigan, US
<b>REVIEW RETURNED</b>	03-Aug-2020

<b>GENERAL COMMENTS</b>	<p>Referee Report</p> <p>COVID-19 and Mortality: A Statistical Analysis of African Countries</p> <p>The article describes the risk factors of COVID-19 cases and deaths for 53 African countries. The authors explore several risk factors to shed some light on the level of preparedness and capacity of health systems in the Africa continent. The authors conclude suggesting that there are some gaps in the health service delivery and more resources should be poured in for countries to be better able to face an epidemic.</p> <p>Major comments</p> <p>1. The overall goal of the study is important. However, in my opinion there is a disconnection between the general purpose of the study and how the authors decide to implement the analysis empirically. It is my understanding that the authors want to analyze the relationship between health indicators and cases/deaths in Africa in order to draw some conclusions on the capacity of African countries to handle the COVID-19 pandemic. In fact, a lot of the introduction and discussion present a discussion on how countries should prepare their health system for this epidemic. There are several ways in which we can answer this question and there are different ways in which we can measure “capacity” or “preparedness” of the health system, especially for an epidemic. Overall, it is unclear why the authors choose certain indicators among the 100 core health indicators. The authors should explain much more in details how they choose certain indicators, and why. More importantly, they should explain how the indicators chosen are important to answer their research questions.</p> <p>Some questions that would be helpful to address:</p> <ul style="list-style-type: none"><li>- How each of the 12 thematic areas correspond to the research questions?</li><li>- Are indicators from all 12 thematic areas chosen or not?</li></ul>
-------------------------	--

	<ul style="list-style-type: none"> <li>- Which indicators represent “preparedness” or “capacity”?</li> <li>- Among others, why insufficient physical activity is a risk factor for deaths? Similarly, why immunizations, breastfeeding, TB/HIV are good risk factors? Other indicators like personnel, service coverage represent good indicators for preparedness to face a health crisis. Similarly, prevalence of overweight could be related to other pre-existing health conditions which can contribute to the risk of dying because of Covid (diabetes, etc). However, there should be a detailed explanation about how the selection of the indicators has been conducted.</li> <li>- A table with the 12 thematic areas, the 100 indicators, the ones chosen and not used might help clarifying the conceptual framework the authors have in mind.</li> </ul> <p>In line with the comment above, the last sentences/paragraph in the introduction should be updated to reflect more information on the indicators chosen, how they map out from the research question, and what is the ultimate goal of the study. Please consider rephrase row 48-49.</p> <p>The authors mentioned in the data sources that all the indicators “directly or indirectly describe the ability of a country to respond to the health needs of the population.” However, it does mean that meeting the health needs of the population prepares the country for a pandemic. A more in depth description is needed.</p> <p>2. The article also lacks details on the analysis: - Which country is missing?</p> <ul style="list-style-type: none"> <li>- Cite a reference for the imputation technique used</li> <li>- For a reader unfamiliar to this: why 10 imputations? what are variance information measures? - Is there a need to assess the model fit with so many methods? No results are reported on the model fit. I would suggest to use one and explain why that measure was chosen or report the results from the model fit.</li> <li>- There is no explanation about why sub-Saharan Africa is separated out in the analysis. What are the reasons of this sub-sample analysis? Why some risk factors are considered for sub-Saharan Africa but not for all the African countries?</li> </ul> <p>3. The discussion is well carried out. However, the last paragraph in page 9 is very vague. In line with comment 1, why the identification of risk factors should help countries to increase the budget allocation? For example, if we find that breastfeeding is low, why a government should invest in breastfeeding to be prepared for COVID-19? The link between risk factors, the preparedness for epidemics and the allocation of resources efficiently is unclear. I do not believe the article can go beyond saying what risk factors are. There are no conclusions that can be drawn on how to improve allocative efficiency of national resources from this study.</p> <p>Minor comments</p> <ul style="list-style-type: none"> <li>- UHC index should be defined in the abstract. For a reader unfamiliar to it, you should be at least spelled out.</li> </ul>
--	--

