Am. J. Trop. Med. Hyg., 104(5), 2021, pp. 1694–1702
doi:10.4269/ajtmh.20-1502

Seeding COVID-19 across Sub-Saharan Africa: An Analysis of Reported Importation Events across 49 Countries

Laura A. Skrip,[1] Prashanth Selvaraj,[1] Brittany Hagedorn,[1] Andre Lin Ouédraogo,[1] Navideh Noori,[1] Amanda Orcutt,[1] Dina Mistry,[1] Jamie Bedson,[2] Laurent Hébert-Dufresne,[3,4] Samuel V. Scarpino,[5,6,7,8,9] and Benjamin M. Althouse[1,10,11]*

[1]*Institute for Disease Modeling, Global Health, Bill & Melinda Gates Foundation, Seattle, Washington;* [2]*Consultant to the Bill & Melinda Gates Foundation, Seattle, Washington;* [3]*Department of Computer Science, Vermont Complex Systems Center, University of Vermont, Burlington, Vermont;* [4]*Department of Computer Science, University of Vermont, Burlington, Vermont;* [5]*Network Science Institute, Northeastern University, Boston, Massachusetts;* [6]*Department of Marine and Environmental Sciences, Northeastern University, Boston, Massachusetts;* [7]*Department of Physics, Northeastern University, Boston, Massachusetts;* [8]*Department of Health Sciences, Northeastern University, Boston, Massachusetts;* [9]*ISI Foundation, Turin, Italy;* [10]*University of Washington, Seattle, Washington;* [11]*New Mexico State University, Las Cruces, New Mexico*

Abstract. The first case of COVID-19 in sub-Saharan Africa (SSA) was reported by Nigeria on February 27, 2020. Whereas case counts in the entire region remain considerably less than those being reported by individual countries in Europe, Asia, and the Americas, variation in preparedness and response capacity as well as in data availability has raised concerns about undetected transmission events in the SSA region. To capture epidemiological details related to early transmission events into and within countries, a line list was developed from publicly available data on institutional websites, situation reports, press releases, and social media accounts. The availability of indicators—gender, age, travel history, date of arrival in country, reporting date of confirmation, and how detected—for each imported case was assessed. We evaluated the relationship between the time to first reported importation and the Global Health Security Index (GHSI) overall score; 13,201 confirmed cases of COVID-19 were reported by 48 countries in SSA during the 54 days following the first known introduction to the region. Of the 2,516 cases for which travel history information was publicly available, 1,129 (44.9%) were considered importation events. Imported cases tended to be male (65.0%), with a median age of 41.0 years (range: 6 weeks–88 years; IQR: 31–54 years). A country's time to report its first importation was not related to the GHSI overall score, after controlling for air traffic. Countries in SSA generally reported with less publicly available detail over time and tended to have greater information on imported than local cases.

INTRODUCTION

The SARS-CoV-2 has spread feverishly across the globe, causing hundreds of thousands of COVID-19 cases and tens of thousands of deaths as of late March 2020.[1] Although detection of cases has centered on Asia, Europe, and the Americas, there so far seems to be a relative paucity of cases across the continent of Africa. This observation during the initial months of the pandemic occurred in the context of regular air traffic in and out of the continent at a time when Africa's share of air passenger traffic, although smaller than that of other regions, is on the rise (2.2% in 2017[2] compared with 1.0% in 2010[3]), in part due to strong economic and development ties with China.[4] This lack of cases may be due to inadequate testing capacity.[5] It is critical to understand how SARS-CoV-2 is introduced into countries[6,7] to anticipate onward transmission of the virus and accurately project infection rates in the overall population, as well as among population subgroups that are most vulnerable to severe morbidity and mortality.[8]

Early in the pandemic, Africa was expected to be uniquely positioned to have the most severe and under-detected outcomes related to COVID-19 infection.[9] The continent's countries are among those most at risk of widespread disease threats, per several indices of epidemic preparedness.[7,10–12] The State Party self-assessment annual reporting (SPAR) database of the WHO assigns scores to countries to assess capacities needed to detect, assess, notify, report, and respond to public health risk and acute events of domestic and international concern.[7,11] Similarly, the Global Health Security Index (GHSI) by Johns Hopkins University [10] uses a variety of healthcare, economic, demographic, and political factors to assess the vulnerability of a country to prevent or contain an infectious disease outbreak in general, whereas the Surgo Foundation's COVID-19 Community Vulnerability Index (CCVI) quantifies vulnerability on the basis of COVID-specific parameters.[12]

Using such indices, recent work has shown most of sub-Saharan Africa (SSA) to be at risk of COVID-19 importation and at reduced capacity to contain outbreaks due to lack of economic and medical resources.[7] Although it appears that the age-groups at the highest risk of severe COVID-19 disease and death (those > 60 years)[8,13] may be proportionately less in many SSA countries than in other parts of the world, the populations in many of these countries are at increased risk of having untreated chronic conditions due to weak health systems.[14] As a result, individuals with cardiovascular diseases or diabetes,[15,16] sickle-cell disease, or conditions associated with immunosuppression,[17,18] which exacerbate the immune response to SARS-CoV-2 infection, may contribute to higher-than-expected mortality for younger age-groups.

Within the first 8 weeks of the first introduction into the region on February 27, 2020, imported seeding events have occurred almost universally in SSA; however, capacity for detection, reporting, and control efforts varies.[7,10] Across the region and with ranging degrees of enforcement,[19,20] countries have implemented suites of preventative interventions, including school closures,[21] curfews, and other social distancing measures,[22] as well as border and airport closures.[23] Countries reporting high numbers of cases since the original seeding in SSA—such as South Africa and Rwanda[24]—were hypothesized to have stronger detection/preparedness

* Address correspondence to Benjamin M. Althouse, Institute for Disease Modeling, Global Health Division Bill & Melinda Gates Foundation, 500 5th Ave. N, Seattle, WA 98109. E-mail: balthouse@idmod.org

systems. Lower observed case counts or delayed reporting of initial cases, relative to the date of first seeding in SSA, could be due to poor detection. It is the goal of this work to systematically collect and present information on COVID-19 cases in all affected countries in SSA during the first several weeks since known importation, with the aim of giving starting points for prediction of onward transmission. We present reported information in the context of countries' pandemic preparedness and COVID-specific vulnerability to underscore the risk of undetected transmission in countries with low numbers of recorded cases to date so that levels of underdetection can be accounted for in such projections.

