

1 **High variation expected in the pace and burden of SARS-CoV-2
2 outbreaks across sub-Saharan Africa**

3
4 Benjamin L. Rice^{1,2}, Akshaya Annapragada³, Rachel E. Baker^{1,4}, Marjolein Bruijning¹, Winfred
5 Dotse-Gborgbortsi⁵, Keitly Mensah⁶, Ian F. Miller¹, Nkengafac Villyen Motaze^{7,8}, Antso
6 Raherinandrasana^{9,10}, Malavika Rajeev¹, Julio Rakotonirina^{9,10}, Tanjona Ramiadantsoa^{11,12,13},
7 Fidisoa Rasambainarivo^{1,14}, Weiyu Yu¹⁵, Bryan T. Grenfell^{1,16}, Andrew J. Tatem⁵, C. Jessica E.
8 Metcalf^{1,16}

- 9
10 1. Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ,
11 USA
12 2. Madagascar Health and Environmental Research (MAHERY), Maroantsetra,
13 Madagascar
14 3. Johns Hopkins University School of Medicine, Baltimore, MD, USA
15 4. Princeton Environmental Institute, Princeton University, Princeton, NJ, USA.
16 5. WorldPop, School of Geography and Environmental Science, University of
17 Southampton, Southampton, UK
18 6. Centre population et Développement CEPED (Université de Paris), Institut Recherche et
19 Développement, Paris, France
20 7. Centre for Vaccines and Immunology (CVI), National Institute for Communicable
21 Diseases (NICD) a division of the National Health Laboratory Service (N HLS), South
22 Africa
23 8. Department of Global Health, Faculty of Medicine and Health Sciences, Stellenbosch
24 University, Cape Town, South Africa
25 9. Faculty of Medicine, University of Antananarivo, Madagascar
26 10. Teaching Hospital of Care and Public Health Analakely, Antananarivo, Madagascar
27 11. Department of Life Science, University of Fianarantsoa, Madagascar
28 12. Department of Mathematics, University of Fianarantsoa, Madagascar
29 13. Department of Integrative Biology, University of Wisconsin-Madison, WI, USA
30 14. Mahaliana Labs SARL, Antananarivo, Madagascar
31 15. School of Geography and Environmental Science, University of Southampton,
32 Southampton, UK
33 16. Princeton School of Public and International Affairs, Princeton University, NJ, USA
34
35
-

36
37 Link to SSA-SARS-CoV-2 online companion tool: <https://labmetcalf.shinyapps.io/covid19-burden-africa/>

38
39 Link to github repository containing data and code: <https://github.com/labmetcalf/SSA-SARS-CoV-2>

40
41 (Changes from the previous submission are highlighted in **ORANGE**)

42

43 **Abstract**

44 A surprising feature of the SARS-CoV-2 pandemic to date is the low burdens reported in sub-
45 Saharan Africa (SSA) countries relative to other global regions. Potential explanations (e.g.,
46 warmer environments¹, younger populations^{2–4}) have yet to be framed within a comprehensive
47 analysis accounting for factors that offset the effects of climate and demography. Here, we
48 synthesize factors hypothesized to be driving the pace of this pandemic and its burden as it
49 moves across SSA, encompassing demographic, comorbidity, climatic, healthcare capacity,
50 intervention efforts implemented to date, and human mobility dimensions of risk. We find large
51 scale diversity in probable drivers, consistent with highly variable outcomes among SSA
52 countries, pointing to a need for caution in interpreting analyses that aggregate across low and
53 middle-income settings. While simulation shows that extensive climatic variation among SSA
54 population centers has little effect on early outbreak trajectories, heterogeneity in connectivity,
55 although rarely considered, is likely an important contributor to variance in the pace of viral
56 spread. The prolonged, asynchronous outbreaks expected in weakly connected settings may
57 result in extended stress to health systems. In addition, the observed variability in comorbidities
58 and access to care will likely modulate the severity of infection: While reported mortality to date
59 remains low, variability in surveillance and death registration systems could obscure excess
60 burden, and we show that even small shifts in the infection fatality ratio towards younger ages
61 can eliminate the protective effect of younger populations. We highlight countries with continued
62 elevated risk of ‘slow pace’, high burden outbreaks. Further empirical data will be central to both
63 understanding and managing the burden of SARS-CoV-2 in SSA. Potential returns on
64 investments in strengthening efforts to monitor patterns in severity over age, the spatial
65 progression of outbreaks sub-nationally, or the relationship between epidemic pace and health
66 system disruptions may be anticipated by our synthesis of context-specific risk in terms of pace
67 and burden.

68
69
70

71 The trajectory of the SARS-CoV-2 pandemic in lower latitude, lower income countries including
72 in Sub-Saharan Africa (SSA) remains uncertain (**Table 1**). To date, reported case counts and
73 mortality in SSA have lagged behind other geographic regions: all SSA countries, with the
74 exception of South Africa, reported less than 75,000 total cases and less than 2,000 deaths as
75 of September 2020⁵ (**Table S1**) - totals far less than observed in Asia, Europe, and the
76 Americas ^{5,6}. However, variation in reporting between countries and some seroprevalence
77 surveys that suggest high rates of local infection ⁷⁻⁹ make it unclear if the relatively few reported
78 cases and deaths to date indicate a generally reduced epidemic potential in SSA ¹⁰.

79

80 Comparisons across SSA populations based on reported infection rates are obscured by
81 heterogeneity in surveillance capacity (e.g., variation in testing rates among countries) and
82 correlation between surveillance and infection reporting ¹¹ (**Figure S1**). Combining reported
83 death counts with assumptions about the probability of mortality given infection ² yields
84 generally low estimates of the percentage of the population expected to have been infected (i.e.,
85 less than 10%) but this varies more than ten-fold between SSA countries and, critically, is sensitive to assumptions about the death reporting rate (**Figure 1A**) and infection fatality ratio (*IFR*, **Figure 1B**). Serology provides an alternative and more
86 direct measure of the percentage infected. Initial serological studies of blood-banks in Kenya (5-
87 10%) ⁷, health care workers in urban Malawi (9-16%) ⁹, or from Niger State in Nigeria (20-30%) ⁸
88 indicate infection rates may be higher in some settings, but only the latter was designed as a
89 representative sample and serology based estimates remain sparse in SSA.

90

91 Given limitations in inference from direct measures of infection and death rates, experience from
92 locations in which the pandemic has progressed more rapidly provides a valuable basis of
93 knowledge to assess the relative risk of populations in SSA and identify those remaining at
94 greatest risk. For example, individuals in lower socio-economic settings have been
95 disproportionately affected in high latitude countries, ^{12,13} indicating poverty as an important
96 determinant of risk. Widespread disruptions to routine health services have been reported ¹⁴⁻¹⁶
97 and are likely to be an important contributor to the burden of the pandemic in SSA ¹⁷. The role of
98 other factors from demography ²⁻⁴ to health system context ¹⁸ and intervention timing ^{19,20} is also
99 increasingly well-characterized.

100

102 **Table 1**

| | |
|-------------------------------|---|
| Background | As the SARS-CoV-2 pandemic expanded globally, reported incidence and mortality remained low in Sub-Saharan Africa (SSA). Yet, a general conclusion that SSA may avoid the high burdens seen elsewhere neglects considerable national and subnational variability in likely drivers of the pandemic's impacts, from its burden to its pace, and does not address variable surveillance and registration systems. |
| Main Findings and Limitations | Synthesizing data on likely drivers of the pace and burden of SARS-CoV-2 in SSA reveals extensive variability in factors that can define the burden once individuals have become infected. Pairing this with simulations of the trajectory of the outbreak indicates little effect of climate but potentially prolonged outbreaks in many settings due to heterogeneities in connectivity, an effect which may be amplified by control efforts. However, although we provide a qualitative overview of the continued potential impact of the pandemic in SSA, quantitative projections remain intractable given a lack of information on the quantitative impact of important risk factors (from how comorbidities might shape the infection fatality ratio to how remoteness will reduce spread). Additionally, uncertainties associated with existing surveillance and mortality registration data impede direct comparison of expectations with national data. |
| Policy Implications | To narrow the range of expectations for country trajectories in SSA for incidence and mortality, strengthening surveillance and registration is necessary. Additional tools for surveillance such as serology, or approaches that quantify excess mortality, will provide important complementary measures. Our national and sub-national analyses point to where returns on investment in strengthening surveillance may yield the greatest returns. Countries with higher comorbidity risk may have most to gain from understanding determinants of mortality; low connectivity countries will benefit from investments in delineating the spatial extent of outbreaks; all countries will benefit from evaluating the intersection between epidemic pace and health system disruptions. |

103

104

105 **Factors expected to increase and decrease SARS-CoV-2 risk in SSA**

106 Characterizing and anticipating the trajectory of ongoing outbreaks in SSA requires considering
 107 variability in known drivers, and how they may interact to increase or decrease risk across
 108 populations in SSA and relative to non-SSA settings (**Figure 2**). For example, while most
 109 countries in SSA have 'young' populations, suggesting a decreased burden (since SARS-CoV-2
 110 morbidity and mortality increase with age ²⁻⁴), prevalent infectious and non-communicable
 111 comorbidities may counterbalance this demographic 'advantage' ^{18,21-23}. Similarly, SSA
 112 countries have health systems that vary greatly in their infrastructure, and dense, resource-
 113 limited urban populations may have fewer options for social distancing ²⁴. Yet, decentralized,

114 community-based health systems that benefit from recent experience with epidemic response
115 (e.g., to Ebola^{25,26}) can be mobilized. Climate is frequently invoked as a potential mitigating
116 factor for warmer and wetter settings¹, including SSA, but climate varies greatly between
117 population centers in SSA and large susceptible populations may counteract any climate forcing
118 during initial phases of the epidemic²⁷. Connectivity, at international and subnational scales,
119 also varies greatly^{28,29} and the time interval between viral introductions and the onset of
120 interventions such as lockdowns will modulate the trajectory¹¹. Finally, burdens of malnutrition,
121 infectious diseases, and many other underlying health conditions are higher in SSA (**Table S2**),
122 and their interactions with SARS-CoV-2 are, as of yet, poorly understood; conversely cross-
123 protection from either SARS-CoV-2 infection or disease as a result of prior infection by
124 widespread circulating coronaviruses remains a possibility.

125

126 The highly variable social and health contexts of countries in SSA will drive location-specific
127 variation in the magnitude of the burden, the time-course of the outbreak, and options for
128 mitigation. Here, we synthesize the range of factors hypothesized to modulate the potential
129 outcomes of SARS-CoV-2 outbreaks in SSA settings by leveraging existing data sources and
130 integrating novel SARS-CoV-2 relevant mobility and climate-transmission models. Data on
131 direct measures and indirect indicators of risk factors were sourced from publicly available
132 databases including from the WHO, World Bank, UNPOP, DHS, GBD, and WorldPop, and
133 newly generated data sets (see **Table S3** for details). We organize our assessment around two
134 aspects that will shape national outcomes and response priorities in the event of widespread
135 outbreaks: i) the burden, or expected severity of the outcome of an infection, which emerges
136 from age, comorbidities, and health systems functioning, and ii) the rate of spread within a
137 geographic area, or pace of the pandemic.

138

139 We group factors that may drive the relative rates of these two features (mortality burden and
140 pace of the outbreak) along six dimensions of risk: (A) Demographic and socio-economic
141 parameters related to transmission and burden, (B) Comorbidities relevant to burden, (C)
142 Climatic variables that may impact the magnitude and seasonality of transmission, (D)
143 Prevention measures deployed to reduce transmission, (E) Accessibility and coverage of
144 existing healthcare systems to reduce burden, and (F) Patterns of human mobility relevant to
145 transmission (**Table S2**).

146

147

148 **National and subnational variability in SSA**

149 National scale variability in SSA among these dimensions of risk often exceeds ranges
150 observed across the globe (**Figure 3A-D**). For example, estimates of access to basic
151 handwashing (i.e., clean water and soap³⁰) among urban households in Mali, Madagascar,
152 Tanzania, and Namibia (62-70%) exceed the global average (58%), but fall to less than 10% for
153 Liberia, Lesotho, Congo DRC, and Guinea-Bissau (**Figure 3D**). Conversely, the range in the
154 number of physicians is low in SSA, with all countries other than Mauritius below the global
155 average (168.78 per 100,000 population) (**Figure 3A**). Yet, estimates are still heterogeneous
156 within SSA, with, for example, Gabon estimated to have more than 4 times the physicians of
157 neighboring Cameroon (36.11 and 8.98 per 100,000 population, respectively). This disparity is
158 likely to interact with social contact rates among the elderly in determining exposure and clinical
159 outcomes (e.g., for variation in household size see **Figure 3E-F**). Relative ranking across
160 variables is also uneven among countries with the result that this diversity cannot be easily
161 reduced (e.g., the first two principal components explain only 32.6%, and 13.1% of the total
162 variance as shown in **Figure S5**), indicating that approaches reliant on a small subset of
163 variables will fail to capture the observed variation among SSA countries.

164

165 **Severity of infection outcome**

166 To first evaluate variation in the burden emerging from the severity of infection outcome, we
167 consider how demography, comorbidity, and access to care might modulate the age profile of
168 SARS-CoV-2 morbidity and mortality²⁻⁴. Subnational variation in the distribution of high risk age
169 groups indicates considerable variability, with higher burden expected in urban settings in SSA
170 (**Figure 4A**), where density and thus transmission are likely higher³¹.