	<ul style="list-style-type: none"> <li>- The last sentence in the abstract “these gaps impact negatively on service delivery, but ...” is not clear. Maybe there is a typo or consider change it.</li> <li>- Similarly, the article summary (strengths and limitations of the study) does not always contain all the information necessary to understand. My suggestion are: to specify the analysis method used in strength (i) and the significant factor in strength (iii) as well as the variance (of what) in limitation (ii).</li> <li>- Update COVID-19 related figures (row 10).</li> <li>- Reference should be “references” (page 10)</li> <li>- The link to references 7 and 8 is the same. Please adjust.</li> <li>- I would move rows 39-44 in the description of the indicators/data above, and drop or move in the statistical methodology the second paragraph.</li> <li>- Either use USA or spell it out in introduction</li> <li>- What is healthy life expectancy? How does it differ from life expectancy? Why did you use chose this indicator?</li> </ul>
--	--

<b>REVIEWER</b>	Dr Anjorin, A. AbdulAzeez Lagos State University, Department of Microbiology (Virology)
<b>REVIEW RETURNED</b>	07-Aug-2020

<b>GENERAL COMMENTS</b>	All suggestions have been included in the manuscript with the major point of re-writing the conclusion to capture the important variables discussed.
-------------------------	--

### VERSION 1 – AUTHOR RESPONSE

Reviewer 1	Responses	
1	Overall, it is unclear why the authors choose certain indicators among the 100 core health indicators. The authors should explain much more in details how they choose certain indicators, and why. More importantly, they should explain how the indicators chosen are important to answer their research questions.	We reviewed the 100 core indicators together with a clinician who advised on the more plausible ones to answer our hypothesis. The identified variables were subsequently subjected to various statistical approaches including testing for correlations and using variable selection regression procedures to arrive at the final list.
2	How each of the 12 thematic areas correspond to the research questions?	While we had 100 potential indicators, their selection into the model was not based on thematic areas per se. We had other thematic areas that we felt as a team were not plausible in answering our research hypothesis
3	Are indicators from all 12 thematic areas chosen or not?	See response in Q2 above
4	Which indicators represent “preparedness” or “capacity”?	Of the 100 indicators we initially identified, we did not classify them into “preparedness” or “capacity”. Our approach was driven by plausibility, knowledge about the field and statistical methodologies to arrive at the

		final variables.
5	Among others, why insufficient physical activity is a risk factor for deaths? Similarly, why immunizations, breastfeeding, TB/HIV are good risk factors? Other indicators like personnel, service coverage represent good indicators for preparedness to face a health crisis. Similarly, prevalence of overweight could be related to other pre-existing health conditions which can contribute to the risk of dying because of Covid (diabetes, etc). However, there should be a detailed explanation about how the selection of the indicators has been conducted.	This paragraph has been updated with additional references to read “More intuitively, we find that insufficient physical activity among adults aged 18+ years increased the risk of COVID-19 mortality. It is well documented in the literature that the lack of physical activity increases the risk of obesity. <sup>19</sup> Recent publications on COVID-19 related mortality, have shown that obesity elevates the risk of mortality. <sup>20 21</sup> Physical inactivity has long been recognized as a risk factor for non-communicable diseases which is a known cause of mortality globally (1, 2). <sup>22 23</sup> Early studies since the onset of the COVID-19 outbreak in China indicated elderly patients and those with co-morbidities particularly diabetes, hypertension and chronic respiratory diseases were at increased risk of mortality from COVID-19 (3, 4). <sup>24 25</sup> Therefore, the finding of insufficient physical activity as a risk factor for death from our study is consistent with earlier studies (1). <sup>23</sup> Similarly, immunization has proved effective in prevention of different types of infectious diseases globally while breastfeeding provides babies with essential nutrients and antibodies to help prevent infections early in life. Hence, establishment of expanded program on immunization (EPI) and programs to encourage breastfeeding globally. Thus we postulate that the population of Africans aged 18+ years who do not undertake sufficient physical activity may be similarly at high risk of mortality.
6	A table with the 12 thematic areas, the 100 indicators, the ones chosen and not used might help clarifying the conceptual framework the authors have in mind.	We have added a Table showing the 12 themes and the variables we considered.
7	In line with the comment above, the last sentences/paragraph in the introduction should be updated to reflect more information on the indicators chosen, how they map out from the research question, and what is the ultimate goal of the study. Please consider rephrase row 48-49.	We have included additional sentences in this paragraph which now reads: “While the indicators were drawn from different thematic areas, our analysis was largely data driven. The analysis is conducted for sub-Saharan Africa (SSA) and entire continent since SSA is known to have a higher burden of disease.”
8	The authors mentioned in the data sources that all the indicators “directly or indirectly describe the ability of a country to respond to the health needs of the population.” However, it does mean that meeting the health needs of the population	This sentence has been expanded to read “These indicators directly or indirectly describe the potential ability of a country’s health system to respond to the health needs of the population and may further determine the extent available services can be expanded to accommodate emergencies.”