METHODS

Data collection and variable definitions. A line list of confirmed cases in 49 SSA countries (Library of Congress list,[25] excluding territories) was developed using data published by national governmental institutions (i.e., ministries of health and public health institutes) through their websites, press releases, and social media accounts, and supplemented with information from news outlets (links to all sources are provided in the Supplemental Table 1). Data on gender, age, travel history (including travel locations and dates of entry into the country where case confirmation occurred), date of reported confirmation, whether a case was due to importation or secondary transmission (both known and community transmission), and information on how the case was detected (e.g., active surveillance monitoring or self-presentation) were recorded for each case. We conducted Google searches using keywords such as "COVID," "Ministry of Health," "situation report," "press release," the names of individual countries, and/or the date in the language of the respective countries (e.g., English, French, Portuguese, and Swahili), although more specific searches in news outlets were performed in English. We also searched Facebook and Twitter for accounts of national governmental institutions. As information was collected from multiple sources, the daily case totals per country in the line list were compared with country-reported totals and/or the WHO situation report totals.[24] The number of entries in the data file for a given date, even in the absence of case-level details, was determined by country-reported totals, where available, or by the WHO situation report totals, when country-reported totals were not available. All data collected for the line list are provided as Supplemental Material.

Reported confirmed cases were assumed to be imported when classified as such by national institutions, or when case information included evidence of recent travel history. For countries with limited data available, information on whether cases were imported or due to local transmission was evaluated from the aggregate information in daily WHO situation reports, when feasible. Cases with uncertain travel history or not enough information to determine their importation status were included in the line list but not given a status.

Imported cases were described in terms of gender, 10-year age categories, and time between arrival in country and date that case confirmation was reported. The difference in the gender ratio for imported cases versus for cases due to local transmission was evaluated using a chi-squared test ($P < 0.05$ considered statistically significant). Temporal trends in the frequency of importation events across SSA and in the continents from which they originated were evaluated.

Preparedness and reporting. The availability of publicly reported data was assessed in terms of an "availability score" or the average number of indicators which were reported or could be inferred per imported case. The indicators included were gender, age, date of case confirmation, travel history, date of arrival in the country, and whether detection was due to active monitoring or self-presentation. These indicators were selected as they have implications for disease severity, timing of infectiousness, origin of infection, and mode of detection for assessing sensitivity of surveillance system during the initial outbreak phase. Such indicators are standard for recreating transmission chains and developing age-of-infection or other mathematical models. Availability was assessed for the first 10 imported cases in each country (or for all cases in countries with fewer than 10 cases reported as of April 21, 2020). To describe changes in availability of individual-level information with increasing incidence, an availability score was calculated and reported for the second, third, and fourth sets of 10 imported cases in countries with more than 10 reported importation events. The first draft of the manuscript was prepared over the month of April 2020.

The relationship between country-level pandemic preparedness and case reporting rates was assessed. Although various indices for assessing country-level preparedness and response exist, we chose to use the GHSI overall score in our analysis as the index captures system-level characteristics which would enable detection of and response to SARS-CoV-2 transmission, reporting of cases, communication of response policies/efforts, and clinical capacity to manage symptomatic cases. In addition, the GHSI includes metrics around governance and compliance with international regulations. For the clustering analysis, we assessed general concordance across metrics of pandemic preparedness, as described in the next section. A Poisson generalized linear model was applied to relate the GHSI overall score with cumulative case counts as of April 21, 2020. To account for countries being in different stages of their outbreaks, we included the number of days since the first reported importation as an offset term in the regression model. We also used a Cox proportional hazards model to consider the daily probability of a country reporting any case since the first introduction into the region, after adjusting for the country's overall GHSI score and flight traffic. Flight traffic was included as the total annual number of domestic and international passengers (in billions) carried by air transport carriers registered in a country. Data reflect those reported to the International Civil Aviation Organization from 1970 to 2018.[26] Data for the most recent year available were included for countries with information from 2016 until present. Thirteen of the 49 SSA countries did not have recent flight traffic data (from 2016 until present) and were excluded from the adjusted Cox model.

Relationship among metrics of reporting likelihood, policy implementation, and reported COVID-19 burden. K-means clustering was conducted to identify groups of countries based on indicators related to increased likelihood of detection and reporting as well as capacity to implement response policies. These included the proportion of the population 60 years or older.[27] The age threshold of 60 years was selected because of evidence to date on increased risk of severe disease in individuals older than 60 years,[8] and the expectation that severe cases would be detected with higher probability due to care seeking. We also accounted for reporting and detection capacity as per the GHSI overall score. In addition, to reflect general capacity to institute and implement policies in the context of variable socioeconomic and political stability, we included gross domestic product (GDP) per capita[28] and the government effectiveness indicator of the Worldwide Governance Indicators.[29]

Because of limited data available for Western Sahara, it was not included in the cluster analysis. The average data availability score (as defined in the previous section) for first 10 importation events, average cumulative reported cases per 100,000 population, and average case fatality as of April 21, 2020 were calculated for each of the clusters. Our outcome measures also included day of policy implementation relative to the first reported introduction into the country for policies around general stay at home orders, school closures, and workplace closures. Data on dates of policy introduction were derived from the Oxford University Coronavirus Government Response Tracker.[30] All outcome metrics were presented as medians and means for the cluster. Differences in average data availability across and between clusters were evaluated using analysis of variance and the Tukey honestly significant difference test.

Last, whereas we used the GHSI overall score in the clustering algorithm, we present descriptive statistics for other indices of pandemic preparedness (i.e., SPAR) or COVID-19 vulnerability (i.e., Surgo Foundation's CCVI) to demonstrate degree of concordance.

