Factors associated with COVID-19 infections and mortality in Africa: a cross-sectional study using publicly available data

Charles Okeahalam,1 Victor Williams,2 Kennedy Otwombe 2,3

ABSTRACT
Introduction 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.

Methods This cross-sectional study uses publicly available core health data for 53 African countries to determine risk factors for cumulative COVID-19 deaths and cases per million in all countries in the continent. Descriptive statistics were determined for the indicators, and a negative binomial regression was used for modelling the risk factors.

Results 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 healthcare (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). An increase in the proportion of infants initiating breast feeding reduced the number of cases per million (p<0.0001), while an increase in higher healthy life expectancy at birth increased the number of cases per million (p=0.0340).

Conclusion Despite its limited resources, Africa’s preparedness and response to the COVID-19 pandemic can be improved by identifying and addressing specific 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.

INTRODUCTION
In January 2020, the SARS-CoV-2 popularly referred to as COVID-19 was identified as the cause of unknown pneumonia in Wuhan, China. Two months later, the WHO classified COVID-19 as a pandemic. COVID-19 is a global health threat. Since March 2020, there has been a rapid increase in the number of cases globally. Globally, as of 30 August 2020, there have been 24,822,800 confirmed cases of COVID-19, including 838,360 deaths, reported to WHO. In a bid to stop the virus from spreading, there has been a global restriction on several activities including travelling. The economic impact of a global shutdown and pressure on frail health systems of many countries across the world presents great uncertainty for the remaining half of the year and the immediate future.

The USA, Brazil, the UK and countries in Western Europe notably Italy, Spain and France account for the highest number of COVID-19 cases and mortality. With the possible exception of Brazil, these countries have advanced health systems, reliable epidemic surveillance and response systems and have been able to mobilise resources with which to respond. However, due to the sudden increase in demand for health personnel, test kits, emergency care equipment and personal protective equipment, the health systems of all affected countries have been put under significant strain. With varying efficacy, the affected countries have increased hospital facilities for critical cases, increased the daily number of tests to identify positive cases and increased provision of personal protective equipment for health personnel.

At present, in comparison with the USA and Europe, Africa has a lower number of cases and lower daily increase in infection. While this may seem an advantage, the WHO continues to express concern about the

Strengths and limitations of this study

- Innovative use of credible publicly available data that can easily be accessed and verified.
- Use of robust negative binomial regression analysis method that allows for modelling of underdispersion in the data.
- Extracted data had missing observations that necessitated excluding some variables.
- The methods used to impune for missing data may have overestimated the variance.

To cite: Okeahalam C, Williams V, Otwombe K. Factors associated with COVID-19 infections and mortality in Africa: a cross-sectional study using publicly available data. BMJ Open 2020;10:e042750. doi:10.1136/bmjopen-2020-042750

Prepublication history for this paper is available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2020-042750).

Received 18 July 2020
Revised 07 October 2020
Accepted 12 October 2020
impact COVID-19 may have on Africa. This is because, from the perspective of capacity, African countries are in a parlous situation relative to Europe, North America and some parts of Asia. The health systems of many African countries have limited ability to roll out widespread community screening and testing, few health personnel and limited emergency facilities to manage critically ill patients with COVID-19. 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. So with the best will and effort, African countries cannot use the same methodology to respond and contain the pandemic to the level of the developed countries that have performed badly, let alone to that of those that have performed well. If they are to use their limited capacity efficiently, it is imperative African countries clearly identify factors that will support that objective and institute effective evidence-based approaches to managing COVID-19. Accordingly, in a bid to support focus on efficient capacity development, this study analyses the relationship between health indicators and COVID-19 deaths and cases in African countries. While the indicators were drawn from different thematic areas, our analysis was largely data driven. We conducted analysis for sub-Saharan African (SSA) countries—the standard multilateral 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.

**METHODOLOGY**

This is a cross-sectional study of the most recent 2020 data for African countries extracted from the WHO Global Health Observatory Repository. Before extraction, the research team reviewed available indicators in the 2018 Global Reference List of 100 Core Health Indicators (plus health-related Sustainable Development Goals (SDGs)) and listed different indicators by thematic areas. 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. Data on confirmed cases of coronavirus and deaths were obtained from the Worldometer Coronavirus Live Update.

**Definition of variables**

Bacille Calmette-Guérin (BCG) immunisation coverage among 1 year olds (%): BCG immunisation coverage among 1 year olds (%).

Nursing and midwifery personnel (per 10 000 population): it is the density of nurses and midwifery personnel per 10 000 people.