	prepares the country for a pandemic. A more in depth description is needed.	
9	2. The article also lacks details on the analysis: - Which country is missing?	We have updated the last sentence in the data management section to read:  “The countries were further categorized into their assigned WHO region and World Bank income group except Somalia that had missing data. ”
10	Cite a reference for the imputation technique used	We have now added the seminal reference for multiple imputations i.e Little RJA, Rubin DB: Statistical analysis with missing data. 2002, John Wiley & Sons, Inc, Hoboken, New Jersey
11	- For a reader unfamiliar to this: why 10 imputations? what are variance information measures?	The multiple imputation seminal work suggests running 5 to 10 imputations since this reduces variability.  Variance information measures are used to assess the fit of the multiple imputations. We have retained the statement below that we had previously written to explain this:  “The fit of the multiple imputation was evaluated using variance information measures including relative efficiency. ”
12	- Is there a need to assess the model fit with so many methods? No results are reported on the model fit. I would suggest to use one and explain why that measure was chosen or report the results from the model fit.	The model fit assessment measures are all generated by default as designed by the developers of the method.  We have now included the following sentence for reporting the MI model fit: “In the multiple imputation estimations, the relative efficiency of variables where imputation was conducted ranged from 95% to 99%.”
13	- There is no explanation about why sub-Saharan Africa is separated out in the analysis. What are the reasons of this sub-sample analysis? Why some	Sub-Saharan African is generally known to have the highest burden of communicable and non-communicable diseases as well as high poverty levels. For this reason, we conducted the analysis including only sub-Saharan Africa as well as all African countries to determine if there would be a difference in their relationship with covid cases and mortality.
14	Why some risk factors are considered for sub-Saharan Africa but not for all the African countries?	The final model variables were purely driven by statistical analysis and varied between sub-Saharan Africa and all Africa.
15	The discussion is well carried out. However, the last paragraph in page 9 is very vague. In line with comment 1, why the identification of risk factors should help countries to increase the budget allocation? For example, if we find that breastfeeding is low, why a government should invest in breastfeeding to be prepared for COVID-	We do agree with the reviewer and further concur that our analysis cannot infer cause and effect for various reasons. The dataset is cross-sectional and was pooled together from various sources. We had no control of the original data collection process. However, there is some relationship shown by the tests of association that we conducted. It may be that the findings are spurious or provides a signal for further research.

	19? The link between risk factors, the preparedness for epidemics and the allocation of resources efficiently is unclear. I do not believe the article can go beyond saying what risk factors are. There are no conclusions that can be drawn on how to improve allocative efficiency of national resources from this study.	
16	UHC index should be defined in the abstract. For a reader unfamiliar to it, you should be at least spelled out.	We have amended based on the reviewer's suggestion to read: "In Sub-Saharan Africa, an increase in the number of nursing and midwifery personnel decreased the risk of COVID-19 deaths ( $p=0.0178$ ) while a unit increase in universal health care (UHC) index of service coverage and prevalence of insufficient physical activity among adults increased the risk of COVID-19 deaths ( $p=0.0432$ and $p=0.0127$ )."
17	The last sentence in the abstract "these gaps impact negatively on service delivery, but ..." is not clear. Maybe there is a typo or consider change it.	This has been reworded for clarity to read: "These gaps impact negatively on service delivery in Africa which requires more nursing personnel and increased UHC coverage to mitigate the effects of COVID-19."
18	Similarly, the article summary (strengths and limitations of the study) does not always contain all the information necessary to understand. My suggestion are: to specify the analysis method used in strength (i) and the significant factor in strength (iii) as well as the variance (of what) in limitation (ii).	This has been updated to read: Section on Strengths and limitation  <ul style="list-style-type: none"> <li>a. Strengths of the study: <ul style="list-style-type: none"> <li>i. Innovative use of credible publicly available data which can easily be accessed and verified</li> <li>ii. Use of multiple imputation technique to address missing observations in the different variables.</li> <li>iii. Use of robust negative binomial regression analysis method which allows for modeling of over-dispersion in the data.</li> </ul> </li> <li>b. Limitations of the study: <ul style="list-style-type: none"> <li>i. Extracted data had missing observations which necessitated excluding some variables.</li> <li>ii. The methods used to impute for missing data may have</li> </ul> </li> </ul>