RESULTS

Overview. Of the 13,201 cases listed between February 27, 2020 and April 21, 2020, individual-level information on the status (i.e., importation or not) for 2,516 reported cases could be determined from available information. Among the individuals with known status, 1,129 importation events were identified (1,129/2,516, 44.9%), whereas 1,387 cases were listed as being infected because of local transmission (i.e., contacts of known travelers or as a result of unexplained community transmission) (1,387/2,516, 55.1%). Imported cases were majority male (343/528, 65.0%), with a median age of 41.0 years (range: 6 weeks–88 years; IQR: 31–54 years) (Figure 1, Table 1). Cases due to local transmission were also majority male (177/307, 57.7%), with a median age of 35.0

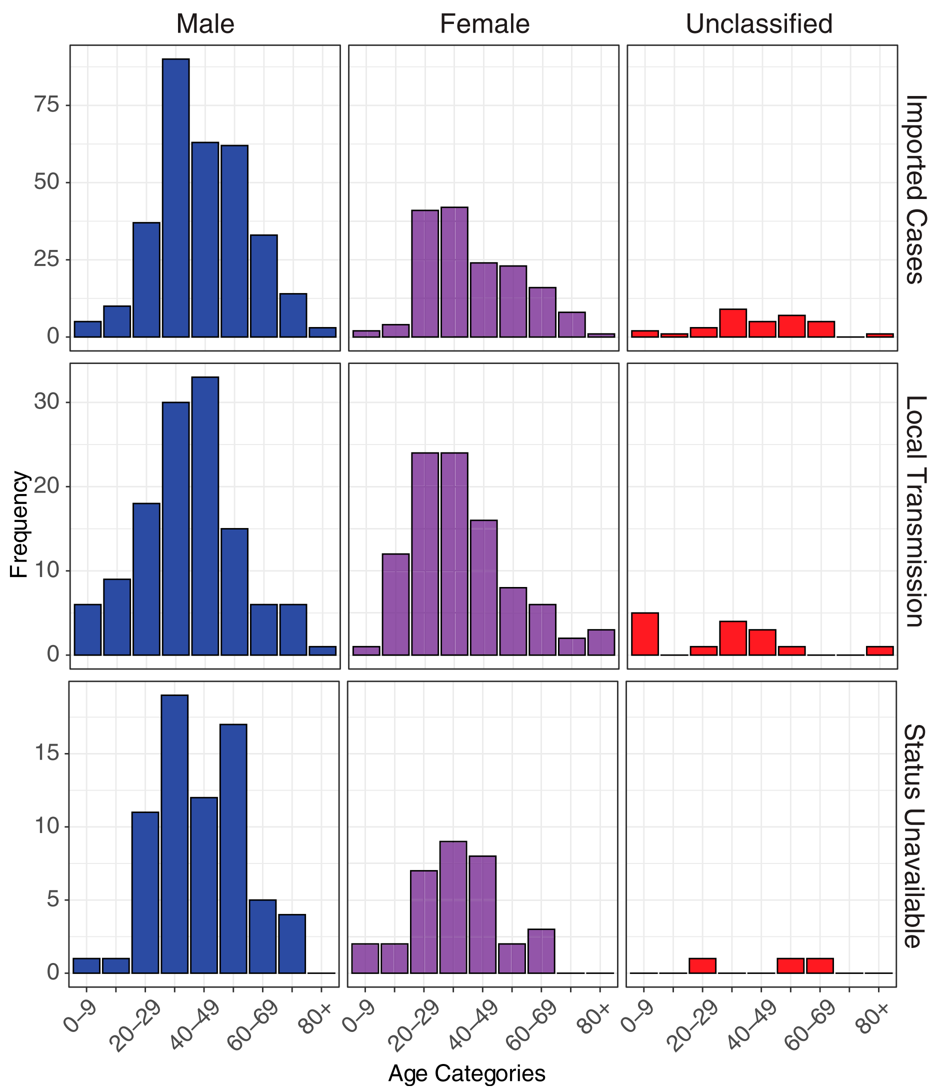

FIGURE 1. Age–gender distribution of confirmed cases of COVID-19 in sub-Saharan Africa by status (imported cases, cases due to local transmission, and cases with status unavailable). Only cases with both age and gender data are reflected.

TABLE 1
Characteristics of confirmed case, by status (importation vs. local transmission)

Characteristic	Overall n = 13,201	Imported n = 1,129	Local transmission n = 1,387	Unknown n = 10,685
Male, n (%)	636 (62.6)	343 (65.0)	177 (57.7)	116 (64.1)
Median age (range)	39 years (28 days–105 years)	41 years (6 weeks–88 years)	35 years (28 days–105 years)	39.5 years (2–78 years)
Mode of detection, n (%)				
Active monitoring	1,327 (93.6)	446 (89.7)	483 (93.1)	396 (99.0)
Self-presentation	91 (6.4)	51 (10.3)	36 (6.9)	4 (1.0)

years (range: 28 days–105 years; IQR: 26–47 years). The proportion of males among imported cases was significantly higher than that among cases due to local transmission (P = 0.043). For imported cases, time between arrival into the country and reporting of the confirmed case status ranged from 0 to 40 days (median: 7; IQR: 4–12) (Figure 2).

Space–time trends. Two SSA countries reported introduction by March 3, 2020 (2/49, 4.1%), eight countries by March 11 (8/49, 16.3%), 31 total countries by March 19 (31/49, 63.3%), 40 total countries as of March 27 (40/49, 81.6%), and 48 total countries as of April 21, 2020 (Figure 3). Comoros and Lesotho remained the only SSA countries not reporting any cases. Among those that could be line-listed, more than 50 importation events were each reported by South Africa (198/1,129, 17.5%), Senegal (85/1,129, 7.5%), Kenya (81/1,129, 7.2%), Rwanda (76/1,129, 6.7%), Cameroon (71/1,129, 6.3%), Ghana (69/1,129, 6.1%), and Uganda (53/1,129, 4.7%); 60.6% of importation events (684/1,129) had a travel history available. Most imported cases reported recent travel from Europe (363/684, 53.1%), with fewer reports of travel to Asia (173/684, 25.3%), and other affected countries in Africa (92/684, 13.5%), the Americas (52/684, 7.6%), or Oceania (2/684, 0.3%) (Figure 4). Two cases with a travel history reported sea travel via a cruise on which they worked (2/684, 0.3%).