UHC index of service coverage: coverage of essential health services such as reproductive, maternal, newborn and child health among others.

Prevalence of insufficient physical activity among adults aged 18+ years: insufficient physical activity was defined as adults not meeting the WHO recommendations on physical activity for health, that is, at least 150 min of moderate intensity or 75 min of vigorous intensity physical activity per week, or any equivalent combination of the two.

Early initiation of breast feeding (%): initiation of breast feeding within the first hour of birth and exclusively breast fed for the first 6 months of life.

Healthy life expectancy at birth (years): this is a life expectancy estimate that applies disability weights to health states to compute the equivalent number of years of good health that a new born can expect.

Life expectancy at birth: this reflects the overall mortality level of a population. It summarises the mortality pattern that prevails across all age groups—children and adolescents, adults and the elderly.

Prevalence of overweight among adults: adults with a body mass index ≥30.

Current health expenditure (CHE) as a percentage of gross domestic product (GDP): this indicates the level of resources channelled to health relative to other uses.

**Statistical methodology**

Data for 32 indicators (or variables) from 12 thematic areas were extracted from the 2018 Global Reference List of 100 Core Health Indicators (table 1). The 12 thematic areas are mortality by age and sex, mortality by cause, morbidity, nutrition, environmental risk factors, non-communicable diseases, immunisation, essential health services, utilisation and access, health workforce, health information and health financing.

Data were extracted in .xls format for each variable and imported into STATA V.15.0 software. 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 analysis. The countries were further categorised into their assigned WHO region and World Bank income group except Somalia that had missing data.

All data on health indicators were continuous and were analysed descriptively using median, IQR and minimum and maximum values.

Of the 53 countries included in the analyses, there were varying proportions of <10% missing data. To address this, we assumed a missing at random mechanism and applied a multiple imputation technique with 10 imputations and summarised the results across all the datasets. The fit of the multiple imputation was evaluated using variance information measures including relative efficiency.

The process of selection of variables for analysis was as follows. First, 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.

Due to its flexibility in allowing for overdispersion, risk factors for cumulative COVID-19 deaths and cases per million were fitted using the negative binomial regression. Both univariate and multivariate regression models were fitted. In the multivariate model, a full model including all the variables was fitted and the final model determined using the backward selection procedure. Regression models were fitted for SSA followed by a sensitivity analyses including all the countries in the continent. Model fit was assessed using the ratio of the deviance, scaled deviance, Pearson $\chi^2$ and scaled Pearson $\chi^2$ divided by the df. Additionally, we also assessed model fit using the cumulative sum of residual plots with 10 000 replications.Deaths and

| Table 1  | Summary of thematic areas of health indicators |
|---------|-----------------------------------------------|
| Thematic area | Indicators used |
|---------|------------------------------------------------|
| Mortality by age and sex | Life expectancy and healthy life expectancy at birth (years). Adult mortality rate between 15 and 60 years of age (adult mortality rate (probability of dying between 15 and 60 years per 1000 population)). Under-five mortality rate (per 1000 live births). Infant mortality rate (per 1000 live births). Neonatal mortality rate (per 1000 live births). |
| Mortality by cause | Maternal mortality ratio (maternal mortality ratio (per 100 000 live births)). TB mortality rate (per 100 000 population). AIDS-related mortality rate. Mortality from unsafe water, unsafe sanitation and lack of hygiene. |
| Morbidity | HIV prevalence rate (per 1000 population). HIV incidence rate (per 1000 population). TB incidence rate (per 100 000 population). |
| Nutrition | Exclusive breastfeeding rate 0–5 months of age (%). Early initiation of breastfeeding (%). |
| Environmental risk factors | Population using safely managed drinking water services (%). Population using safely managed sanitation services (also: population with handwashing facility with soap and water) (%). Air pollution level in cities ($\mu g/m^3$). |
| Non-communicable diseases | Tobacco use among persons aged 15+ years (SDG 3.a.1) (also: adolescents) (%). Raised blood pressure among adults (18+ years) (%). Overweight and obesity in adults (also: school-age children and adolescents) (%). Raised blood glucose/diabetes among adults (%). Insufficient physical activity in adults (also: adolescents) (%). |
| Immunisation | Immunisation coverage rate by vaccine for each vaccine in the national schedule (%). |
| Essential health services | Coverage of essential health services (%). |
| Utilisation and access | Health facility density and distribution (also: access to emergency surgery) (per 10 000 population). Hospital bed density (per 10 000 population). Access to a core set of relevant essential medicines (%). |
| Health workforce | Health worker density and distribution (per 10 000 population). |
| Health information* | Completeness of reporting by facilities (also: completeness and timeliness for notifiable diseases). |
| Health financing | Total current expenditure on health as % of gross domestic product (also: total capital expenditure on health as % of current+capital expenditure on health). Public domestic sources of current spending on health as % of current health expenditure. External source of current spending on health (% of current expenditure on health). |

*Data not available.
SDG, Sustainable Development Goals; TB, tuberculosis.
cases per million were those reported in the Worldometer as of 29 May 2020.