		overestimated the variance
19	Update COVID-19 related figures (row 10).	The COVID-19 figures have been updated to reflect the most recent numbers as at 30 <sup>th</sup> August. The amended sentences now read “Since March 2020, there has been a rapid increase in the number of cases globally. Globally, as of 30 <sup>th</sup> August 2020, there have been 24,822,800 confirmed cases of COVID-19, including 838,360 deaths, reported to WHO. <sup>1 2</sup> In a bid to stop the virus from spreading, there has been a global restriction on several activities including travelling. <sup>3 4</sup> ”
20	Reference should be “references” (page 10)	This has been updated to “References”
21	The link to references 7 and 8 is the same. Please adjust.	We tested the links and can confirm they are different: Ref 7: <a href="https://apps.who.int/gho/data/node.home">https://apps.who.int/gho/data/node.home</a> Ref 8: <a href="https://www.who.int/healthinfo/indicators/2018/en/">https://www.who.int/healthinfo/indicators/2018/en/</a>
22	I would move rows 39-44 in the description of the indicators/data above, and drop or move in the statistical methodology the second paragraph.	These have now been moved in the statistical methodology section
23	Either use USA or spell it out in introduction	This has been amended to “United States of America (USA)”
24	What is healthy life expectancy? How does it differ from life expectancy? Why did you use chose this indicator?	We have now defined both Life Expectancy and Healthy Life Expectancy  Both variables were tested in the statistical model and the health life expectancy came out better in model performance compared with life expectancy.
	<b>Reviewer 2</b>	<b>Responses</b>
1	COVID-19 and mortality: A statistical analysis of African countries	A cross-sectional study of COVID-19 and mortality in African countries
2	The current COVID-19 pandemic is a global threat. This elicits questions on the level of preparedness and capacity of health systems to respond to emergencies. Relative to other parts of the world, Africa has poorly developed health systems with limited capacity to respond to health crises. Africa is particularly disadvantaged.	The current pandemic is a global threat. This elicits questions on the level of preparedness and capacity of health systems to respond to emergencies relative to other parts of the world.
3	In Sub-Saharan Africa, an increase in the number of nursing and midwifery	In Sub-Saharan Africa, an increase in the number of nursing and midwifery personnel decreased the risk of COVID-19 deaths (p=0.0178) while a unit increase in