171

172 Comorbidities and access to clinical care also vary across SSA (e.g., for diabetes prevalence
173 and hospital bed capacity see **Figure 4B**). In comparison to settings where previous SARS-
174 CoV-2 infection fatality ratio (*IFR*) estimates have been reported, mortality due to
175 noncommunicable diseases in SSA increases more rapidly with age (**Figure S6**) suggesting risk
176 for an elevated *IFR* in some settings. Conversely, an analysis of the reported age-specific death
177 data available from Kenya and South Africa suggested low *IFRs* in comparison to non-SSA
178 countries³². Comparison of empirical age profiles of mortality more broadly across SSA is
179 currently limited by the small number of total deaths reported to date for many countries (e.g.,
180 30 of 48 SSA countries have reported fewer than 100 total deaths as of September 2020) and
181 incomplete associated age data. Consequently, we use global *IFR* by age estimates and

182 explore the potential effect on mortality of deviations from the expected baseline *IFR* in diverse
183 SSA settings.

184

185 Small shifts (e.g., of 2-10 years) in the *IFR* profile result in large effects on expected mortality for
186 a given level of infection. For example, Chad, Burkina Faso, and the Central African Republic,
187 while among the youngest SSA countries, have a relatively high prevalence of diabetes and
188 relatively low density of hospital beds. Given the age structure of these countries, a slight shift in
189 the *IFR* by age profile towards higher mortality in middle-aged groups (e.g., ages 50-60) would
190 result in mortality increasing to a rate that would exceed a majority of the other, 'older' SSA
191 countries at the unshifted baseline (**Figure 4C**, see supplement for details of methods).

192 Generally, minor shifts in the *IFR* lead to differences larger than the magnitude of the difference
193 expected from differing age structures for countries in SSA.

194

195 Although there is greater access to care in older populations by some metrics (**Figure 3A**,
196 correlation between age and the number of physicians per capita, $r = 0.896$, $p < 0.001$), access
197 to clinical care is highly variable overall (**Figure 4D**) and maps poorly to indicators of
198 comorbidity (**Figure 4E**). Empirical data are urgently needed to assess the extent to which the
199 *IFR*-age-comorbidity associations observed elsewhere are applicable to SSA settings with
200 reduced access to advanced care. Yet both surveillance and mortality registration³³ are
201 frequently under-resourced in SSA, complicating both evaluating and anticipating the burden of
202 the pandemic, and underscoring the urgency of strengthening existing systems²⁶.

203

204 **Pandemic pace**

205 Next, we turn to the pace of the pandemic within each country. The frequency of viral
206 introduction to each country, likely governed by international air travel in SSA³⁴, determines
207 both the timing of the first infections and the number of initial infection clusters that seed
208 subsequent outbreaks. The relative importation risk among SSA cities and countries was
209 assessed by compiling data from 108,894 flights arriving at 113 international airports in SSA
210 from January to April 2020 (**Figure 5A**), stratified by the SARS-CoV-2 status at the departure
211 location on the day of travel (**Figure 5B**). A small subset of SSA countries received a
212 disproportionately large percentage (e.g., South Africa, Ethiopia, Kenya, Nigeria together
213 contribute 47.9%) of the total travel from countries with confirmed SARS-CoV-2 infections, likely
214 contributing to variation in the pace of the pandemic across settings (and consistent with those
215 four countries together contributing 74.7% of all reported cases in SSA)^{34,35}.

216

217 Once local chains of infection are established, the rate of spread within countries will be shaped
218 by efforts to reduce spread, such as handwashing and other non-pharmaceutical interventions
219 (**Figure 3D**), population contact patterns including mobility and urban crowding³¹ (e.g., **Figure**
220 **3C**), and potentially the effect of climatic variation¹. Where countries fall across this spectrum of
221 pace will shape interactions with lockdowns and determine the length and severity of disruptions
222 to routine health system functioning.

223

224 Subnational connectivity varies greatly across SSA, both between subregions of a country and
225 between cities and their rural periphery (e.g., as indicated by travel time to the nearest city over
226 50,000 population, **Figure 5C**). As expected, in stochastic simulations using estimates of viral
227 transmission parameters and mobility, a smaller cumulative proportion of the population is
228 infected at a given time in countries with larger populations in less connected subregions
229 (**Figure 5D**), and including NPIs tends to reduce this proportion still further (**Figure S10**). At the
230 national level, susceptibility declines more slowly and more unevenly in such settings (e.g.,
231 Ethiopia, South Sudan, Tanzania) due to a lower probability of introductions and re-introductions
232 of the virus locally; an effect amplified by lockdowns (**Figure S10A**). It remains unclear whether
233 the more prolonged, asynchronous epidemics expected in these countries or the overlapping,
234 concurrent epidemics expected in countries with higher connectivity (e.g. Malawi, Kenya,
235 Burundi) will be a greater stress to health systems. Outbreak control efforts are likely to be
236 further complicated during prolonged epidemics if they intersect with seasonal events such as
237 temporal patterns in human mobility³¹ or other infections (e.g., malaria).

238

239 Turning to climate, despite extreme variation among cities in SSA (**Figure 5E**), large epidemic
240 peaks are expected in all cities (**Figure 5F**), even from models incorporating interventions and
241 transmission rates that decline in response to warmer, more humid local climates (climate
242 dependent variation in transmission rate for coronaviruses inferred from endemic circulation in
243 the US, but robust to parameter value choice; see methods). After accounting for differences in
244 the date of introductions, simulated climate forcing generates a maximum of only 6-7 weeks
245 variation in the time to epidemic peaks, with peaks generally expected earlier in more southerly,
246 colder, drier, cities (e.g., Windhoek and Maseru) and later in more humid, coastal cities (e.g.,
247 Bissau, Lomé, and Lagos). Reductions in transmission due to control efforts, as expected,
248 prolong the time to epidemic peak (**Figure 5F**). Apart from these slight shifts in timing, large
249 susceptible populations overwhelm the effects of climate²⁷, and earlier suggestions that Africa's

250 generally more tropical environment alone may provide a protective effect¹ are not supported by
251 evidence.

252

253 **Context-specific preparedness in SSA**

254 Our synthesis emphasizes striking country to country variation in drivers of the pandemic in SSA
255 (**Figure 3**), indicating continued variation in the burden (**Figure 4**) and pace (**Figure 5**) is to be
256 expected even across low income settings. As small perturbations in the age profile of mortality
257 could drastically change the national level burden in SSA (**Figure 4**), building expectations for
258 the risk for each country requires monitoring for deviations in the pattern of morbidity and
259 mortality over age. Transparent and timely communication of these context-specific risk patterns
260 could aid community engagement in efforts to reduce transmission, help motivate population
261 behavioral changes, and guide existing networks of community case management.

262

263 Because the largest impacts of SARS-CoV-2 outbreaks may be through indirect effects on
264 routine health provisioning, understanding how existing programs may be disrupted differently
265 by acute versus longer outbreaks is crucial to planning resource allocation. For example,
266 population immunity will decline proportionally with the length of disruptions to routine
267 vaccination programs ³⁶, resulting in more severe consequences in areas with prolonged
268 epidemic time courses.

269

270 Others have suggested that this crisis presents an opportunity to unify and mobilize across
271 existing health programs (e.g., for HIV, TB, Malaria, and other NCDs) ²⁶. While this may be a
272 powerful strategy in the context of acute, temporally confined crises, long term distraction and
273 diversion of resources ³⁷ may be harmful in settings with extended, asynchronous epidemics. A
274 higher risk of infection among healthcare workers during epidemics ^{38,39} may amplify this risk.

275

276 As evidenced by failures in locations where the epidemic progressed rapidly (e.g., USA),
277 effective governance and management prior to reaching large case counts likely yields the
278 largest rewards. Generalizing across SSA is difficult as the time course and estimates of the
279 effect of intervention policies have varied greatly (**Figure S10**), but Mauritius ⁴⁰ and Rwanda ⁴¹,
280 for example, have reported extremely low incidence thanks in part to a well-managed early
281 response.

282

283

284 **Conclusions**

285 The burden and time-course of SARS-CoV-2 is expected to be highly variable across sub-
286 Saharan Africa. Simulations show that variation in international and subnational connectivity are
287 expected to be important determinants of pace, but variability in reporting regimes makes it
288 difficult to compare observations to date with expectations (**Figure S7-8**). As the outbreak
289 continues to unfold, critically evaluating this mapping (e.g., **Figure S9**) can focus surveillance
290 efforts to areas expected to have prolonged epidemic trajectories and relatively high mortality
291 burdens. Additional immunological surveys and country-specific analyses of the age profile of
292 mortality are urgently needed in SSA and will likely be a powerful lens for understanding the
293 current landscape of population risk ⁴². When considering hopeful futures with the possibility of a
294 SARS-CoV-2 vaccine, it is imperative that vaccine distribution be equitable, and in proportion
295 with need. Understanding factors that both drive spatial variation in vulnerable populations and
296 temporal variation in pandemic progression could help approach these goals in SSA.

297

298 **Online Content**

299 Methods and additional figures are available in the supplementary materials. In addition, high resolution
300 maps and further visualizations of the risk indicators and simulations studied here can be accessed online
301 through an interactive tool:

302 Link to SSA-SARS-CoV-2 online companion tool: <https://labmetcalf.shinyapps.io/covid19-burden-africa/>

303

304 **Data Availability**

305 All data have been deposited into a publicly available GitHub repository:

306 Link to GitHub repository containing data and code: <https://github.com/labmetcalf/SSA-SARS-CoV-2>

307

308 **Code Availability**

309 All code has been deposited into the publicly available GitHub repository (same as above):

310 Link to GitHub repository containing data and code: <https://github.com/labmetcalf/SSA-SARS-CoV-2>

311

312 **Acknowledgements**

313 REB is supported by the Cooperative Institute for Modeling the Earth System (CIMES). AA acknowledges
314 support from the NIH Medical Scientist Training Program 1T32GM136577. AJT is funded by the BMGF
315 (OPP1182425, OPP1134076 and INV-002697). MB is funded by NWO Rubicon grant 019.192EN.017.

316

317 **Author contributions**

318 Conceptualization: BLR, AA, REB, MB, WWD, KM, IFM, NVM, AR, MR, JR, TR, FR, WY, BTG, CJT,
319 CJEM; Data curation: BLR, MR, MB, WWD, WY; Formal analysis: BLR, AA, MB, MR, REB; Methodology:
320 BLR, MR, MB, REB, CJEM, BTG; Software and Shiny app online tool: BLR, MR, MB, REB, WY;
321 Visualization: BLR, MR, MB, REB, WY; Writing – original draft: BLR, CJEM; Writing – reviewing and
322 editing: BLR, AA, REB, MB, WWD, KM, IFM, NVM, AR, MR, JR, TR, FR, WY, BTG, CJT, CJEM

323

324 **Additional Information**

325 Supplementary Information is available for this paper. Correspondence and requests for materials should
326 be addressed to BLR (b.rice@princeton.edu)

327

328 **Data and materials availability**

329 All materials are available in the online content

330

331 **Competing interests**

332 The authors declare no competing interests

333

334 Supplementary Materials Outline:

A1 Reported SARS-CoV-2 case counts, mortality, and testing in sub-Saharan Africa as of Sept. 2020

Table S1: Sub-Saharan Africa country codes, case counts, and testing

Figure S1: Variation between SSA countries in testing and reporting rates

A2 Synthesizing factors hypothesized to increase or decrease SARS-CoV-2 epidemic risk in SSA

Table S2: Dimensions of risk and expected direction of effect on SARS-CoV-2 transmission or burden in sub-Saharan Africa (SSA) relative to higher latitude countries

Table S3: Variables and data sources

Figure S2: Year of most recent data available for variables compared between global regions

Figure S3: Variation among sub-Saharan African countries in determinants of SARS-CoV-2 risk by variable (a subset of variables is shown in Figure 3 in the main text)

Figure S4: Variation among sub-Saharan African countries in determinants of SARS-CoV-2 mortality risk by category (subsets of variables are shown in Figure 4 in the main text)

Data File 1: Data for all compiled indicators

A3 Principal component analysis (PCA) of variables considered

Figure S5: PCAs of all variables and category specific subsets of variables

Data File 2: GDP, GINI Index, and tests completed data for PCA visualizations

A4 Evaluating the burden emerging from the severity of infection outcome

Table S4: Sources of age-stratified infection fatality ratio (*IFR*) estimates

Figure S6: Age profiles of comorbidities in sub-Saharan Africa countries

A5 International air travel to sub-Saharan Africa

Table S5: Arrivals to SSA airports by the number of passenger seats and status of the SARS-CoV-2 pandemic at the origin at the time of travel

A6 Interventions and subnational connectivity among countries in sub-Saharan Africa

Metapopulation model methods

Figure S7: Pace of the outbreak

Figure S8: Cases and testing vs. the pace of the outbreak

Figure S9: Bivariate example of expected pace versus expected burden at the national level in SARS-CoV-2 outbreaks in sub-Saharan Africa

Figure S10: Indicators of the impact of intervention policies and waning immunity, and their effects on spatial spread in sub-Saharan Africa

A7 Modeling epidemic trajectories in scenarios where transmission rate depends on climate

Data on climate variation in SSA

Climate model methods

Sensitivity analyses

Figure S11: Transmission climate-dependency and sensitivity to R_{0max} and R_{0min} value selection