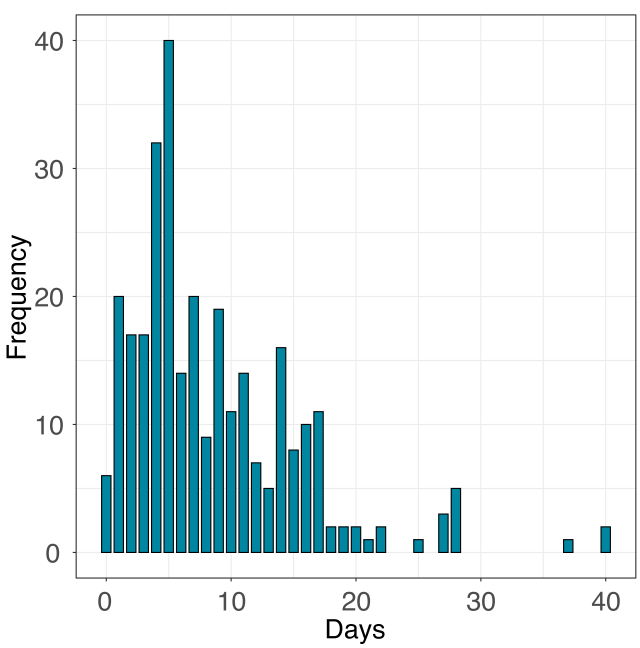

FIGURE 2. Distribution of time between arrival in country and reporting of case confirmation status.

Preparedness and reporting. On average, 3.73 of the six indicators were available for the first 10 imported cases (or the total number of reported importation events for countries with less than 10 total) per country (SD: 1.21, range: 1.0–6.0). For the 21 countries with more than 10 identified importation events, 2.90 indicators were available for the subsequent 10 importation events (SD: 1.52, range: 1.0–5.6); an average of 2.93 indicators were available for the next set of 10 importation events in the 14 countries reporting more than 20 cases (SD: 1.72, range: 1.0–5.7); data availability remained consistently under three average indicators for subsequent imported cases in the 11 countries reporting more than 30 cases (mean = 2.80, SD: 1.42, range: 1.0–5.2).

The median time until reporting an introduction for a country in SSA was 19 days after the first importation into the region (95% CI: 15–24 days, mean = 19.65, SD: 7.81). As per the results of the Cox proportional hazards analysis, time to the first reported importation was significantly associated with the GHSI overall score (HR: 1.01, 95% CI: 1.00–1.02, P = 0.042), although the statistical significance of the relationship was lost after adjusting for scaled number of flight passengers per year (HR: 1.00, 95% CI: 0.99–1.02, P = 0.796). The hazard of reporting an introduction was 13.6% higher with each billion increase in the number of annual air traffic passengers (95% CI: 2.88–25.4, P = 0.012), after adjusting for the GHSI overall score.

Total case counts reported as of April 21, 2020 were related to the GHSI overall score. For each 10-unit increase in the GHSI overall score, cumulative reported case counts increased by a factor of 1.30, after controlling for the number of days since a country reported its first importation event (95% CI: 1.29–1.31, P < 0.001).

Relationship among metrics of reporting likelihood, policy implementation, and reported COVID-19 burden. For the K-means statistical clustering analysis, assessment of within-cluster sum of squares suggested the use of four clusters (i.e., groupings of countries) (Table 2). The governance metric, along with GDP per capita, was found to be more significantly represented within the clustering (Supplemental Figure). Countries in cluster D had the highest median values across all metrics expected to be associated with greater detection—that is, they tended to have the highest proportions of their populations older than 60 years, highest GHSI score, highest per capita GDP, and highest governance score. On average, these countries were reporting the greatest case incidence both at 8 weeks post-introduction into SSA and as of the time the manuscript was prepared, although the results across clusters were not significantly different (all global and pairwise P-values > 0.05) (Table 3). By contrast, countries in cluster C tended to have

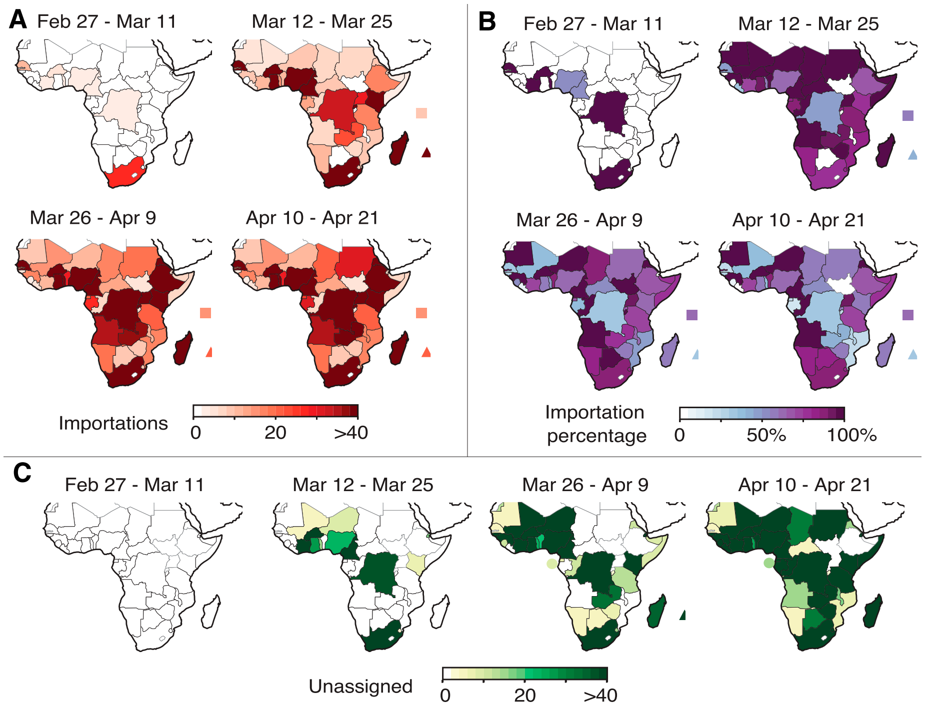

FIGURE 3. (**A**) Cumulative numbers of imported cases for which status (importation vs. local) is known, by week; (**B**) overall percentage of cases (with known status) that are importations; and panel (**C**), number of cases for which status could not be determined from available data. Mauritius and Seychelles are marked by a triangle and square, respectively, given the scale of the maps. Note: In the fourth time period, importation ratios in (**B**) calculated from the line list were adjusted to reflect aggregate information in three countries for which it was available (South Africa, Senegal, and Mauritius).

lower GDP, GHSI score, governance scores, and proportions of their populations 60 years or older. These countries were reporting relatively lower incidence. No statistically significant differences were observed in timing of policy implementation, relative to country-specific introduction, or data availability scores across clusters.