	<p>personnel decreased the risk of COVID-19 deaths (<math>p=0.0178</math>) while a unit increase in UHC index of service coverage and prevalence of insufficient physical activity among adults increased the risk of COVID-19 deaths (<math>p=0.0432</math> and <math>p=0.0127</math>).</p>	<p>universal health care (UHC) index of service coverage and prevalence of insufficient physical activity among adults increased the risk of COVID-19 deaths (<math>p=0.0432</math> and <math>p=0.0127</math>).</p>
4	<p>Despite its limited resources, Africa's preparedness and response to the COVID-19 pandemic can be improved by identifying and addressing <i>specific</i> gaps in the funding of health services delivery. These gaps impact negatively on service delivery but appear to have received limited funding and policy priority.</p>	<p>Despite its limited resources, Africa's preparedness and response to the COVID-19 pandemic can be improved by identifying and addressing <i>specific</i> gaps in the funding of health services delivery. These gaps impact negatively on service delivery in Africa which requires more nursing personnel and increased UHC coverage to mitigate the effects of COVID-19. .</p>
5	<p>a. Strengths of the study:  i. Use of a robust statistical analysis method  ii. Identification of key evidence-based factors that might mitigate COVID-19 deaths and infections in Africa.  iii. Identification of a significant factor in the population which may offer protection against COVID-19.  iv. Innovative use of publicly available data which can easily be accessed and verified  b. Limitations of the study:  i. Some countries had missing data in some of the indicators analyzed.  ii. Use of multiple imputation technique which may have caused a slight overestimate of the variance</p>	<p>This has been updated to read:  Section on Strengths and limitation</p> <p>c. Strengths of the study:  i. Innovative use of credible publicly available data which can easily be accessed and verified  ii. Use of multiple imputation technique to address missing observations in the different variables.  iii. Use of robust negative binomial regression analysis method which allows for modeling of over-dispersion in the data.  d. Limitations of the study:  i. Extracted data had missing observations which necessitated excluding some variables.</p> <p>The methods used to impute for missing data may have overestimated the variance</p>
6	<p>Two months later the World Health Organisation (WHO) classified COVID-19 as a pandemic. COVID-19 is a</p>	<p>Two months later, the World Health Organisation (WHO) classified COVID-19 as a pandemic. COVID-19 is a global health threat.</p>



	global health threat.	
7	COVID-19 is a global health threat.	The sentence "COVID-19 is a global health threat." is now deleted.
8	In a bid to stop the virus spreading,	The sentence has now been amended to read "In a bid to stop the virus from spreading,...."
9	While this may seem an advantage, the WHO continues to express concern about the impact COVID may have on Africa.	The sentence has now been amended to read "While this may seem an advantage, the WHO continues to express concern about the impact COVID-19 may have on Africa."
10	And the easing of restrictions on social and economic activities is likely to lead to an increase in the number of cases of COVID-19 in African countries.	The sentence has been rewritten for clarity to read "The easing of restrictions on social and economic activities is likely to lead to a second wave of cases of COVID-19 in African countries."
11	It is the first paper of this type from the African continent.	This sentence has now been omitted.
12	Data for 32 indicators (or variables) from 12 thematic areas was extracted from the 2018 Global Reference List of 100 Core Health Indicators.	This has been amended to read "Data for 32 indicators (or variables) from 12 thematic areas were extracted from the 2018 Global Reference List of 100 Core Health Indicators."
13	Data were extracted in .xls format for each variable and imported into STATA 15.0 software (StataCorp LLC College Station, TX). For each variable, the most recent data for all countries included in the study was retained with the corresponding year and country name and saved in .dta format. The different variables were merged using the country name as the unique identifier to obtain the final data set used for the analysis.	This paragraph has been amended to read "Data were extracted in .xls format for each variable and imported into STATA 15.0 software (StataCorp LLC College Station, TX). For each variable, the most recent data for all countries included in the study were retained with the corresponding year and country name in .dta format. The different variables were merged using the country name as the unique identifier to obtain the final data set used for the analyses."
14	All data on health indicators were continuous and were analysed descriptively using medians, interquartile ranges (IQR) and the range of the measures were determined by the minimum and maximum values.	This sentence has been amended to read "All data on health indicators were continuous and were analysed descriptively using median, interquartile range (IQR) and minimum and maximum values."