335

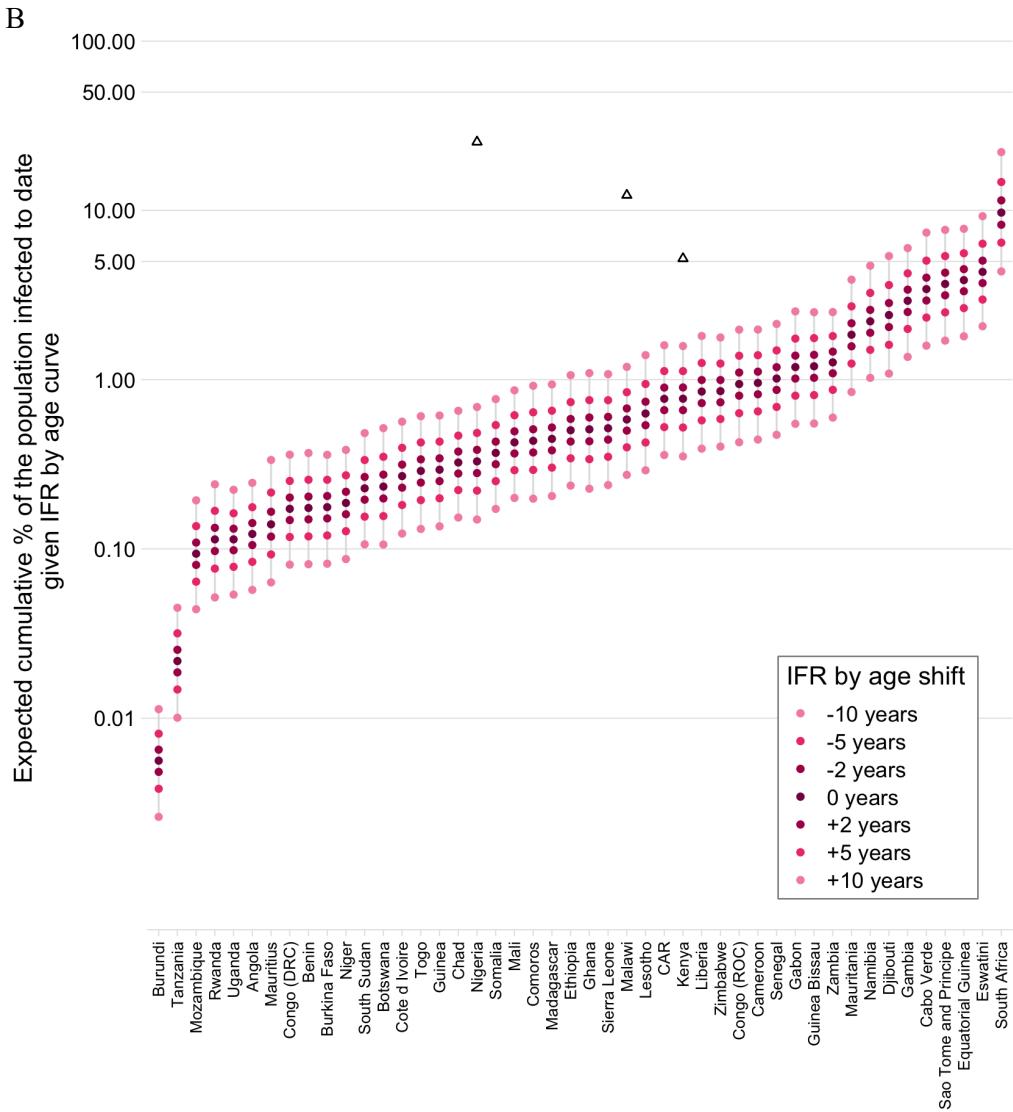
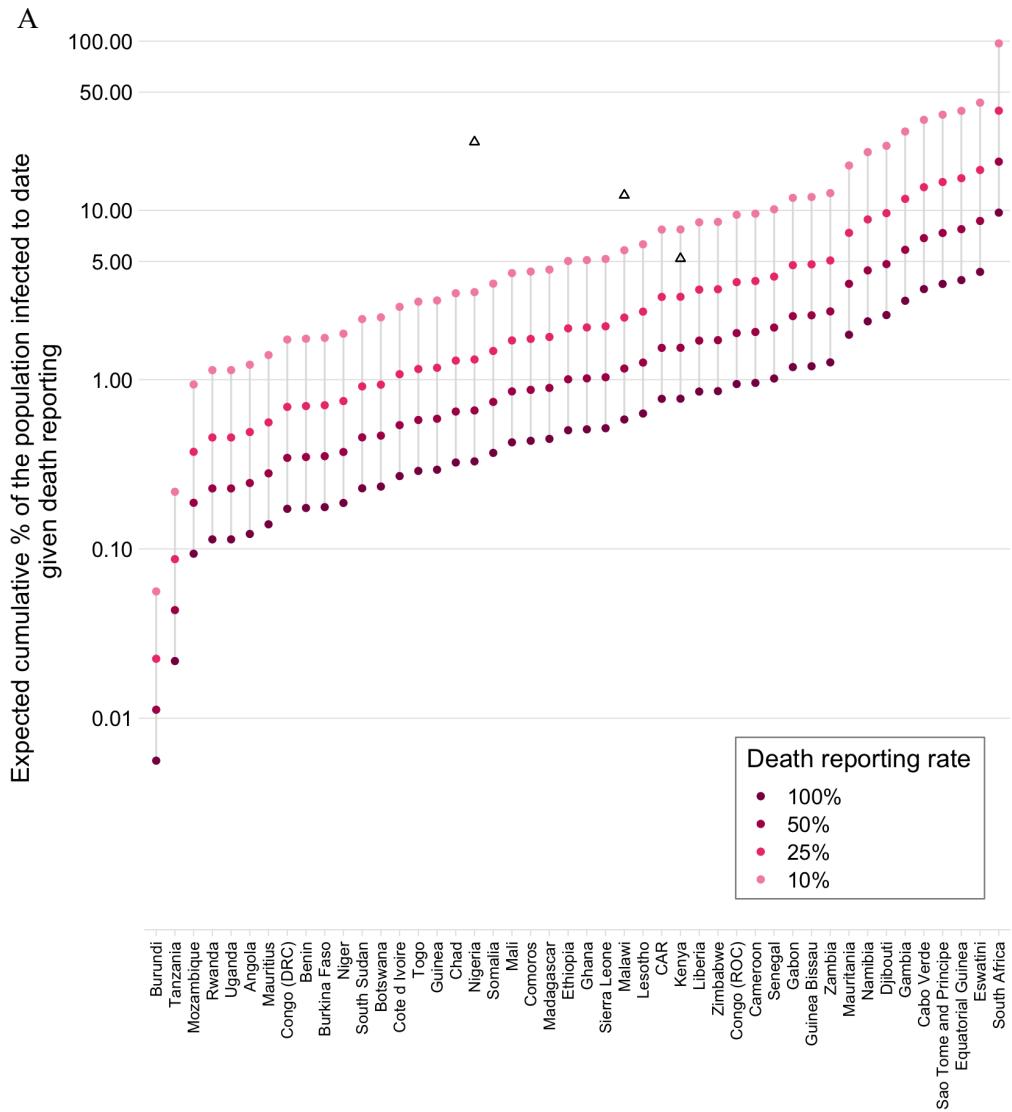


Figure 1 | Variation in the cumulative percentage of the population infected in sub-Saharan Africa countries as expected from reported mortality totals

The expected percentage of a country's population infected given the number of reported deaths to date, country-specific age structure, and a range of death reporting completeness scenarios (panel A), or a range of infection fatality ratio (*IFR*) scenarios (panel B). The global *IFR*-age curves were fit to existing age-stratified *IFR* estimates (see methods, **Table S4**) and, for panel B, shifted towards younger or older ages by the specified number of years to simulate higher or lower *IFRs*, respectively. Conservatively, we assume no variation in infection rates by age (infections skewed towards older age groups would result in a higher average *IFR* and thus

a lower expected percentage of the population infected for a given number of deaths). Reported case and death counts current as of September 2020 (sourced from Africa CDC, see **Table S1**); Data from Eritrea and the Seychelles not shown as they have zero reported deaths as of September 2020. Comparisons to serological surveys (unfilled triangles) available from blood-banks in Kenya ⁷, health care workers in urban Malawi ⁸, and a subnational cluster-stratified random sample from Niger State in Nigeria ⁹ are shown.

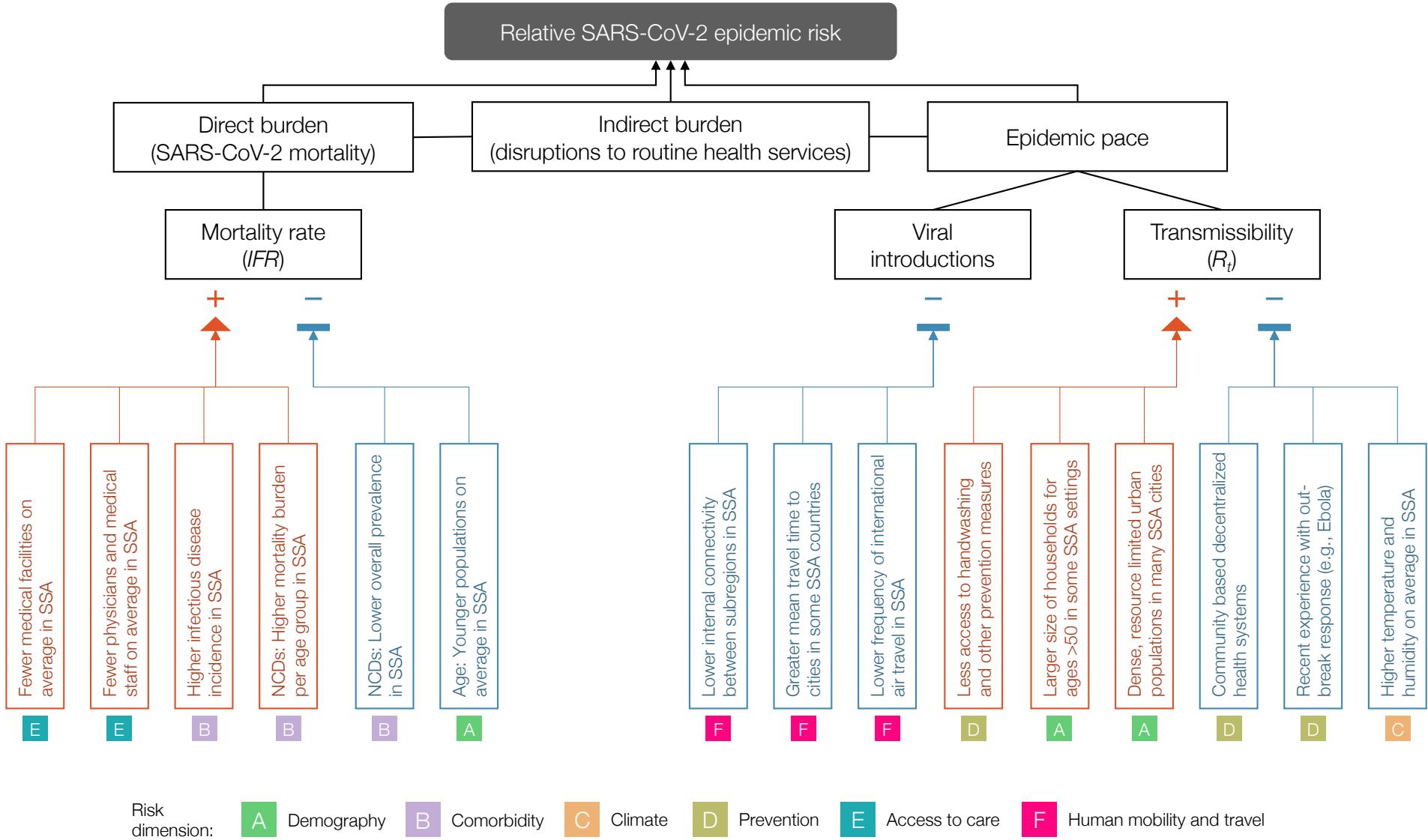
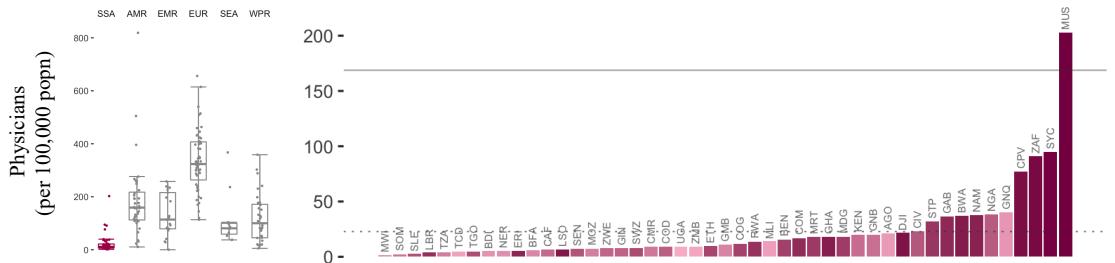


Figure 2 | Hypothesized modulators of relative SARS-CoV-2 epidemic risk in sub-Saharan Africa

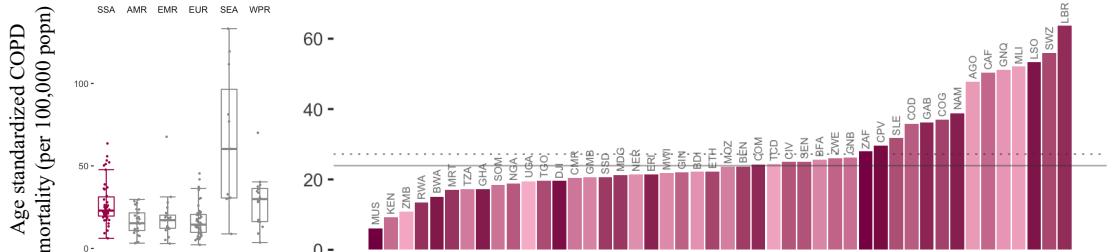
Factors hypothesized to increase (red) or decrease (blue) mortality burden or epidemic pace within sub-Saharan Africa, relative to global averages, are grouped in six categories or dimensions of risk (A-F). In this framework, epidemic pace is determined by person to person transmissibility (which can be defined as the time-varying effective reproductive number, R_t) and introduction and geographic spread of the virus via human mobility.