There was general, qualitative agreement across metrics of pandemic preparedness and COVID-specific vulnerability, with country clusters that had higher median GHSI scores also tending to have higher SPAR scores and lower CCVI scores (Supplemental Table 2).

DISCUSSION

We describe the seeding events that have been associated with more than 13,200 COVID-19 cases reported within 54 days of the first known introduction into the SSA region. Most documented importation events were associated with a recent travel history to Europe. As of April 21, 2020, eight countries—Burkina Faso, Cameroon, Côte d'Ivoire, Ghana, Guinea, Niger, Nigeria, and South Africa—had reported at least 500 confirmed cases each, whereas two countries had reported no cases.[31] Regional heterogeneity in reported case counts may be due to current surveillance efforts, particularly around air travel passengers, and not solely due to heterogeneity in transmission. For instance, lower reported incidence may be due to reduced international air travel and the inability to surveil all air passengers and informal border crossings, and future analyses should account for country-specific surveillance strategies for both importations and community transmission to better understand the context. Overall, importations into a country may take weeks or months to seed community spread at noticeable levels,[32] particularly if countries had already begun implementing border closure measures ahead of known introduction. Past outbreaks have evidenced that importations and subsequent transmission will go unnoticed without active monitoring through quarantine and testing to identify infections among travelers,[33] and risk communication directed at population subgroups most likely interacting with international travelers to encourage vigilance around symptom onset.

We found that existing preparedness indicators, namely, the GHSI overall score, were associated with the number of cases reported (i.e., the most well-prepared countries as per the GHSI overall score were reporting more cases) but not with risk of earlier introduction. The sensitivity and utility of the GHSI score and other composite indices in capturing COVID-specific differences in countries' vulnerability to widespread transmission warrant further investigation and analysis, accounting for the different efforts being implemented across the continent to prevent disease spread.[20,23] The interplay of capacity for detection and response and the policies being implemented requires consideration—for instance, high GHSI score countries may have instituted border closures that reduce importation and therefore likelihood of detection. Clusters of countries with high metrics of preparedness (or low vulnerability) did not necessarily show clear trends in terms of timing of policy adoption. Accordingly, at a macro-level, the COVID-19 pandemic has led to shifts in global health spending as countries have turned inward[34] such that national political leadership, accountability, and prioritization of health spending may be important to incorporate in a more nuanced index for COVID-19. Moreover, the indices are not nuanced to

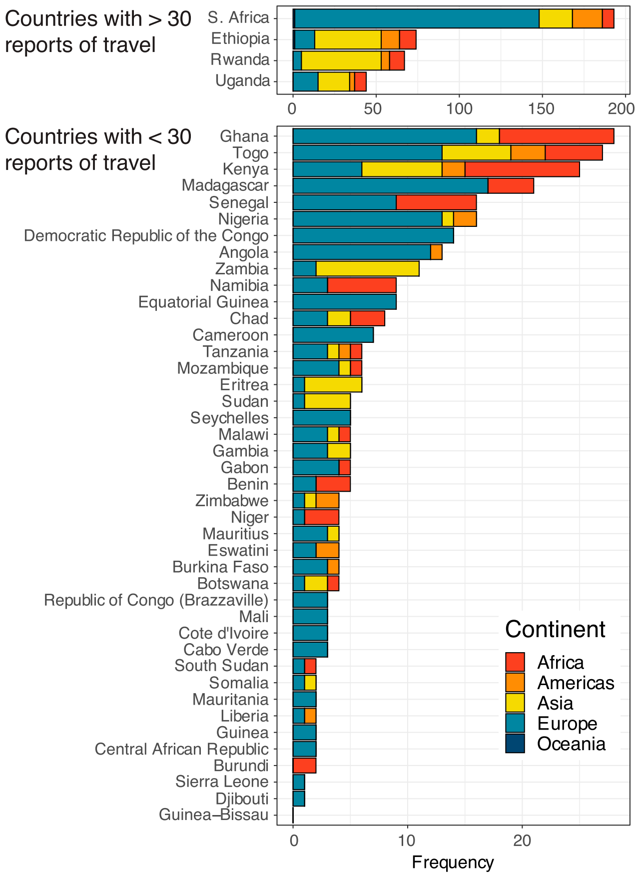

FIGURE 4. Travel history of imported cases by country of residence and continent of travel, for cases with documented travel history. Multiple entries are shown for confirmed cases with a reported travel itinerary that included two or more continents.

include behavioral characteristics (e.g., mobility and resistance to behavior adoption due to denial) of disease-specific relevance. The situation in SSA warrants context-specific consideration in terms of policies around timing of control efforts relative to reported case counts.

The COVID-19 situation is among the first pandemics in which high-capacity computing technology can be leveraged to quickly (and responsibly) disseminate large-scale information globally.[35,36] Information on importation events was extracted from online reporting platforms, such as situation reports, press releases from ministries of health, and dashboards, and social media accounts of political leaders and governmental institutions. Our results include information from 49 countries and more than 2,500 cases; the line list generated from publically released information reflects a growing trend intended to better connect populations with updates on the situation locally and globally.[37] In addition to informing the general population, use of online media to share data on cases is facilitating rapid efforts at better understanding the epidemiology of the disease and its control. With data that provide improved understanding of the novel coronavirus SARS-CoV-2 and of where transmission clusters are commonly occurring, surveillance teams can work with other pillars of the response to implement effective strategies to identify individuals most at risk of infection. Surveillance teams can also promote social distancing measures that are contextually relevant, and recommend personal protective measures based on understanding of common transmission routes. Transparent communication of national recommendations and strategies will benefit other countries looking for best practices during an unprecedented public health emergency.