15	Of the 53 countries included in the analysis,	This sentence has been amended to read "Of the 53 countries included in the analyses,...."
16	Of the 53 countries included in the analyses, there were varying proportions of missing data with the majority missing less than 10% of the data	This sentence has been amended to read "Of the 53 countries included in the analyses, there were varying proportions of < 10% missing data."
17	In the multivariate model. A full model was fitted first and then structurally reduced using the backward selection procedure to arrive at the final model.	This sentences have been amended to read "In the multivariate model, a full model including all the variables was fitted and the final model was determined using the backward selection procedure."
18	All statistical analysis was conducted using SAS Enterprise Guide 7.15 (SAS Institute Inc., NC, USA) using the procedures MEANS, GENMOD, MI, MIANALYZE.	This sentence has been amended to read "All statistical analyses were conducted using SAS Enterprise Guide 7.15 (SAS Institute Inc., NC, USA)."
19	In summary, our findings are that, controlling for BCG immunization coverage in sub-Saharan Africa, building health capacity by increasing the number of nursing and midwifery personnel will reduce COVID-19 deaths.	This sentence has been reworded to read "In summary, building health capacity by increasing the number of nursing and midwifery personnel will reduce COVID-19 deaths in sub-Saharan Africa."
20	However, we also find that higher UHC index of service coverage and prevalence of insufficient physical activity in adults $\geq 18$ years increases fatalities.	This sentence has been reworded to read "However, we also found that higher UHC index of service coverage and prevalence of insufficient physical activity in adults $\geq 18$ years increases fatalities."
21	This is not paradoxical but may point to the fact that the number of people who could contract the virus is augmented by longer life expectancy.	This sentence has been omitted
22	African governments will need to strengthen their overall health care systems and in doing so, specific focus needs to be placed on enhancing human resource capacity such as nurses, medical doctors and laboratory personnel.	This sentence has been amended for clarity to read "African governments need to strengthen the overall health care systems and in doing so, specific focus needs to be placed on enhancing human resource capacity such as nurses, medical doctors and laboratory personnel."
23	No major differences relative to the first	The sentence has been expanded to include the implication. It now reads "No major differences relative to

	regression were observed.	the first regression were observed suggesting that importation of infections similarly occurred in other sub-Saharan African countries.”
24	Additionally, from a statistical perspective, testing identifies cases (which might otherwise have been unobserved) and severe cases result in deaths - which are then more accurately recorded.	The sentence has been amended to read “From a statistical perspective, testing identifies cases which may be managed before getting severe resulting in deaths.”
25	how? Please refer to the evidence in the result	This paragraph has been updated with additional references to read “More intuitively, we find that insufficient physical activity among adults aged 18+ years increased the risk of COVID-19 mortality. It is well documented in the literature that the lack of physical activity increases the risk of obesity. <sup>19</sup> Recent publications on COVID-19 related mortality, have shown that obesity elevates the risk of mortality. <sup>20 21</sup> Physical inactivity has long been recognized as a risk factor for non-communicable diseases which is a known cause of mortality globally (1, 2). <sup>22 23</sup> Early studies since the onset of the COVID-19 outbreak in China indicated elderly patients and those with co-morbidities particularly diabetes, hypertension and chronic respiratory diseases were at increased risk of mortality from COVID-19 (3, 4). <sup>24 25</sup> Therefore, the finding of insufficient physical activity as a risk factor for death from our study is consistent with earlier studies (1). <sup>23</sup> Similarly, immunization has proved effective in prevention of different types of infectious diseases globally while breastfeeding provides babies with essential nutrients and antibodies to help prevent infections early in life. Hence, establishment of expanded program on immunization (EPI) and programs to encourage breastfeeding globally. Thus we postulate that the population of Africans aged 18+ years who do not undertake sufficient physical activity may be similarly at high risk of mortality.
26	from around the world that have shown older people at higher risk for COVID-19 infection.	This sentence has been updated to read: “...from around the world that have shown older people are at higher risk for COVID-19 infection.”
27	Please conclude based on the overall summary of the key health system indicators for African countries highlighted in both the results and your discussion so as to justify the study aim. You may choose to include the financial burden based on the CHE of GDP estimated analyses described in Table 1	The conclusion has been updated to read “In conclusion, our study findings showed a relationship between COVID-19 cases and deaths with health capacity, breast feeding, life expectancy (as a proxy for age) and healthcare funding. Timely identification of the key evidence-based factors that might mitigate COVID-19 infections and deaths in Africa is pertinent for better management of the current and future pandemics. This may include investing in healthcare capacity building, infrastructure, disease surveillance, public health laboratories and all other aspects that relate to health as elucidated in the WHO International Health Regulations. <sup>28</sup>

## VERSION 2 – REVIEW

<b>REVIEWER</b>	Elisa Maria Maffioli University of Michigan, USA
<b>REVIEW RETURNED</b>	25-Sep-2020