SARS-CoV-2 mortality (determined by the infection fatality ratio, IFR) is modulated by demography, comorbidities (e.g., non-communicable diseases (NCDs)), and access to care. Overall burden is a function of direct burden and indirect effects due to, for example, socio-economic disruptions and disruptions in health services such as vaccination and infectious disease control. **Table S2** contains details and the references used as a basis to draw the hypothesized modulating pathways.

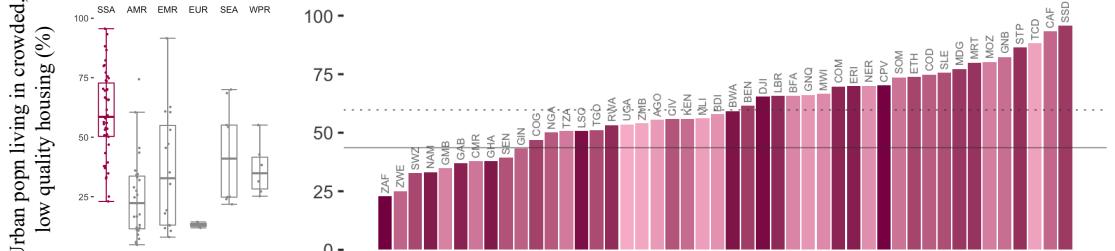
A



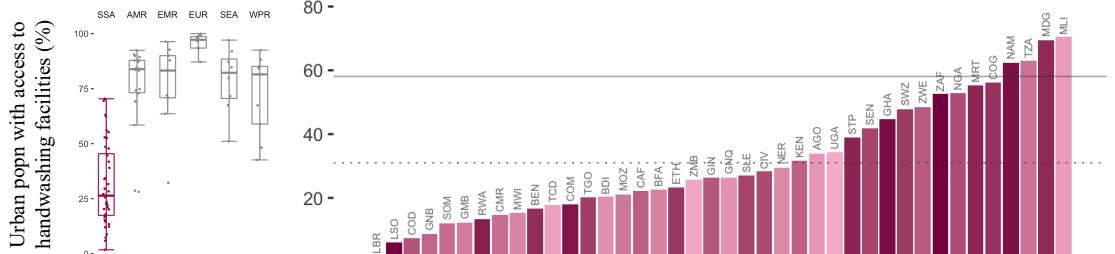
B



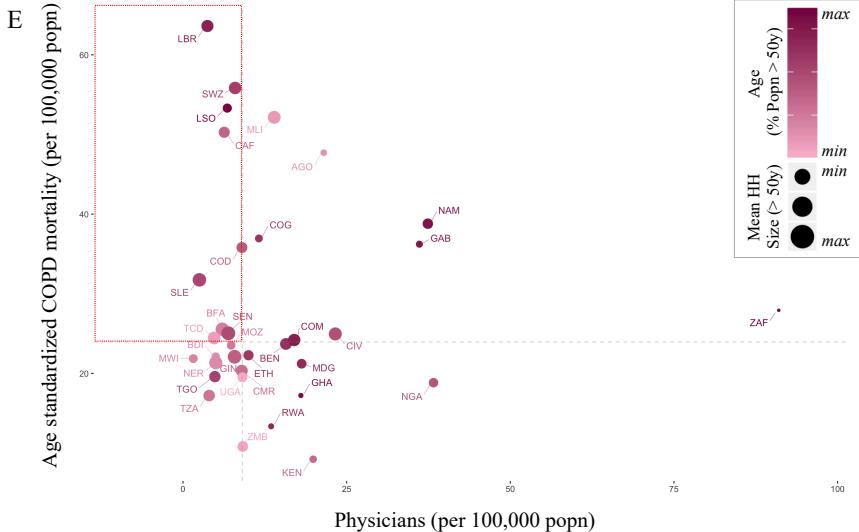
C



D



E



F

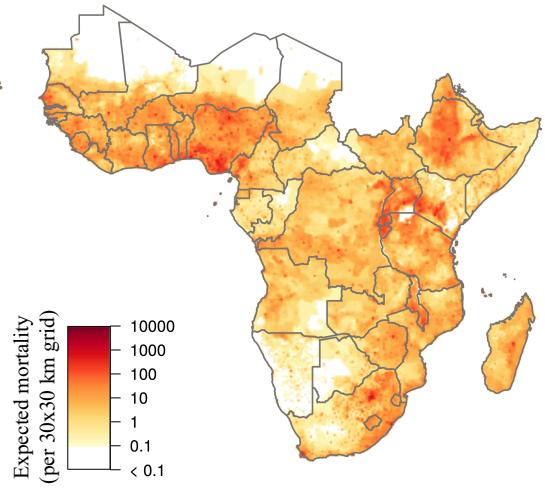


Figure 3 | Variation among sub-Saharan African countries in select determinants of SARS-CoV-2 risk

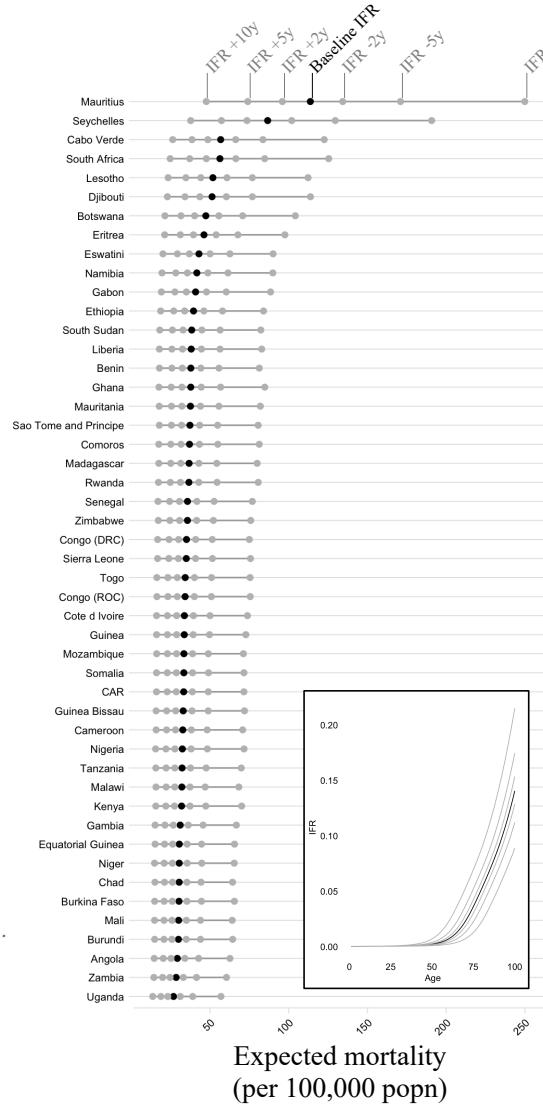
A-D: At right, SSA countries are ranked from least to greatest for each indicator; bar color shows population age structure (% of the population above age 50). Solid horizontal lines show the global mean value; dotted lines show the mean among SSA countries. At left, boxplots show median and interquartile range, grouped by geographic region, per WHO: sub-Saharan Africa (SSA); Americas Region (AMR); Eastern Mediterranean Region (EMR); Europe Region (EUR); Southeast Asia Region (SEA); Western Pacific Region (WPR).

E-F: Dot size shows mean household (HH) size for HHs with individuals over age 50; dashed lines show median value among SSA countries; quadrants of greatest risk are outlined in red (e.g., fewer physicians and greater age standardized Chronic Obstructive Pulmonary Disease (COPD) mortality). See Table S3, Figure S3, and the [\[SSA-SARS-CoV-2-tool\]](#) for full description and visualization of all variables.

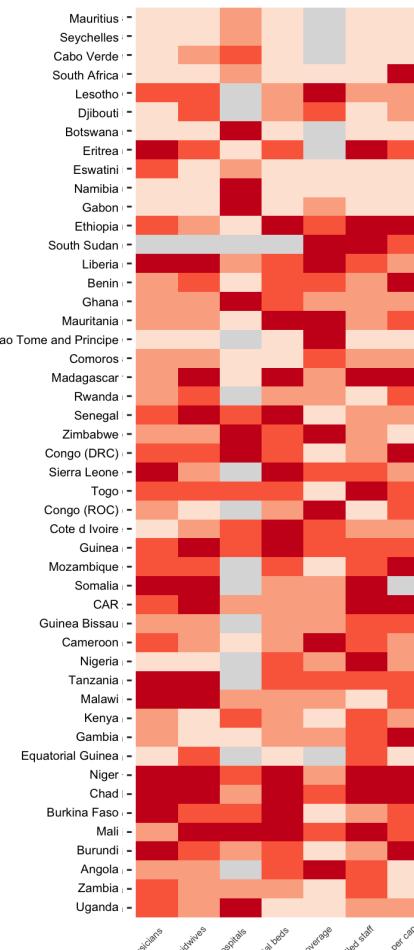
A | Baseline mortality risk from demographic structure



C | Range in mortality under simulated *IFR* scenarios



D | Indicators of access to care at national level



E | Indicators of comorbidity burden at national level

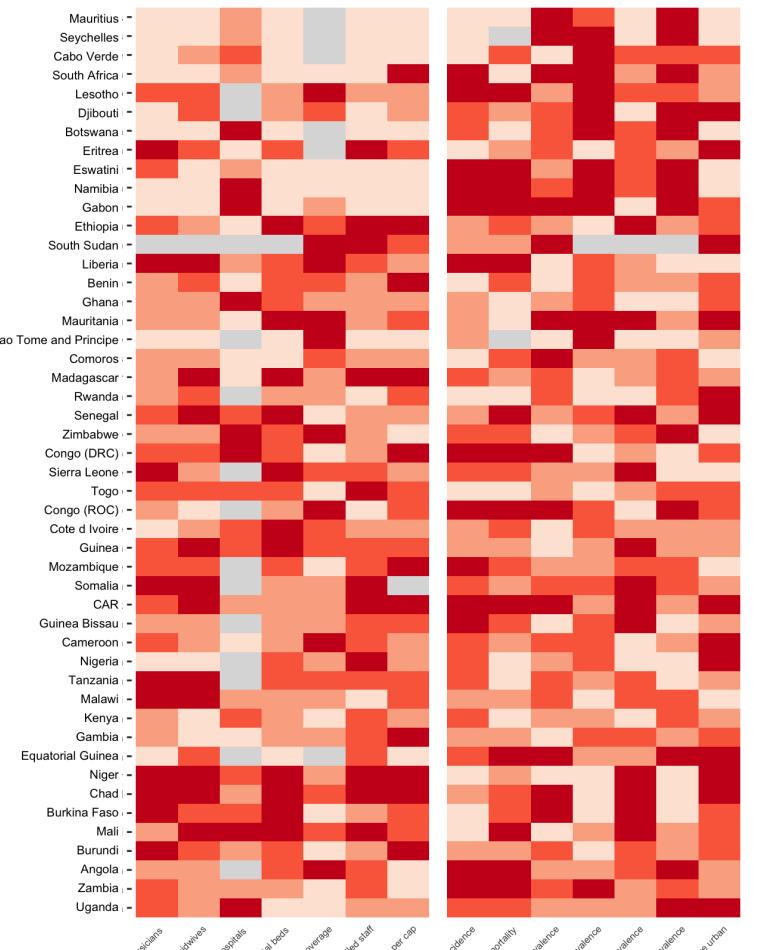


Figure 4 | Variation in expected burden for SARS-CoV-2 outbreaks in sub-Saharan Africa

A: Expected mortality in a scenario where cumulative infection reaches 20% across age groups and the infection fatality ratio (*IFR*) curve is fit to existing age-stratified *IFR* estimates (see methods, **Table S4**). **B:** National level variation in comorbidity and access to care variables, for e.g., diabetes prevalence among adults and the number of hospital beds per 100,000 population for sub-Saharan African countries. **C:** The range in mortality per 100,000 population expected in scenarios where cumulative infection rate is 20% and *IFR* per age is the baseline (black) or shifted \pm 2, 5, or 10 years (gray). Inset, the *IFR* by age curves for each scenario.

D-E: Select national level indicators; estimates of reduced access to care (e.g., fewer hospitals) or increased comorbidity burden (e.g., higher prevalence of raised blood pressure) shown with darker red for higher risk quartiles (see **Figure S4** for all indicators). Countries missing data for an indicator (NA) are shown in gray. For comparison between countries, estimates are age-standardized where applicable (see **Table S3** for details). See the [[SSA-SARS-CoV-2-tool](#)] for high resolution maps for each variable and scenario.