All statistical analyses were conducted using SAS Enterprise Guide V.7.15.

Patient and public involvement
This study used publicly available health indicators and aggregated COVID-19 cases and deaths. No patients were involved.

RESULTS
Characteristics
A total of 53 countries from Africa were included in this analysis (table 2). The median number of cumulative COVID-19 deaths across the African continent was 12 (IQR: 3–49), cumulative deaths per million (2, IQR: 0.2–6.0), cumulative COVID-19 cases (820, IQR: 295–2216), cumulative cases per million (63, IQR: 17.0–243.0), CHE was 5.3 (IQR: 3.9–6.9) and BCG immunisation coverage in 1 year olds was 91.5% (IQR: 84.0–96.5). The median life expectancy was 63.2 years (IQR: 59.8–66.1), healthy life expectancy was 55.7 years (IQR: 52.3–57.9), medical doctors per 10 000 population was 1.6 (IQR: 0.8–4.0), nursing and midwifery personnel per 10 000 was 8.8 (IQR: 4.4–15.5) and universal healthcare (UHC) index of service coverage 0.5 (IQR: 0.4–0.5).

Risk factors for SSA
Cumulative COVID-19 deaths
In the multivariate regression, modelling risk factors for COVID-19 deaths in SSA and controlling for BCG immunisation coverage, a unit increase in the number of nursing and midwifery personnel decreased the risk of death by 0.0426 (p=0.0178), whereas a unit increase in the UHC index of service coverage and prevalence of insufficient physical activity among adults aged 18+ years increased the risk of COVID-19 deaths by 4.7049 (p=0.0432) and 0.0830 (p=0.0127), respectively (table 3).

Cumulative cases per million
A unit increase in the proportion of infants initiating breast feeding reduces the number of cases per million by 0.0563 (p<0.0001), whereas an increase in higher healthy life expectancy at birth increases the number of COVID-19 cases per million by 0.0417 (p=0.0340).

Risk factors for all the African countries
Cumulative COVID-19 deaths
Early initiation of breast feeding (beta=−0.0514, p value=0.0027) was associated with a lower risk of death, whereas increasing healthy life expectancy at birth was associated with a higher risk of death (beta=0.1059, p value=0.0285).
Cumulative cases per million
An increase in the current health expenditure as a percentage of GDP (beta=−0.1739, p value=0.0397) and the percentage initiated early on breast feeding (beta=−0.0460, p value=0.0061) was associated with a decrease in cases of COVID-19 per million (table 3).

In the multiple imputation estimations, the relative efficiency of variables where imputation was conducted ranged from 95% to 99%, whereas the scaled deviance values for the regression models were within the acceptable range.

DISCUSSION
This study considers the factors associated with COVID-19 deaths and infection cases per million from 53 of 54 African countries. In summary, building health capacity by increasing the number of nursing and midwifery personnel will reduce COVID-19 deaths in SSA. However, we also found that higher UHC index of service coverage and prevalence of insufficient physical activity in adults ≥18 years increases fatalities. Additionally, having a population with a high proportion of people where breast feeding had been initiated early provided protection. A higher healthy life expectancy increased the risk of cases. When evaluating all the African countries, the initiation of early breast feeding was protective against death, whereas higher healthy life expectancy increased the risk of death. We also find that an increase in current health expenditure levels provide additional protection against COVID-19 infections.

Delving deeper into our results, the finding that an increase in the number of nursing and midwifery personnel reduced the risk of COVID-19 related mortality in SSA echo findings from other parts of the world where the benefits of preparedness and available capacity have been shown to confer benefits in the management and care of patients.16 17 Where capacity shortages exist in the health sector, there is potential for poor outcomes in the management and care of patients with COVID-19. Major gaps still exist in the implementation of the WHO International Health Regulations, notably increasing capacity of hospitals and appropriate injection of financial resources.17 The results confirm the point made earlier that unlike the other COVID-19 affected parts of the world, Africa has less capacity to handle a major outbreak of COVID-19. Therefore, African governments need to strengthen the overall healthcare systems, and in doing so, specific focus needs to be placed on enhancing human resource capacity such as nurses, medical doctors and laboratory personnel.