<b>GENERAL COMMENTS</b>	<p>The authors fully address all the comments. I just have one additional request related to their points 2, 7 and 14. It is still unclear to me why some indicators are chosen and similarly why they explore some risk factors for SSA and others for full Africa.</p> <p>In their answers they mentioned: our analysis was largely data-driven; Our approach was driven by plausibility, knowledge about the field and statistical methodologies to arrive at the final variables; We had other thematic areas that we felt as a team were not plausible in answering our research hypothesis.... I think these points should be discussed in the manuscript. More specifically, the choice of indicators to be analyzed in an empirical model are not data-driven. The researcher chose to include or not certain variables. So, can the authors explain their criteria in choosing the indicators? Which is their process of selection (influenced by experience, etc)? If they did not follow thematic areas, how did they choose? I think this would be beneficial to fully grasp the research. Otherwise it seems the authors simply choose (almost random) indicators or decide to present the ones where they found results.</p>
-------------------------	--

<b>REVIEWER</b>	Dr AbdulAzeez A. Anjorin Lagos State University, Nigeria
<b>REVIEW RETURNED</b>	08-Sep-2020

<b>GENERAL COMMENTS</b>	Pg. 8 lines 53 and 56: (1, 2) and (3, 4), are these references or what? Please delete (1, 2) and (3, 4) respectively to conform with the journal's referencing format.
-------------------------	--

## VERSION 2 – AUTHOR RESPONSE

### Response to reviewers' comments

<p>Please leave your comments for the authors below Pg. 8 lines 53 and 56: (1, 2) and (3, 4), are these references or what? Please delete (1, 2) and (3, 4) respectively to conform with the journal's referencing format.</p>	<p>We thank the reviewer for picking this error. The bracketed references have now been deleted and the sentence now reads</p> <p>“Physical inactivity has long been recognized as a risk factor for non-communicable diseases which is a known cause of mortality globally.<sup>22 23</sup> Early studies since the onset of the COVID-19 outbreak in China indicated elderly patients and those with co-morbidities particularly diabetes, hypertension and chronic respiratory diseases were at increased risk of mortality from COVID-19.<sup>24 25</sup>”</p>
<p>Please leave your comments for the authors below The authors fully address all the comments. I just have one additional request related to their points 2, 7 and 14. It is still unclear to me why some indicators are chosen and similarly why they explore some risk factors for SSA and others for</p>	<p>On why we conducted analysis for SSA and all of Africa, we have amended the last paragraph of the introduction section to read</p> <p>“We conducted analysis for sub-Saharan African (SSA) countries - the standard multilateral</p>

<p>full Africa.</p> <p>In their answers they mentioned: our analysis was largely data-driven; Our approach was driven by plausibility, knowledge about the field and statistical methodologies to arrive at the final variables; We had other thematic areas that we felt as a team were not plausible in answering our research hypothesis....</p> <p>I think these points should be discussed in the manuscript. More specifically, the choice of indicators to be analyzed in an empirical model are not data-driven. The researcher chose to include or not certain variables. So, can the authors explain their criteria in choosing the indicators? Which is their process of selection (influenced by experience, etc)? If they did not follow thematic areas, how did they choose? I think this would be beneficial to fully grasp the research. Otherwise it seems the authors simply choose (almost random) indicators or decide to present the ones where they found results.</p>	<p>institutions data classification of Africa and also for all the countries on the continent including those north of the Sahara, usually classified with the Middle East region. We present our analysis by SSA countries due to their higher burden of disease and all of Africa.”</p> <p>In the statistical methodology section, we have reworded our explanation on how we arrived at the final variables from the larger list. The amendment now reads:</p> <p>“The process of selection of variables for analysis was as follows. Firstly, the team reviewed all the core publicly available health indicators. Then the plausibility of the explanatory power of these variables in the context of this study was subjected to various statistical approaches. These include the use of univariate and multivariate regression selection procedures. This approach enabled the identification of the final variables.”</p>
--	--

**VERSION 3 – REVIEW**

<b>REVIEWER</b>	Elisa Maria Maffioli University of Michigan
<b>REVIEW RETURNED</b>	08-Oct-2020
<b>GENERAL COMMENTS</b>	No further requests.