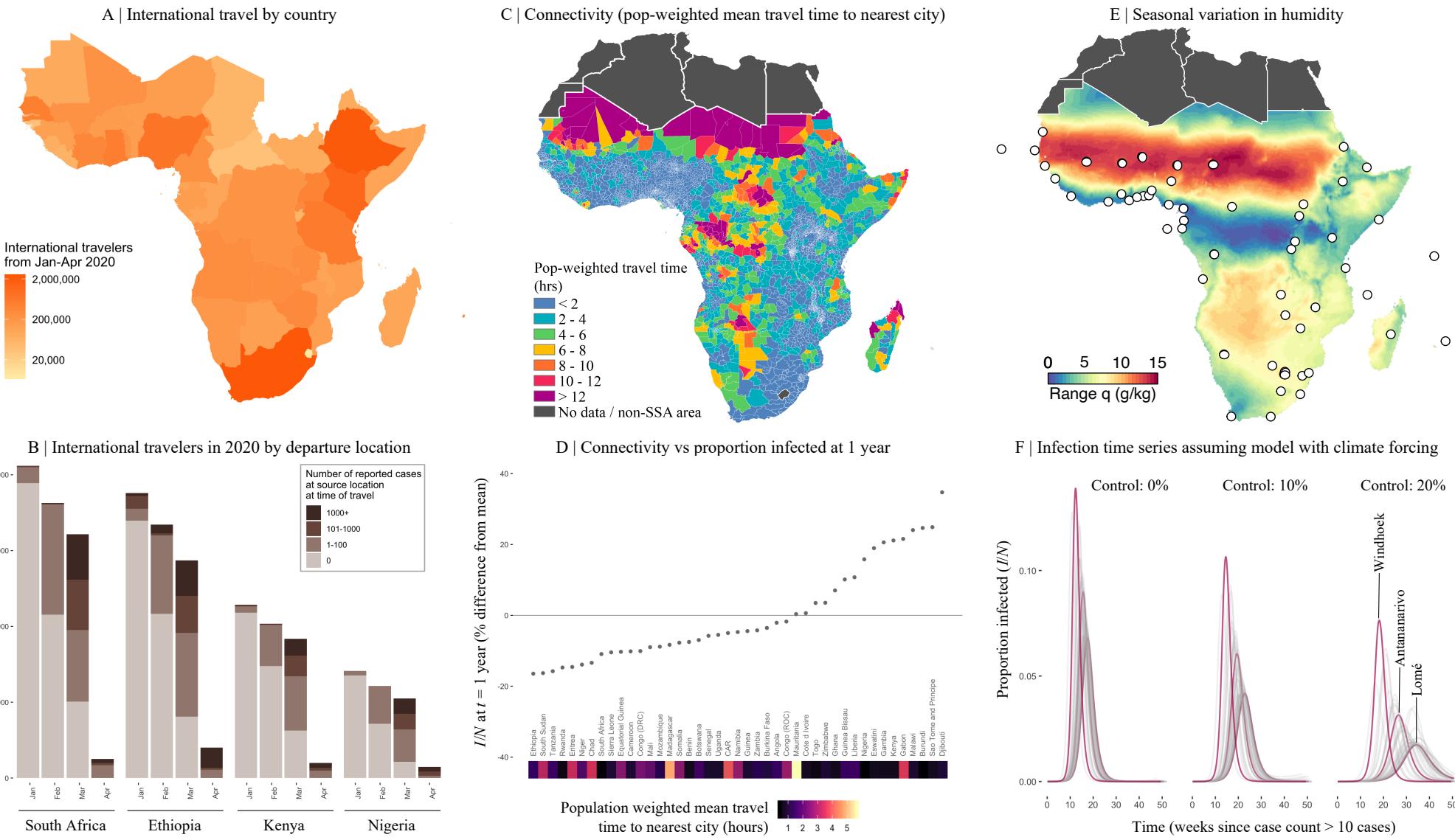


Figure 5 | Variation in connectivity and climate in sub-Saharan Africa and expected effects on SARS-CoV-2

A: International travelers to sub-Saharan Africa (SSA) from January to April 2020, as inferred from the number of passenger seats on arriving aircraft. **B:** For the four countries with the most arrivals, the proportion of arrivals by month coming from countries with 0, 1-100, 101-1000, and 1000+ reported SARS-CoV-2 infections at the time of travel (see Table S5 for all others).

C: Connectivity within SSA countries as inferred from average population weighted mean travel time to the nearest urban area greater than 50,000 population.

D: Mean travel time at the national level and variation in the fraction of the population expected to be infected (I/N) in the first year from stochastic simulations (see methods). **E:** Climate variation across SSA as shown by seasonal range in specific humidity, q (g/kg) (max average q - min average q). **F:** The effect of local seasonality and control efforts (R_0 decreases by 0%, i.e., unmitigated, 10%, or 20%) on the timing of epidemic peaks (max I/N) in SSA cities (with three exemplar cities highlighted in pink, see methods)

336 **Figure legends:**

337

338 **Figure 1**

339 **Variation in the cumulative percentage of the population infected in sub-Saharan Africa**
 340 **countries as expected from reported mortality totals**

341 The expected percentage of a country's population infected given the number of reported deaths to date, country-
 342 specific age structure, and a range of death reporting completeness scenarios (panel **A**), or a range of infection
 343 fatality ratio (*IFR*) scenarios (panel **B**). The global *IFR*-age curves were fit to existing age-stratified *IFR* estimates (see
 344 methods, **Table S4**) and, for panel B, shifted towards younger or older ages by the specified number of years to
 345 simulate higher or lower *IFRs*, respectively. Conservatively, we assume no variation in infection rates by age
 346 (infections skewed towards older age groups would result in a higher average *IFR* and thus a lower expected
 347 percentage of the population infected for a given number of deaths). Reported case and death counts current as of
 348 September 2020 (sourced from Africa CDC, see **Table S1**); Data from Eritrea and the Seychelles not shown as they
 349 have zero reported deaths as of September 2020. Comparisons to serological surveys (unfilled triangles) available
 350 from blood-banks in Kenya ⁷, health care workers in urban Malawi ⁹, and a subnational cluster-stratified random
 351 sample from Niger State in Nigeria ⁸ are shown.

352

353 **Figure 2**

354 **Hypothesized modulators of relative SARS-CoV-2 epidemic risk in sub-Saharan Africa**

355 Factors hypothesized to increase (red) or decrease (blue) mortality burden or epidemic pace within sub-Saharan
 356 Africa, relative to global averages, are grouped in six categories or dimensions of risk (A-F). In this framework,
 357 epidemic pace is determined by person to person transmissibility (which can be defined as the time-varying effective
 358 reproductive number, R_t) and introduction and geographic spread of the virus via human mobility. SARS-CoV-2
 359 mortality (determined by the infection fatality ratio, *IFR*) is modulated by demography, comorbidities (e.g., non-
 360 communicable diseases (NCDs)), and access to care. Overall burden is a function of direct burden and indirect
 361 effects due to, for example, socio-economic disruptions and disruptions in health services such as vaccination and
 362 infectious disease control. **Table S2** contains details and the references used as a basis to draw the hypothesized
 363 modulating pathways.

364

365 **Figure 3**

366 **Variation among sub-Saharan African countries in select determinants of SARS-CoV-2**
 367 **risk**

368 **A-D:** At right, SSA countries are ranked from least to greatest for each indicator; bar color shows population age
 369 structure (% of the population above age 50). Solid horizontal lines show the global mean value; dotted lines show
 370 the mean among SSA countries. At left, boxplots show median and interquartile range, grouped by geographic
 371 region, per WHO: sub-Saharan Africa (SSA); Americas Region (AMR); Eastern Mediterranean Region (EMR);
 372 Europe Region (EUR); Southeast Asia Region (SEA); Western Pacific Region (WPR). **E-F:** Dot size shows mean
 373 household (HH) size for HHs with individuals over age 50; dashed lines show median value among SSA countries;
 374 quadrants of greatest risk are outlined in red (e.g., fewer physicians and greater age standardized Chronic
 375 Obstructive Pulmonary Disease (COPD) mortality). See Table S3, Figure S3, and the [[SSA-SARS-CoV-2-tool](#)] for full
 376 description and visualization of all variables.

377

378

379 **Figure 4**

380 **Variation in expected burden for SARS-CoV-2 outbreaks in sub-Saharan Africa**

381 A: Expected mortality in a scenario where cumulative infection reaches 20% across age groups and the infection
 382 fatality ratio (*IFR*) curve is fit to existing age-stratified *IFR* estimates (see methods, **Table S4**). B: National level
 383 variation in comorbidity and access to care variables, for e.g., diabetes prevalence among adults and the number of
 384 hospital beds per 100,000 population for sub-Saharan African countries. C: The range in mortality per 100,000
 385 population expected in scenarios where cumulative infection rate is 20% and *IFR* per age is the baseline (black) or
 386 shifted \pm 2, 5, or 10 years (gray). Inset, the *IFR* by age curves for each scenario. D-E: Select national level indicators;
 387 estimates of reduced access to care (e.g., fewer hospitals) or increased comorbidity burden (e.g., higher prevalence
 388 of raised blood pressure) shown with darker red for higher risk quartiles (see **Figure S4** for all indicators). Countries
 389 missing data for an indicator (NA) are shown in gray. For comparison between countries, estimates are age-
 390 standardized where applicable (see **Table S3** for details). See the [[SSA-SARS-CoV-2-tool](#)] for high resolution maps
 391 for each variable and scenario.

392

393 **Figure 5**

394 **Variation in connectivity and climate in sub-Saharan Africa and expected effects on**
 395 **SARS-CoV-2**

396 A: International travellers to sub-Saharan Africa (SSA) from January to April 2020, as inferred from the number of
 397 passenger seats on arriving aircraft. B: For the four countries with the most arrivals, the proportion of arrivals by
 398 month coming from countries with 0, 1-100, 101-1000, and 1000+ reported SARS-CoV-2 infections at the time of
 399 travel (see **Table S5** for all others). C: Connectivity within SSA countries as inferred from average population
 400 weighted mean travel time to the nearest urban area greater than 50,000 population. D: Mean travel time at the
 401 national level and variation in the fraction of the population expected to be infected (*I/N*) in the first year from
 402 stochastic simulations (see methods). E: Climate variation across SSA as shown by seasonal range in specific
 403 humidity, q (g/kg) (max average q - min average q). F: The effect of local seasonality and control efforts (R_0
 404 decreases by 0%, i.e., unmitigated, 10%, or 20%) on the timing of epidemic peaks (max *I/N*) in SSA cities (with three
 405 exemplar cities highlighted in pink, see methods).

406

407

408 **A1 | Reported SARS-CoV-2 case counts, mortality, and testing in sub-**
409 **Saharan Africa as of September 2020**

410
411 *1.1. Variables and data sources for testing data*

412
413 The numbers of reported cases, deaths, and tests for the 48 studied sub-Saharan Africa (SSA)
414 countries (**Table S1**) were sourced from the Africa Centers for Disease Control (CDC)
415 dashboard on September 23, 2020 (and previously on June 30, 2020)
416 (<https://africacdc.org/covid-19/>). Africa CDC obtains data from the official Africa CDC Regional
417 Collaborating Centre and member state reports. Differences in the timing of reporting by
418 member states results in some variation in recency of data within the centralized Africa CDC
419 repository, but the data should broadly reflect the relative scale of testing and reporting efforts
420 across countries.

421
422 The countries or member states within SSA in this study follow the United Nations and Africa
423 CDC listed regions of Southern, Western, Central, and Eastern Africa (not including Sudan).
424 From the Northern Africa region, Mauritania is included in SSA.

425
426 For comparison to non-SSA countries, the number of reported cases in other geographic
427 regions were obtained from the Johns Hopkins University Coronavirus Resource Center on
428 September 23, 2020 (<https://coronavirus.jhu.edu/map.html>).

429
430 Case fatality ratios (CFRs) were calculated by dividing the number of reported deaths by the
431 number of reported cases and expressed as a percentage. Positivity was calculated by dividing
432 the number of reported cases by the number of reported tests. Testing and case rates were
433 calculated per 100,000 population using population size estimates for 2020 from the United
434 Nations Population Division ⁴³. As reported confirmed cases are likely to be a significant
435 underestimate of the true number of infections, CFRs may be a poor proxy for the infection
436 fatality ratio (IFR), defined as the proportion of infections that result in mortality ⁴.

437
438 *1.2 Variation in testing and mortality rates*

439
440 Testing rates among SSA countries varied by multiple orders of magnitude as of June 2020 and
441 remain highly variable as of September 2020. The number of tests completed per 100,000
442 population ranged from 19.84 in Burundi to 13,508.13 in Mauritius in June 2020 and from 65.98
443 in Congo (DRC) to 18,321.83 in Mauritius in September 2020 (**Figure S1A**). Tanzania (6.50
444 tests per 100,000 population) has not reported new tests, cases or deaths to the Africa CDC
445 since April 2020. The number of reported infections (i.e., positive tests) was strongly correlated
446 with the number of tests completed both in June 2020 (Pearson's correlation coefficient, $r =$
447 0.9667, $p < 0.001$) and in September 2020 ($r = 0.9689$, $p < 0.001$) (**Figure S1B**). As of June
448 2020, no deaths due to SARS-CoV-2 were reported to the Africa CDC for five SSA countries
449 (Eritrea, Lesotho, Namibia, Seychelles, Uganda). As of September 2020, still no deaths due to
450 SARS-CoV-2 were reported to the Africa CDC for two of those countries (Eritrea and
451 Seychelles). Among countries with at least one reported death, CFR varied from 0.22% in

452 Rwanda to 8.54% in Chad in June 2020 and from 0.21% in Burundi to 6.96% in Chad in
453 September 2020 (**Figure S1C**). Limitations in the ascertainment of infection rates and the rarity
454 of reported deaths (e.g., median number of reported deaths per SSA country was 25.5 as of
455 June 2020 and 71.0 as of September 2020), indicate that the data are insufficient to determine
456 country specific *IFRs* and *IFR* by age profiles for most countries. As a result, global *IFR* by age
457 estimates were used for the subsequent analyses in this study.