Our findings also illustrate that a higher UHC index of service coverage increases the risk of fatalities. While counterintuitive, this may be driven by the countries in SSA with the largest economies that have also had the largest number of cases and mortality: South Africa and Nigeria. A large number of people from these two

### Table 3 Risk factors for COVID-19 deaths and cases per million in Africa

| Variables                                      | Univariate |          | Multivariate |          |
|------------------------------------------------|------------|----------|--------------|----------|
|                                                | Beta (SE)  | P value  | Beta (SE)    | P value  |
| **(A) Sub-Saharan Africa**                     |            |          |              |          |
| Risk factors for deaths                         |            |          |              |          |
| BCG immunisation coverage among 1 year olds (%) | −0.0302 (0.0189) | 0.1106   | −0.0293 (0.0191) | 0.1242   |
| Nursing and midwifery personnel (per 10 000 population) | −0.0324 (0.0343) | 0.3467   | −0.0426 (0.0178) | 0.0171   |
| UHC index of service coverage (SCI)            | 4.8486 (1.7709) | 0.0062   | 4.7049 (2.3268) | 0.0432   |
| Prevalence of insufficient physical activity among adults aged 18+ years | 0.1201 (0.0243) | <0.0001  | 0.0830 (0.0333) | 0.0127   |
| Risk factors for cases per million              |            |          |              |          |
| Early initiation of breast feeding (%)          | −0.0534 (0.0143) | 0.0002   | −0.0563 (0.0136) | <0.0001  |
| Healthy life expectancy (HALE) at birth (years) | 0.0917 (0.0410) | 0.0251   | 0.0870 (0.0415) | 0.0373   |
| Prevalence of overweight among adults           | 0.0958 (0.0390) | 0.0140   | 0.0417 (0.0340) | 0.2214   |
| **(B) All the African countries**               |            |          |              |          |
| Risk factors for deaths (all countries)         |            |          |              |          |
| Early initiation of breast feeding (%)          | −0.0437 (0.0183) | 0.0205   | −0.0514 (0.0171) | 0.0027   |
| HALE at birth (years)                           | 0.0865 (0.0465) | 0.0626   | 0.1059 (0.0483) | 0.0285   |
| Risk factors for cases per million (all countries) |          |          |              |          |
| Current health expenditure as percentage of gross domestic product | −0.1913 (0.0896) | 0.0328   | −0.1739 (0.0845) | 0.0397   |
| Early initiation of breast feeding (%)          | −0.0476 (0.0169) | 0.0049   | −0.0460 (0.0167) | 0.0061   |

BCG, Bacille Calmette-Guérin; UHC, universal healthcare.
countries travel internationally, in particular to Europe, and it is likely that they have contracted the disease there and have imported it back into their countries.\(^3\) To assess the robustness of this finding, we conducted a sensitivity analysis by removing the data on South Africa and Nigeria and re-estimated the model. No major differences relative to the first regression were observed suggesting that importation of infections similarly occurred in other SSA countries.

Additionally, that cohort of the population that travels internationally is generally able to afford COVID-19 test kits and conduct more tests. Testing is a critical requirement for effective management of COVID-19. From a statistical perspective, testing identifies cases that may be managed before getting severe resulting in deaths. This increases the number of observations on deaths. This view is similar to that reported in a previous paper on the vulnerability of African countries to COVID-19 and their preparedness to mitigate.\(^18\)

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.\(^19\) Recent publications on COVID-19 related mortality have shown that obesity elevates the risk of mortality.\(^20\) Physical inactivity has long been recognised as a risk factor for non-communicable diseases, which is a known cause of mortality globally.\(^22\) Early studies since the onset of the COVID-19 outbreak in China indicated elderly patients and those with comorbidities particularly diabetes, hypertension and chronic respiratory diseases were at increased risk of mortality from COVID-19.\(^24\) Therefore, the finding of insufficient physical activity as a risk factor for death from our study is consistent with earlier studies.\(^1\) Similarly, immunisation has proved effective in prevention of different types of infectious diseases globally while breast feeding provides babies with essential nutrients and antibodies to help prevent infections early in life. Hence, establishment of expanded programme on immunisation and programmes to encourage breast feeding 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.