Competing priorities and resources, not uncommon throughout the world, could lead to delays in providing information to decision-makers. In SSA specifically, external reporting via public electronic platforms may also be challenged by infrastructural and capacity limitations, whereas data collected and disseminated internally for the response may have the granularity needed for COVID-related policy decisions. The significance of data collection, both for internal purposes and public consumption, extends beyond information on direct consequences of COVID-19 and includes indirect consequences—for example, if data from routine surveillance and immunization activities or care-seeking behaviors become less available to capture detrimental impacts on routine care, as has been documented during previous

TABLE 2

Indicators included in clustering analysis to reflect likelihood of detection and policy implementation*

Cluster	Countries in cluster	Population†	Global health security index	% Population aged 60+	Governance	Per capita gross domestic product
A	Cameroon, Cote d'Ivoire, Ethiopia, Gambia, Ghana, Kenya, Liberia, Madagascar, Namibia, Nigeria, Sierra Leone, Uganda, and Zimbabwe	25,500,324 (39,895,376)	77.00 (77.69)	5.3% (5.2%)	–0.63 (–0.74)	$8,282.12 ($11,120.17)
B	Angola, Benin, Burkina Faso, Djibouti, Eswatini, Gabon, Guinea, Malawi, Mali, Mauritania, Mozambique, Niger, and Togo	13,290,444 (13,404,422)	40.00 (38.38)	4.7% (4.9%)	–0.78 (–0.79)	$5,227.76 ($15,353.65)
C	Burundi, Central African Republic, Chad, Comoros, the Democratic Republic of the Congo, Equatorial Guinea, Eritrea, Guinea-Bissau, the Republic of the Congo (Brazzaville), Somalia, South Sudan, and Sudan	7,697,334 (16,739,091)	20.50 (21.58)	4.9% (4.9%)	–1.63 (–1.69)	$4,751.00 ($11,877.35)
D	Botswana, Cabo Verde, Lesotho, Rwanda, Senegal, Seychelles, and South Africa	2,197,426 (12,127,394)	51.00 (63.86)	8.1% (8.0%)	0.29 (0.13)	$18,148.24 ($53,560.56)

* Within-cluster distributions presented as means (medians).
† Population size given for reference, not included as a metric in the clustering analysis.

Table 3
Outcome metrics on reported case burden by cluster

Cluster	Days until stay home policy	Days until workplace closure policy	Days until school closure policy	Case fatality (first 8 weeks)*	Incidence per 100,000 (first 8 weeks)*	Incidence per 100,000 (cumulative to present)	Data availability score	Probability of any case†
A	12.00 (13.92)	9.00 (8.33)	3.00 (4.46)	0.03 (0.04)	0.56 (1.31)	82.41 (126.98)	4.00 (4.22)	1.00
B	11.00 (13.46)	13.00 (16.85)	3.00 (3.62)	0.04 (0.05)	1.39 (9.18)	48.91 (171.81)	3.20 (3.35)	1.00
C	14.00 (17.22)	12.00 (11.67)	12.00 (8.71)	0.00 (0.04)	0.30 (1.68)	38.77 (101.62)	3.67 (3.58)	0.92
D	8.00 (6.71)	15.00 (37.83)	15.00 (3.57)	0.01 (0.02)	2.59 (4.97)	175.52 (592.07)	4.60 (4.19)	0.86

* Calculated based on case counts through April 21, 2020 (8 weeks post-introduction into the region).
† Probability of any case per cluster was determined as the number of countries in the cluster with at least one reported COVID-19 case as of April 21, 2020 divided by the total number of countries in the cluster.

large-scale outbreaks.[38–40] Data on the burden of COVID-19, relative to the social and economic burden of control measures, are critical for informed decision-making. Notably, information from health agencies has guided mathematical modeling efforts to project case counts and assess potential impact of interventions, such as social distancing.[41–43] The robustness of modeling efforts to inform decision-making depends on the quality of data used to develop them. Support to data teams could enhance data collection and data quality so that modeling and other analyses can be performed effectively and efficiently. It could also enhance the capacity of how to understand and handle missing information, as many different factors may impact the completeness and timeliness of case information—for example, care-seeking among infected individuals, healthcare availability, testing capacity, and poor coordination of data collection across institutions. Investing in data collection and analysis capacity could have long-term benefits beyond COVID-specific surveillance.

Limitations. While we attempted to search all available sources, data included in our line list are likely not exhaustive. We also recognize that news sources used to supplement official sources may be less reliable in the information they provide. Information was extracted from available text material only. Videos of press conferences or radio announcements were not considered unless their content appeared in published press articles. By opting to maintain an individual-level database, we were unable to reflect some information provided by national authorities in aggregate form. Where possible, we reconciled totals across individual entries with aggregate information or adjusted for the latter, such as with the importation percentages for South Africa, Senegal, and Mauritius. Furthermore, data were not reported consistently for some indicators. For instance, we note that data being disseminated to the public are often for purposes of information sharing and education rather than to guide detailed decision-making. Moreover, with increasing incidence, less information is provided publicly about cases at the individual level, leading to lack of information on gender, age, and status (importation versus local transmission) for most of the cases in South Africa, for instance. This precluded our investigation into whether cases with unclassified status may resemble known importations versus locally transmitted cases. We further note that data availability, particularly on public sites, is related to factors beyond detection capacity alone, as factors related to governance, accessibility of healthcare, and individual-level behavior decisions influence probabilities of importation, testing, and sharing of information. Despite the limitations identified here, our study is the first to generate a comprehensive dataset of COVID-19 cases for the SSA region to aid researchers in initiating analyses with potential to support local decision-making.

CONCLUSION

With the introduction of COVID-19 into SSA, subregional variation in the incidence and reported ratio of imported versus locally infected cases has been observed. The availability and quality of publicly released information also vary significantly. As of 54 days post-introduction into the SSA region, countries with high scores on preparedness indicators, high likelihood of severe cases, and high ranking for government effectiveness were not necessarily reporting with more per-case information or higher incidence. The lack of information on imported cases could signal the potential for undetected transmission and highlight the complexity of detection and response for a largely asymptomatic infection in the context of SSA.