458

Table S1

459

**Sub-Saharan Africa country country codes, case counts, and testing as of September 23,
460 2020**

| Country Name | Country Code | Cases ^a | Deaths ^a | Tests ^a | Population ^b | Cases per 100k ^c | Tests per 100k ^c | Positivity (%) | CFR (%) |
|-------------------------|--------------|--------------------|---------------------|--------------------|-------------------------|-----------------------------|-----------------------------|----------------|---------|
| Angola | AGO | 4363 | 59 | 67623 | 32866268 | 13.28 | 205.75 | 6.45 | 1.35 |
| Benin | BEN | 2294 | 40 | 144210 | 12123198 | 18.92 | 1189.54 | 1.59 | 1.74 |
| Botswana | BWA | 2567 | 13 | 156054 | 2351625 | 109.16 | 6636.01 | 1.64 | 0.51 |
| Burkina Faso | BFA | 1929 | 56 | 43607 | 20903278 | 9.23 | 208.61 | 4.42 | 2.90 |
| Burundi | BDI | 476 | 1 | 34723 | 11890781 | 4.00 | 292.02 | 1.37 | 0.21 |
| Cameroon | CMR | 20690 | 416 | 453441 | 26545864 | 77.94 | 1708.14 | 4.56 | 2.01 |
| Cabo Verde | CPV | 5412 | 54 | 78957 | 555988 | 973.40 | 14201.21 | 6.85 | 1.00 |
| Central Africa Republic | CAF | 4802 | 62 | 31723 | 4829764 | 99.43 | 656.82 | 15.14 | 1.29 |
| Chad | TCD | 1164 | 81 | 26741 | 16425859 | 7.09 | 162.80 | 4.35 | 6.96 |
| Comoros | COM | 470 | 7 | 3025 | 869595 | 54.05 | 347.86 | 15.54 | 1.49 |
| Côte d'Ivoire | CIV | 19430 | 120 | 125972 | 26378275 | 73.66 | 477.56 | 15.42 | 0.62 |
| Congo (DRC) | COD | 10555 | 271 | 59093 | 89561404 | 11.79 | 65.98 | 17.86 | 2.57 |
| Djibouti | DJI | 5407 | 61 | 75166 | 988002 | 547.27 | 7607.88 | 7.19 | 1.13 |
| Equatorial Guinea | GNQ | 5018 | 83 | 58603 | 1402985 | 357.67 | 4177.02 | 8.56 | 1.65 |
| Eritrea | ERI | 364 | 0 | 12269 | 3546427 | 10.26 | 345.95 | 2.97 | 0.00 |
| Eswatini | SWZ | 5343 | 108 | 42834 | 1160164 | 460.54 | 3692.06 | 12.47 | 2.02 |
| Ethiopia | ETH | 71083 | 1141 | 1226297 | 114963583 | 61.83 | 1066.68 | 5.80 | 1.61 |
| Gabon | GAB | 8716 | 54 | 157828 | 2225728 | 391.60 | 7091.07 | 5.52 | 0.62 |
| Gambia | GMB | 3542 | 110 | 16759 | 2416664 | 146.57 | 693.48 | 21.13 | 3.11 |
| Ghana | GHA | 46153 | 299 | 474120 | 31072945 | 148.53 | 1525.83 | 9.73 | 0.65 |
| Guinea | GIN | 10434 | 65 | 82365 | 13132792 | 79.45 | 627.17 | 12.67 | 0.62 |
| Guinea Bissau | GNB | 2303 | 39 | 1846 | 1967998 | 117.02 | 93.80 | 124.76 | 1.69 |
| Kenya | KEN | 37348 | 664 | 523998 | 53771300 | 69.46 | 974.49 | 7.13 | 1.78 |
| Lesotho | LSO | 1554 | 35 | 18216 | 2142252 | 72.54 | 850.32 | 8.53 | 2.25 |
| Liberia | LBR | 1338 | 82 | 17874 | 5057677 | 26.45 | 353.40 | 7.49 | 6.13 |

461

462

463

(Table S1 continued) September 23, 2020

| Country Name | Country Code | Cases ^a | Deaths ^a | Tests ^a | Population ^b | Cases per 100k ^c | Tests per 100k ^c | Positivity (%) | CFR (%) |
|-----------------------|--------------|--------------------|---------------------|--------------------|-------------------------|-----------------------------|-----------------------------|----------------|---------|
| Madagascar | MDG | 16191 | 227 | 76137 | 27691019 | 58.47 | 274.95 | 21.27 | 1.40 |
| Malawi | MWI | 5746 | 179 | 51456 | 19129955 | 30.04 | 268.98 | 11.17 | 3.12 |
| Mali | MLI | 3034 | 130 | 40733 | 20250834 | 14.98 | 201.14 | 7.45 | 4.28 |
| Mauritania | MRT | 7425 | 161 | 78815 | 4649660 | 159.69 | 1695.07 | 9.42 | 2.17 |
| Mauritius | MUS | 367 | 10 | 233011 | 1271767 | 28.86 | 18321.83 | 0.16 | 2.72 |
| Mozambique | MOZ | 7262 | 49 | 128805 | 31255435 | 23.23 | 412.10 | 5.64 | 0.67 |
| Namibia | NAM | 10663 | 117 | 91902 | 2540916 | 419.65 | 3616.88 | 11.60 | 1.10 |
| Niger | NER | 1193 | 69 | 16582 | 24206636 | 4.93 | 68.50 | 7.19 | 5.78 |
| Nigeria | NGA | 57724 | 1102 | 484051 | 206139587 | 28.00 | 234.82 | 11.93 | 1.91 |
| Congo (ROC) | COG | 5005 | 89 | 35219 | 5518092 | 90.70 | 638.25 | 14.21 | 1.78 |
| Rwanda | RWA | 4779 | 27 | 481283 | 12952209 | 36.90 | 3715.84 | 0.99 | 0.56 |
| São Tomé and Príncipe | STP | 908 | 15 | 6193 | 219161 | 414.31 | 2825.78 | 14.66 | 1.65 |
| Senegal | SEN | 14816 | 304 | 170087 | 16743930 | 88.49 | 1015.81 | 8.71 | 2.05 |
| Seychelles | SYC | 143 | 0 | 5200 | 98340 | 145.41 | 5287.78 | 2.75 | 0.00 |
| Sierra Leone | SLE | 2183 | 72 | 25976 | 7976985 | 27.37 | 325.64 | 8.40 | 3.30 |
| Somalia | SOM | 3465 | 98 | 20516 | 15893219 | 21.80 | 129.09 | 16.89 | 2.83 |
| South Africa | ZAF | 665188 | 16206 | 4083757 | 59308690 | 1121.57 | 6885.60 | 16.29 | 2.44 |
| South Sudan | SSD | 2664 | 49 | 19433 | 11193729 | 23.80 | 173.61 | 13.71 | 1.84 |
| Tanzania ^d | TZA | 509 | 21 | 3880 | 59734213 | 0.85 | 6.50 | 13.12 | 4.13 |
| Togo | TGO | 1701 | 41 | 85422 | 8278737 | 20.55 | 1031.82 | 1.99 | 2.41 |
| Uganda | UGA | 7064 | 70 | 463618 | 45741000 | 15.44 | 1013.57 | 1.52 | 0.99 |
| Zambia | ZMB | 14491 | 332 | 146329 | 18383956 | 78.82 | 795.96 | 9.90 | 2.29 |
| Zimbabwe | ZWE | 7725 | 227 | 63185 | 14862927 | 51.97 | 425.12 | 12.23 | 2.94 |

464

^a Data from Africa CDC as of September 23, 2020 (<https://africacdc.org/covid-19/>)^b Data from UN Population Division UNPOP (2019 revision) estimates of population by single calendar year (2020)⁴³^c Rates per 100,000 population^d Tanzania has not reported tests, cases, or deaths since April 29, 2020

465

466

467

468

469 **Table S1**470 **Sub-Saharan Africa country country codes, case counts, and testing as of June 30, 2020**

| Country Name | Country Code | Cases ^a | Deaths ^a | Tests ^a | Population ^b | Cases per 100k ^c | Tests per 100k ^c | Positivity (%) | CFR (%) |
|-------------------------|--------------|--------------------|---------------------|--------------------|-------------------------|-----------------------------|-----------------------------|----------------|---------|
| Angola | AGO | 267 | 11 | 22895 | 32866268 | 0.81 | 69.66 | 1.17 | 4.12 |
| Benin | BEN | 1187 | 19 | 20014 | 12123198 | 9.79 | 165.09 | 5.93 | 1.60 |
| Botswana | BWA | 89 | 1 | 36868 | 2351625 | 3.78 | 1567.77 | 0.24 | 1.12 |
| Burkina Faso | BFA | 959 | 53 | 9040 | 20903278 | 4.59 | 43.25 | 10.61 | 5.53 |
| Burundi | BDI | 170 | 1 | 2359 | 11890781 | 1.43 | 19.84 | 7.21 | 0.59 |
| Cameroon | CMR | 12592 | 313 | 80000 | 26545864 | 47.43 | 301.37 | 15.74 | 2.49 |
| Cabo Verde | CPV | 1165 | 12 | 22665 | 555988 | 209.54 | 4076.53 | 5.14 | 1.03 |
| Central Africa Republic | CAF | 3429 | 45 | 23208 | 4829764 | 71.00 | 480.52 | 14.78 | 1.31 |
| Chad | TCD | 866 | 74 | 4633 | 16425859 | 5.27 | 28.21 | 18.69 | 8.55 |
| Comoros | COM | 293 | 7 | 1173 | 869595 | 33.69 | 134.89 | 24.98 | 2.39 |
| Côte d'Ivoire | CIV | 9101 | 66 | 48340 | 26378275 | 34.50 | 183.26 | 18.83 | 0.73 |
| Congo (DRC) | COD | 6939 | 167 | 24657 | 89561404 | 7.75 | 27.53 | 28.14 | 2.41 |
| Djibouti | DJI | 4656 | 53 | 46108 | 988002 | 471.25 | 4666.79 | 10.10 | 1.14 |
| Equatorial Guinea | GNQ | 2001 | 32 | 16000 | 1402985 | 142.62 | 1140.43 | 12.51 | 1.60 |
| Eritrea | ERI | 191 | 0 | 7943 | 3546427 | 5.39 | 223.97 | 2.40 | 0.00 |
| Eswatini | SWZ | 781 | 11 | 11094 | 1160164 | 67.32 | 956.24 | 7.04 | 1.41 |
| Ethiopia | ETH | 5846 | 103 | 250604 | 114963583 | 5.09 | 217.99 | 2.33 | 1.76 |
| Gabon | GAB | 5209 | 40 | 34774 | 2225728 | 234.04 | 1562.37 | 14.98 | 0.77 |
| Gambia | GMB | 45 | 2 | 2947 | 2416664 | 1.86 | 121.94 | 1.53 | 4.44 |
| Ghana | GHA | 17351 | 112 | 294867 | 31072945 | 55.84 | 948.95 | 5.88 | 0.65 |
| Guinea | GIN | 5291 | 30 | 33737 | 13132792 | 40.29 | 256.89 | 15.68 | 0.57 |
| Guinea Bissau | GNB | 1614 | 21 | 8056 | 1967998 | 82.01 | 409.35 | 20.03 | 1.30 |
| Kenya | KEN | 6190 | 144 | 167417 | 53771300 | 11.51 | 311.35 | 3.70 | 2.33 |
| Lesotho | LSO | 27 | 0 | 3000 | 2142252 | 1.26 | 140.04 | 0.90 | 0.00 |
| Liberia | LBR | 768 | 34 | 6125 | 5057677 | 15.18 | 121.10 | 12.54 | 4.43 |

471

472

473

474

(Table S1 continued) June 30, 2020

| Country Name | Country Code | Cases ^a | Deaths ^a | Tests ^a | Population ^b | Cases per 100k ^c | Tests per 100k ^c | Positivity (%) | CFR (%) |
|-----------------------|--------------|--------------------|---------------------|--------------------|-------------------------|-----------------------------|-----------------------------|----------------|---------|
| Madagascar | MDG | 2138 | 20 | 21444 | 27691019 | 7.72 | 77.44 | 9.97 | 0.94 |
| Malawi | MWI | 1152 | 13 | 13369 | 19129955 | 6.02 | 69.89 | 8.62 | 1.13 |
| Mali | MLI | 2147 | 113 | 12869 | 20250834 | 10.60 | 63.55 | 16.68 | 5.26 |
| Mauritania | MRT | 4149 | 126 | 39398 | 4649660 | 89.23 | 847.33 | 10.53 | 3.04 |
| Mauritius | MUS | 341 | 10 | 171792 | 1271767 | 26.81 | 13508.13 | 0.20 | 2.93 |
| Mozambique | MOZ | 859 | 5 | 28586 | 31255435 | 2.75 | 91.46 | 3.00 | 0.58 |
| Namibia | NAM | 183 | 0 | 8706 | 2540916 | 7.20 | 342.63 | 2.10 | 0.00 |
| Niger | NER | 1074 | 67 | 6555 | 24206636 | 4.44 | 27.08 | 16.38 | 6.24 |
| Nigeria | NGA | 24567 | 565 | 130164 | 206139587 | 11.92 | 63.14 | 18.87 | 2.30 |
| Congo (ROC) | COG | 1245 | 40 | 11790 | 5518092 | 22.56 | 213.66 | 10.56 | 3.21 |
| Rwanda | RWA | 900 | 2 | 137751 | 12952209 | 6.95 | 1063.53 | 0.65 | 0.22 |
| São Tomé and Príncipe | STP | 713 | 13 | 17773 | 219161 | 325.33 | 8109.56 | 4.01 | 1.82 |
| Senegal | SEN | 6698 | 108 | 76343 | 16743930 | 40.00 | 455.94 | 8.77 | 1.61 |
| Seychelles | SYC | 77 | 0 | 704 | 98340 | 78.30 | 715.88 | 10.94 | 0.00 |
| Sierra Leone | SLE | 1427 | 60 | 9973 | 7976985 | 17.89 | 125.02 | 14.31 | 4.20 |
| Somalia | SOM | 2894 | 90 | 11807 | 15893219 | 18.21 | 74.29 | 24.51 | 3.11 |
| South Africa | ZAF | 138134 | 2456 | 1567084 | 59308690 | 232.91 | 2642.25 | 8.81 | 1.78 |
| South Sudan | SSD | 2006 | 37 | 10630 | 11193729 | 17.92 | 94.96 | 18.87 | 1.84 |
| Tanzania ^d | TZA | 509 | 21 | 3880 | 59734213 | 0.85 | 6.50 | 13.12 | 4.13 |
| Togo | TGO | 642 | 14 | 30316 | 8278737 | 7.75 | 366.19 | 2.12 | 2.18 |
| Uganda | UGA | 870 | 0 | 186200 | 45741000 | 1.90 | 407.07 | 0.47 | 0.00 |
| Zambia | ZMB | 1531 | 21 | 53370 | 18383956 | 8.33 | 290.31 | 2.87 | 1.37 |
| Zimbabwe | ZWE | 567 | 6 | 66712 | 14862927 | 3.81 | 448.85 | 0.85 | 1.06 |