We also found that where a large population is exposed to early initiation of breast feeding, there was protection against COVID-19 infection. Research on the long-term benefits of breast feeding is growing and recent evidence suggests a protective effect against some chronic diseases in adulthood.\(^26\) It may be that this protective effect extends to non-severe cases of COVID-19.

Furthermore, in SSA and all the African countries combined, higher healthy life expectancy (life expectancy that accounts for disabilities) in this study was associated with a higher risk of COVID-19 infection. Our findings are in tandem with previously published work on risk factors for COVID-19 infection from around the world that have shown older people are at higher risk for COVID-19 infection. Being older is associated with lower immunity and inflammatory reactions and a higher risk of comorbidities such as diabetes and hypertension among others. Such factors predispose older cohorts of the population to COVID-19 infection.

This study is not without limitations. The data used were sourced from publicly available repositories and therefore we had no control in sampling, study design and data collection processes. The study we conducted is cross-sectional and therefore cannot deduce causality. Since COVID-19 statistics are updated on a daily basis, the findings we report may vary with updated data. Despite these challenges, we applied robust statistical analysis methods to alleviate potential biases.

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.\(^28\)

**Contributors** CO developed the initial research concept, developing the hypothesis and methodology. WV did the data extraction and merging of the different data sets. KO participated in developing the hypothesis and conducted the data analysis. All the authors participated in interpreting the results and writing the manuscript.

**Funding** This work was funded by Africa Growth Holdings Group (AGH) GROUP South Africa – Quantitative Analysis of the Impact of COVID-19 in Africa – (Grant number: 002/20).

**Competing interests** None declared.

**Patient consent for publication** Not required.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data are available in a public, open access repository. Data used for analysis are available in the dryad data repository reference. DOI: 10.5061/dryad.cnp59cp2r.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

**ORCID iD**

Kennedy Otswombe http://orcid.org/0000-0002-7433-4383

**REFERENCES**

1 Huang R, Xia J, Chen Y, et al. A family cluster of SARS-CoV-2 infection involving 11 patients in Nanjing, China. *Lancet Infect Dis* 2020;20:534–5.

2 World Health Organisation. Coronavirus disease (COVID-19) dashboard Geneva: WHO, 2020. Available: https://covid19.who.int/

3 World Health Organisation. Novel coronavirus (2019-nCoV) SITUATION REPORT - 1. 21 January 2020 Geneva: WHO, 2020. Available: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200121-sitrep-1-2019-ncov.pdf?sfvrsn=20a99c10_4

4 WHO. Critical preparedness, readiness and response actions for COVID-19: interim guidance, 22 March 2020. Geneva: WHO, 2020.
Remuzzi A, Remuzzi G. COVID-19 and Italy: what next? The Lancet 2020;395:1255–60.
10 Velavan TP, Meyer CG. The COVID-19 epidemic. Trop Med Int Health 2020;25:278–80.
11 World Health Organisation. Global health observatory data repository Geneva: WHO, 2020. Available: https://apps.who.int/gho/data/node.home
12 WHO. Global reference list of 100 core health indicators [plus health-related SDGs], Geneva: WHO; 2018, 2020. Available: https://www.who.int/healthinfo/indicators/2018/en/
13 Worldometer. Covid-19 coronavirus pandemic: Worldometer, 2020. Available: https://www.worldometers.info/coronavirus/
14 WHO. Global reference list of 100 core health indicators, 2018. Available: https://www.who.int/healthinfo/indicators/2018/en/
15 WHO. Global recommendations on physical activity for health, 2010. Available: https://apps.who.int/iris/bitstream/handle/10665/44399/9789241599979_eng.pdf?sequence=1
16 UNICEF W. Capture the moment – early initiation of breastfeeding: the best start for every newborn, 2018. Available: https://www.unicef.org/publications/files/UNICEF_WHO_Capture_the_moment_EIBF_.2018.pdf
17 WHO. Healthy life expectancy (HALE) at birth, 2020. Available: https://www.who.int/gho/health_burden_disease/life_tables/hale_text/en/
18 WHO. Definitions and metadata, 2006. Available: https://www.who.int/whosis/whostat2006DefinitionsAndMetadata.pdf
19 Little RJA, Rubin DB. Statistical analysis with missing data. Hoboken, New Jersey: John Wiley & Sons Inc., 2002.
20 Nkengasong JN, Mankoula W. Looming threat of COVID-19 infection in Africa: act collectively, and fast. Lancet 2020;395:841–2.
21 Paintsil E. COVID-19 threatens health systems in sub-Saharan Africa: the eye of the crocodile. J Clin Invest 2020;130:2741–4.