Received November 23, 2020. Accepted for publication February 1, 2021.

Published online March 8, 2021.

Note: Supplemental tables and figure appear at www.ajtmh.org.

Acknowledgments: We thank Kurt Frey for his help in collecting data and Assaf Oron for his helpful comments on the manuscript. We also acknowledge the tireless efforts of country teams to collect and disseminate information as a way to inform local and global stakeholders. The American Society of Tropical Medicine and Hygiene has waived the Open Access fee for this article due to the ongoing COVID-19 pandemic.

Financial support: L. A. S., P. S., B. H., A. L. O., N. N., D. M., A. O., and B. M. A. were funded by the Bill & Melinda Gates Foundation through the Global Good Fund. L. H.-D. acknowledges support from the National Institutes of Health 1P20 GM125498-01 Centers of Biomedical Research Excellence Award. S. V. S. is supported by startup funds provided by Northeastern University.

Disclosure: Data are available with the manuscript.

Authors' addresses: Laura A. Skrip, Prashanth Selvaraj, Brittany Hagedorn, Andre Lin Ouédraogo, Navideh Noori, Amanda Orcutt, and Dina Mistry, Department of Epidemiology, Institute for Disease Modeling, Seattle, WA, E-mails: lskrip@idmod.org, pselvaraj@idmod.org, bhagedorn@idmod.org, aouedraogo@idmod.org, nnoori@idmod.org, aorcutt@idmod.org, and dmistry@idmod.org. Jamie Bedson, Bill & Melinda Gates Foundation, Seattle, WA, E-mail: bma85@uw.edu. Laurent Hébert-Dufresne, Department of Computer Science, University of Vermont, Burlington, VT, E-mail: laurent.hebert-dufresne@uvm.edu. Samuel V. Scarpino, College of Science, Northeastern University D'Amore-McKim School of Business, Boston, MA, E-mail: s.scarpino@northeastern.edu. Benjamin M. Althouse, Department of Epidemiology, Institute for Disease Modeling, Seattle, WA, Information School, University of Washington, Seattle, WA, and Department of Biology, New Mexico State University, Las Cruces, NM, E-mail: balthouse@idmod.org.