475

^a Data from Africa CDC as of June 30, 2020 (<https://africacdc.org/covid-19/>)^b Data from UN Population Division UNPOP (2019 revision) estimates of population by single calendar year (2020)⁴³^c Rates per 100,000 population^d Tanzania has not reported tests, cases, or deaths since April 29, 2020

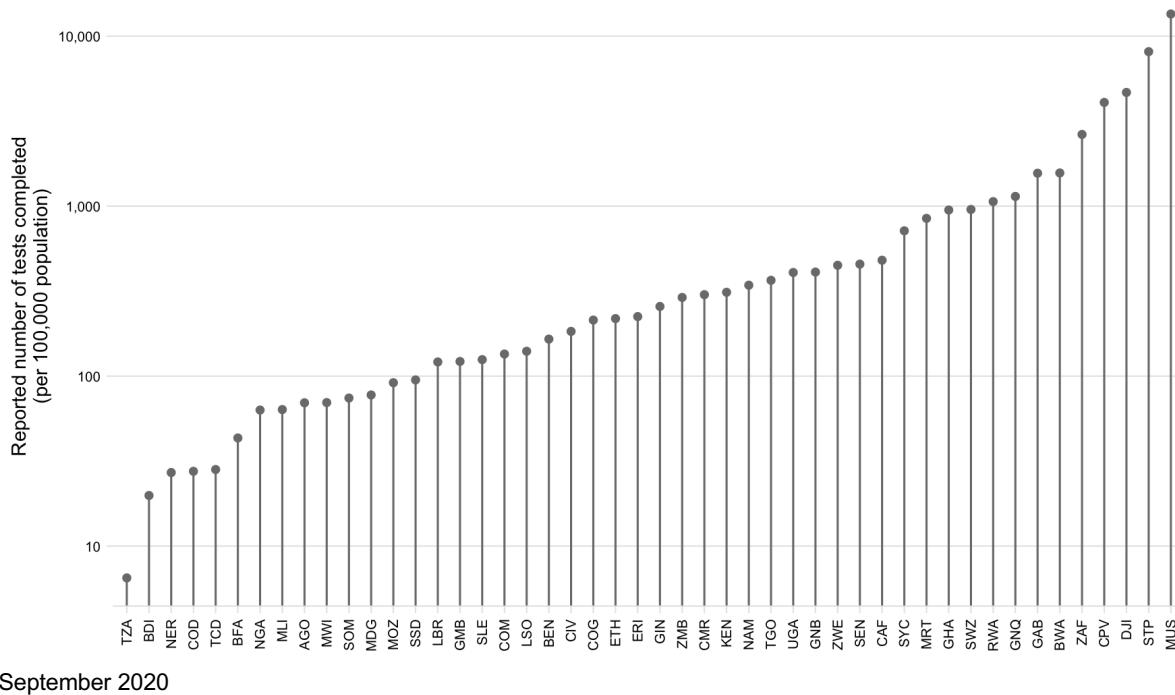
480

481 **Figure S1**

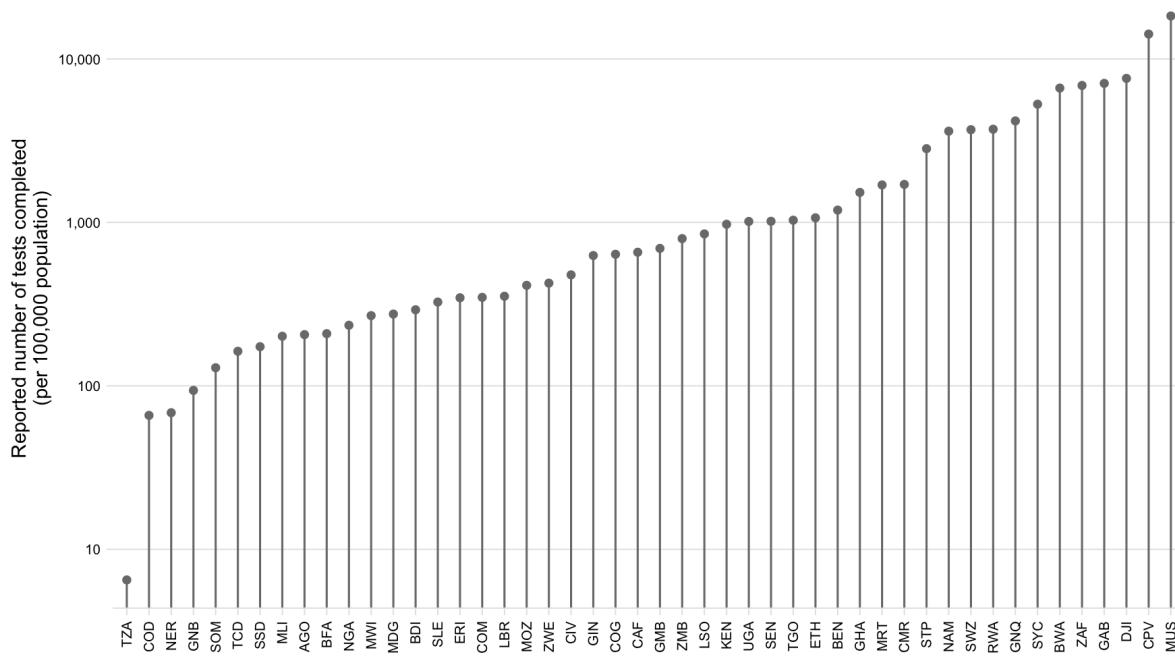
482 **Variation between SSA countries in testing and reporting rates**

483 A: Reported number of tests completed per country

484 June 2020

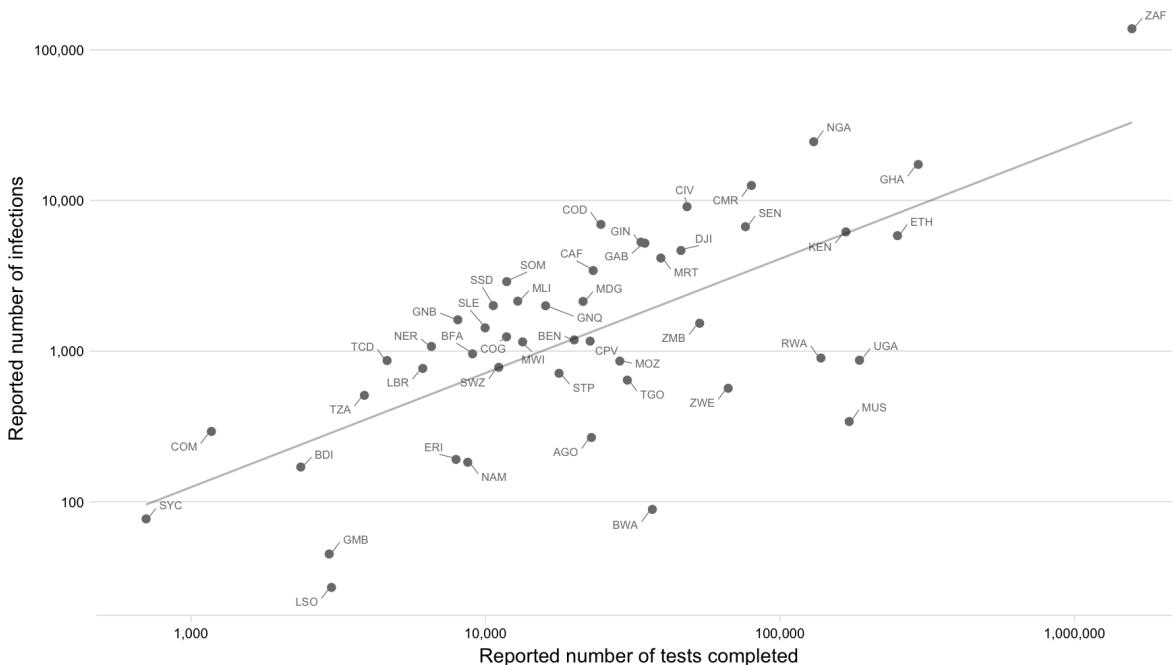


485
486

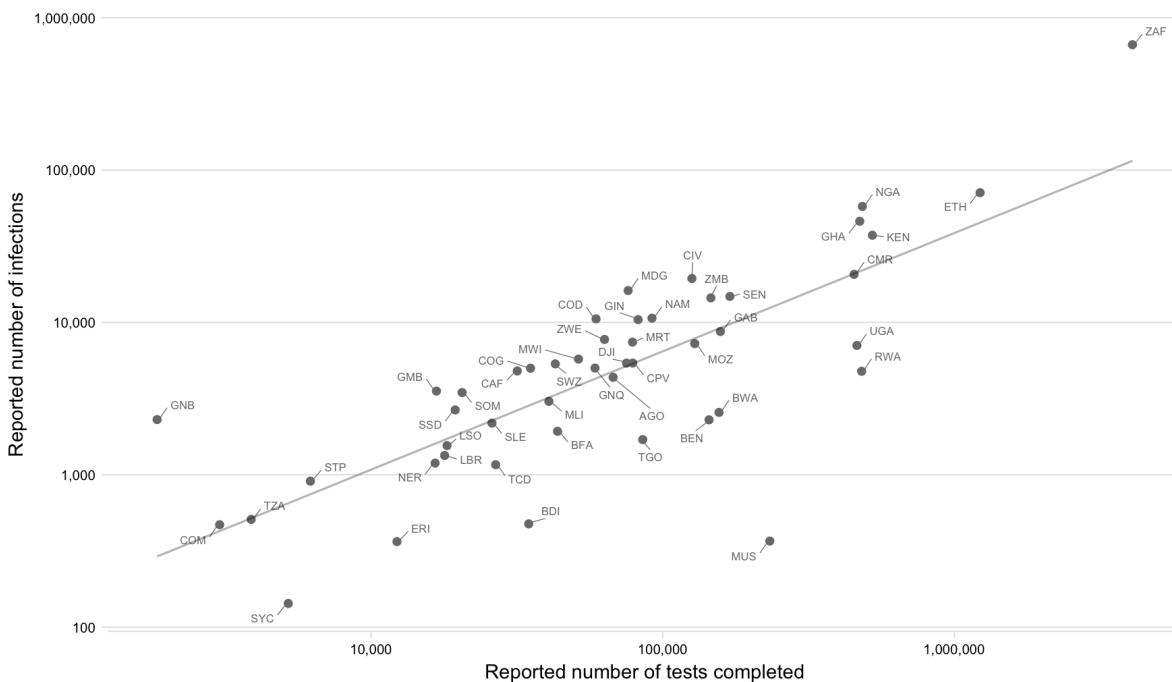


487
488
489

490 (Figure S1 continued)

491 **B:** Number of infections (I) per reported number of tests (T); line shows linear regression.492 June 2020: $I = 8.454 \times 10^{-2} \times T - 8.137 \times 10^2$ ($R^2 = 0.933, p < 0.001$)

493

494 September 2020: $I = 1.523 \times 10^{-1} \times T - 1.090 \times 10^4$ ($R^2 = 0.937, p < 0.001$)

495

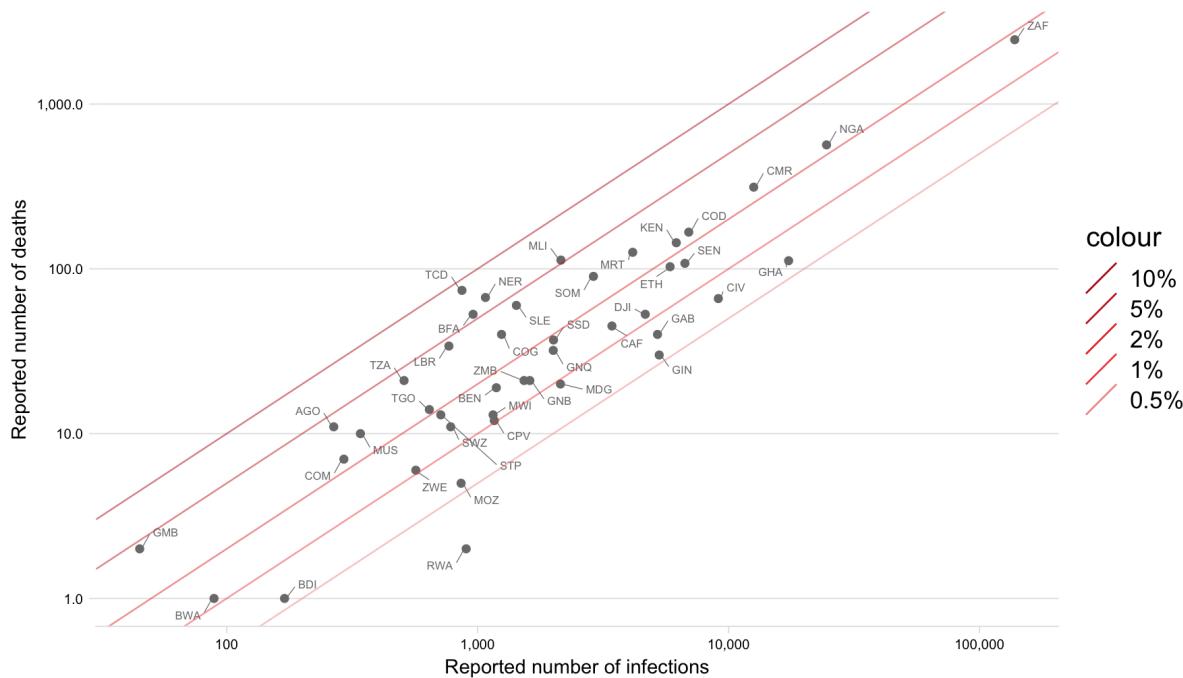
496

497 (Figure S1 continued)