REFERENCES

1. Johns Hopkins Coronavirus Resource Center, 2020. *Johns Hopkins Coronavirus Resource Center*. Available at: https://coronavirus.jhu.edu/map.html. Accessed March 29, 2020.
2. Association IAT, Others, 2018. *Traveler Numbers Reach New Heights. IATA World Air Transport Statistics Released Montreal: IATA*. Available at: https://www.iata.org/en/pressroom/pr/2018-09-06-01/. Accessed March 30, 2020.
3. CE Schlumberger, 2010. *Open Skies for Africa: Implementing the Yamoussoukro Decision*. The World Bank. Available at: https://www.worldbank.org/en/topic/transport/publication/open-skies-for-africa. Accessed March 30, 2020.
4. Sun IY, Jayaram K, Kassiri O, 2017. *Dance of the Lions and Dragons JUNE 2017 How Are Africa and China Engaging, and How Will the Partnership Evolve?* New York, NY: McKinsey & Company. Available at: https://www.mckinsey.com/featured-insights/middle-east-and-africa/the-closest-look-yet-at-chinese-economic-engagement-in-africa. Accessed March 30, 2020.
5. The Washington Post, 2020. Africa Has 1.2 Billion People and Only Six Labs that Can Test for Coronavirus. How Quickly Can They Ramp up? *The Washington Post*. Available at: https://www.washingtonpost.com/world/africa/africa-has-12-billion-people-and-only-six-labs-that-can-test-for-coronavirus-how-quickly-can-they-ramp-up/2020/02/05/68af96de-4758-11ea-91ab-ce439aa5c7c1_story.html. Accessed March 30, 2020.
6. Candido DDS, Watts A, Abade L, Kraemer MUG, Pybus OG, Croda J, Oliveira W, Khan K, Sabino EC, Faria NR, 2020. Routes for COVID-19 importation in Brazil. *J Trav Med 27:* taaa042.
7. Gilbert M et al., 2020. Preparedness and vulnerability of African countries against importations of COVID-19: a modelling study. *Lancet 395:* 871–877.
8. Wu Z, McGoogan JM, 2020. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72,314 cases from the Chinese Center for Disease Control and Prevention. *JAMA 323:* 1239–1242.
9. Nordling L, 2020. "A ticking time bomb": scientists worry about coronavirus spread in Africa. *Science 371*. Available at: https://www.sciencemag.org/news/2020/03/ticking-time-bomb-scientists-worry-about-coronavirus-spread-africa.
10. Global Health Security Index, 2019. *Building Collective Action and Accountability*. Center for Health Security, Johns Hopkins Bloomberg School of Public Health. Available at: https://www.ghsindex.org/report-model/. Accessed March 30, 2020.
11. World Health Organization, Electronic State Parties Self-Assessment Annual Reporting Tool (e-SPAR), 2019. Available at: https://extranet.who.int/e-spar. Accessed March 27, 2020.
12. Surgo Foundation, 2020. *Bringing Greater Precision to the COVID-19 Response*. Available at: https://precisionforcovid.org/ccvi. Accessed November 23, 2020.
13. Guan WJ et al.; China Medical Treatment Expert Group for COVID-19, 2020. Clinical characteristics of coronavirus disease 2019 in China. *N Engl J Med 382:* 1708–1720.
14. de-Graft Aikins A, Unwin N, Agyemang C, Allotey P, Campbell C, Arhinful D, 2010. Tackling Africa's chronic disease burden: from the local to the global. *Glob Health 6:* 5.
15. Azevedo M, Alla S, 2008. Diabetes in sub-saharan Africa: Kenya, Mali, Mozambique, Nigeria, South Africa and Zambia. *Int J Diabetes Dev Ctries 28:* 101–108.
16. Atun R et al., 2017. Diabetes in sub-Saharan Africa: from clinical care to health policy. *Lancet Diabetes Endocrinol 5:* 622–667.
17. Mbithi A et al.; KAIS Study Group, 2014. Tuberculosis and HIV at the national level in Kenya: results from the second Kenya AIDS indicator survey. *J Acquir Immune Defic Syndr 66 (Suppl 1):* S106–S115.
18. Murray CJL et al., 2014. Global, regional, and national incidence and mortality for HIV, tuberculosis, and malaria during 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet 384:* 1005–1070.
19. Mpungu P, 2020. Kenya Police Under Fire Over "Excessive Force" as Curfew Begins. *Al Jazeera*. Available at: https://www.aljazeera.com/news/2020/03/kenya-police-fire-excessive-force-curfew-begins-200328101357933.html. Accessed March 30, 2020.
20. Time, 2020. Coronavirus prevention measures turn violent in parts of Africa as nations go into lockdown. *Time*. Available at: https://time.com/5811945/coronavirus-prevention-africa/. Accessed March 30, 2020.
21. UNESCO, 2020. *COVID-19 Educational Disruption and Response*. UNESCO. Available at: https://en.unesco.org/themes/education-emergencies/coronavirus-school-closures. Accessed March 30, 2020.
22. Rfi COW, 2020. Africa: SA in lockdown, curfews in the west as continent braces for COVID-19 wave. *RFI*. Available at: http://www.rfi.fr/en/africa/20200324-lockdown-curfew-restrictions-africa-coronavirus-covid-19. Accessed March 30, 2020.
23. Salcedo A, Cherelus G, 2020. Coronavirus Travel Restrictions, Across the Globe. *The New York Times*. Available at: https://www.nytimes.com/article/coronavirus-travel-restrictions.html. Accessed March 30, 2020.
24. WHO, 2020. *Coronavirus Disease 2019 (COVID-19) Situation Report*. Geneva, Switzerland: World Health Organization. Available at: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200327-sitrep-67-covid-19.pdf?sfvrsn=b65f68eb_4. Accessed March 30, 2020.
25. The Library of Congress, 2010. *List of Sub-saharan African Countries*. Available at: https://www.loc.gov/rr/amed/guide/afr-countrylist.html. Accessed March 27, 2020.
26. World Bank, 2020. *Air Transport, Passengers Carried*. Available at: https://data.worldbank.org/indicator/is.air.psgr. Accessed March 30, 2020.
27. Global Health Observatory, World Health Organization, 2020. *Probability (%) of Dying between Age 30 and Exact Age 70 from Any of Cardiovascular Disease, Cancer, Diabetes, or Chronic Respiratory Disease (Global Strategy for Women's, Children's and Adolescents' Health)*. Available at: https://apps.who.int/gho/data/node.imr.NCDMORT3070?lang=en. Accessed March 30, 2020.
28. World Bank, 2020. *GDP Per Capita (Current US$)*. Available at: https://data.worldbank.org/indicator/. Accessed March 30, 2020.
29. Kaufmann D, Kraay A, Mastruzzi M, 2010. *The Worldwide Governance Indicators: Methodology and Analytical Issues*. Available at: https://papers.ssrn.com/abstract=1682130. Accessed March 30, 2020.
30. Coronavirus Government Response Tracker, 2020. Available at: https://www.bsg.ox.ac.uk/research/research-projects/coronavirus-government-response-tracker. Accessed March 30, 2020.
31. World Health Organization, 2020. *Novel Coronavirus (2019-nCoV) Situation Reports*. Available at: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports. Accessed March 30, 2020.
32. Chinazzi M et al., 2020. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science 368:* 395–400.
33. World Health Organization, 2020. *Updated WHO Recommendations for International Traffic in Relation to COVID-19 Outbreak*. Available at: https://www.who.int/news-room/articles-detail/updated-who-recommendations-for-international-traffic-in-relation-to-covid-19-outbreak. Accessed April 1, 2020.
34. Organization for Economic Co-operation and Development, 2020. *COVID-19 and Africa: Socio-Economic Implications and Policy Responses*. Available at: https://www.oecd.org/coronavirus/policy-responses/covid-19-and-africa-socio-economic-implications-and-policy-responses-96e1b282/. Accessed August 11, 2020.
35. Dong E, Du H, Gardner L, 2020. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect Dis 20:* P533–P534.
36. Ienca M, Vayena E, 2020. On the responsible use of digital data to tackle the COVID-19 pandemic. *Nat Med26:* 463–464.
37. Atlantic Council, 2020. African Leaders Respond to coronavirus... On Twitter - Atlantic Council. *Atlantic Council*. Available at:

https://atlanticcouncil.org/blogs/africasource/african-leaders-respond-to-coronavirus-on-twitter/. Accessed March 28, 2020.

38. Sun X et al., 2017. Impact of the ebola outbreak on routine immunization in western area, Sierra Leone–a field survey from an ebola epidemic area. *BMC Public Health 17:* 363.
39. Gigova R, 2019. Measles has claimed more than twice as many lives as Ebola in DRC, CNN, 28 November 2019. Available at: https://edition.cnn.com/2019/11/28/africa/measles-ebola-drc/index.html. Accessed April 1, 2020.
40. Parpia AS, Ndeffo-Mbah ML, Wenzel NS, Galvani AP, 2016. Effects of response to 2014–2015 ebola outbreak on deaths from malaria, HIV/AIDS, and tuberculosis, west Africa. *Emerg Infect Dis J 22:* 433.
41. Institute for Health Metrics and Evaluation, 2020. *COVID-19 Projections*. Available at: https://covid19.healthdata.org/projections. Accessed March 30, 2020.
42. Hellewell J, Abbott S, Gimma A, Bosse NI, Jarvis CI, Russell TW, Munday JD, Kucharski AJ, Edmunds WJ; Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group, Funk S, Eggo RM, 2020. Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. *Lancet Glob Health 8:* e488–e496.
43. Koo JR, Cook AR, Park M, Sun Y, Sun H, Lim JT, Tam C, Dickens BL, 2020. Interventions to mitigate early spread of SARS-CoV-2 in Singapore: a modelling study. *Lancet Infect Dis 20:* 678–688.