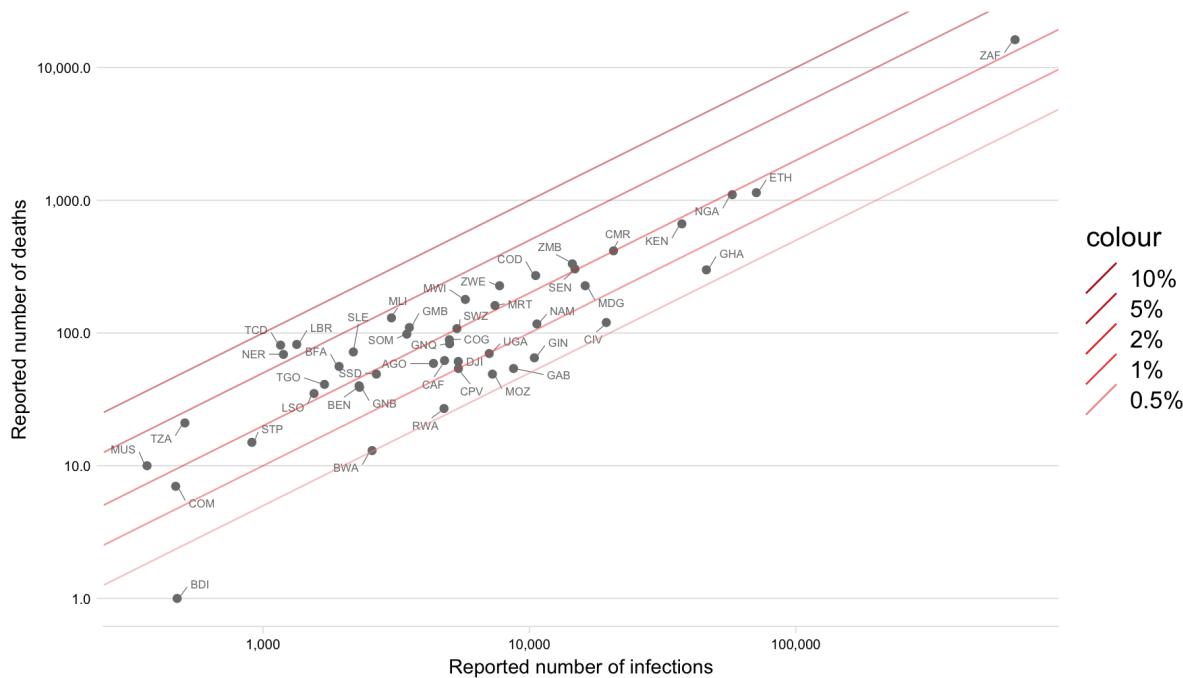
498 C: Reported infections and deaths for sub-Saharan African countries with case fatality ratios (CFRs) shown as

499 diagonal lines

500 June 2020



501 September 2020



503

504

505

506 **A2 | Methods: Synthesizing factors hypothesized to increase or**
507 **decrease SARS-CoV-2 epidemic risk in SSA**

509 *2.1. Variable selection and data sources for variables hypothesized to associate with an*
510 *increased probability of severe clinical outcomes for an infection*

512 To characterize epidemic risk, defined as potential SARS-CoV-2 related morbidity and mortality,
513 we first synthesized factors hypothesized to influence risk in SSA settings (**Table S2**). Early
514 during the pandemic, evidence suggested that age was an important risk factor associated with
515 morbidity and mortality associated with SARS-CoV-2 infection ⁴⁴, a pattern subsequently
516 confirmed across settings ^{2,13,45}. Associations between SARS-CoV-2 mortality and comorbidities
517 including hypertension, diabetes, and cardiovascular disease emerged early ⁴⁴; and have been
518 observed across settings, with further growing evidence for associations with obesity ^{13,46},
519 severe asthma ¹³, and respiratory effects of pollution ⁴⁷. Specific to Africa, vulnerability scores
520 based on these hypothesized associations or combinations of risks factors have been
521 developed (e.g., ^{48,49}).

522
523 Many possible sources of bias complicate interpretation of these associations ⁵⁰, and while they
524 provide a useful baseline, inference is also likely to change as the pandemic advances. To
525 reflect this, our analysis combines a number of high level variables likely to broadly encompass
526 these putative risk factors (e.g., non-communicable disease (NCD) related mortality and health
527 life expectancy) with more specific measures encompassed in evidence to date (e.g.,
528 prevalence of diabetes, obesity, and respiratory illness such as Chronic Obstructive Pulmonary
529 Disease (COPD)). We also include measures relating to infectious diseases, undernourishment,
530 and anemia given their interaction and effects in determining health status in these settings ⁵¹.
531 Although interactions with such infectious diseases have been suggested, evidence is limited to
532 date, barring with HIV, where effects have been suggested to be minor ⁵². We also note that the
533 key concern raised around such infections to date is associated with disruption to routine
534 screening (e.g., for malaria ⁵³), treatment ⁵⁴, or prevention programs ⁵⁵.

535
536 Data on the identified indicators were sourced in May 2020 from the World Health Organization
537 (WHO) Global Health Observatory (GHO) database (<https://www.who.int/data/gho>), World Bank
538 (<https://data.worldbank.org/>), and other sources detailed in **Table S3**. National level
539 demographic data (population size and age structure) was sourced from United Nations World
540 Population Prospects (UNPOP) ⁴³ and data on subnational variation in demography was
541 sourced from WorldPop ²⁹. Household size data was defined by the mean number of individuals
542 in a household with at least one person aged > 50 years, taken from the most recently available
543 demographic health survey (DHS) data ⁵⁶. All country level data for all indicators can be found
544 online at the **SSA-SARS-CoV-2-tool** (<https://labmetcalf.shinyapps.io/covid19-burden-africa/>).

545
546 Comparisons of national level estimates sourced from WHO and other sources are affected by
547 variation within countries and variation in the uncertainty around estimates from different
548 geographical areas. To assess potential differences in data quality between geographic areas
549 we compared the year of most recent data for variables (**Figure S2**). The mean (range varied

550 from 2014.624 to 2014.928 by region) and median year (2016 for all regions) of the most recent
551 data varied little between regions. To account for uncertainty associated in the estimates
552 available for a single variable, we also include multiple variables per category (e.g.,
553 demographic and socio-economic factors, comorbidities, access to care) to avoid reliance on a
554 single metric. This allows exploring variation between countries across a broad suite of
555 variables likely to be indicative of the different dimensions of risk.

556
557 Although including multiple variables that are likely to be correlated (see PCA methods below
558 for further discussion) would bias inference of cumulative risk in a statistical framework, we do
559 not attempt to quantitatively combine risk across variables for a country, nor project risk based
560 on the variables included here. Rather, we characterize the magnitude of variation among
561 countries for these variables (see **Figure 3** in the main text for a subset of the variables; **Figure**
562 **4B** for bivariate risk maps following⁵⁷) and then explore the range of outcomes that would be
563 expected under scenarios where *IFR* increases with age at different rates (see **Figure 4** in the
564 main text).

565
566 2.2. *Variable selection and data sources for variables hypothesized to modulate the rate of viral*
567 *spread*

568
569 In addition to characterizing variation among factors likely to modulate burden, we also
570 synthesize data sources relevant to the rate of viral spread, or pace, for the SARS-CoV-2
571 pandemic in SSA. Factors hypothesized to modulate viral transmission and geographic spread
572 include climatic factors (e.g., specific humidity), access to prevention measures (e.g.,
573 handwashing), and human mobility (e.g., international and domestic travel). **Table S2** outlines
574 the dimensions of risk selected and references the previous studies relevant to the selection of
575 these factors.

576
577 Climate data was sourced from the global, gridded ERA5 dataset⁵⁸ where model data is
578 combined with global observation data (see **Section A7** for details).

579
580 International flight data was obtained from a custom report from OAG Aviation Worldwide (UK)
581 and included the departure location, airport of arrival, date of travel, and number of passenger
582 seats for flights arriving to 113 international airports in SSA (see **Section A5**).

583
584 As an estimate of connectivity within subregions of countries, the population weighted mean
585 travel time to the nearest city with a population greater than 50,000 was determined; details are
586 provided in **Section A6**. To obtain a set of measures that broadly represent connectivity within
587 different countries in the region, friction surfaces from ref²⁸ were used to obtain estimates of the
588 connectivity between different administrative level 2 units within each country. Details of this,
589 alongside the metapopulation model framework used to simulate viral spread with variation in
590 connectivity are in **Section A6**.

591
592 **Figure 3** in the main text shows variation among SSA countries for four of the variables; **Figure**
593 **S3** shows variation for all variables. **Figure 4** in the main text shows variation for a subset of the

594 comorbidity and access to care indicators as a heatmap; **Figure S4** shows variation for all the
595 variables (both also available online at the [SSA-SARS-CoV-2-tool](#)
596 (<https://labmetcalf.shinyapps.io/covid19-burden-africa/>)).

597

Table S2

598

**Hypothesized dimensions of risk and expected direction of effect on SARS-CoV-2
599 transmission or burden in sub-Saharan Africa (SSA) relative to higher latitude countries**

| Dimension of risk | Factors hypothesized to decrease transmission or burden in sub-Saharan Africa relative to other geographic areas | Factors hypothesized to increase transmission or burden in sub-Saharan Africa relative to other geographic areas |
|---|---|---|
| (A) Demographic and socio-economic characteristics in SSA | Younger populations, and thus a smaller proportion of individuals in the older age groups that experience the highest mortality ²⁻⁴ | A larger proportion of urban populations living in dense settings, which may result in higher transmission ⁵⁹ ; higher contact with older individuals as a result of multi-generation households ¹⁸ |
| (B) Comorbidities in SSA | Lower rates of some comorbidities that have been associated with risk of worse outcomes, e.g., obesity ^{13,46} | Higher rates of NCDs such as hypertension or COPD ⁴⁴ , which are associated with worse outcomes; and a potential role for as yet undescribed interactions e.g., with anemia, or high prevalence infectious diseases |
| (C) Climate in SSA | Warmer, wetter climates on average driving reduced transmission ^{1,60} | |
| (D) Capacity to deploy prevention measures in SSA | Experience with previous outbreak response which may yield more rapid and nimble approaches to reducing transmission ^{25,26} | Lower access to handwashing ^{61,62} and other prevention options such as self-isolation ⁶³ , increasing transmission Subregions of countries with reduced governance infrastructure ⁶⁴ |
| (E) Access to healthcare in SSA | | Larger variation in access to and coverage of health systems ⁶⁵ including fewer medical staff and facilities such as hospital beds ¹⁸ increasing burden Increased vulnerability to disruption of routine health services (e.g., ³⁶) Limited testing capacity ⁶⁶ reducing the capacity to identify and interrupt chains of transmission |
| (F) Human mobility and travel in SSA | Fewer viral importations due to reduced frequency of international travel ^{34,35} Decreased rate of internal spread due to less connectivity within countries ⁶⁷ | |

600

Table S3**Variables and data sources for indicators of SARS-CoV-2 epidemic risk in sub-Saharan Africa**

| ID | Variable | Source | Hypothesized association(s) with SARS-CoV-2 outcomes |
|--|---|------------------------|---|
| <i>(A) Demographic and socio-economic characteristics</i> | | | |
| A1 | Human population size; Proportion of population over age 50 (%) from the UN Population Division UNPOP (2019 revision) estimates of population by single calendar year (2020), age, and country | UN ⁴³ | Morbidity and mortality observed to increase with age (e.g., ²⁻⁴) |
| A2 | Subnational spatial variation in the distribution of the human population and age structure | WorldPop | |
| A3 | Household size: Mean household size for households with an individual over age 50 | DHS | Proxy for social contact rate for the elderly population at higher risk for SARS-CoV-2 morbidity and mortality ¹⁸ |
| A4 | Proportion of households with an individual over age 50 | DHS | |
| A5 | Health life expectancy (HALE) at age 60 (years) | WHO | Proxy for baseline health status of elderly population |
| A6 | Proportion of population below the poverty line (%) | World Bank | More severe clinical outcomes associated with poverty; A proxy for access to advanced care ^{68,69} |
| A7 | Proportion of the urban population living in crowded, low quality housing (defined as households lacking one or more of the following conditions: access to improved water, access to improved sanitation, sufficient living area, and durability of housing) (%) | World Bank | Indicator of capacity for prevention (e.g., through handwashing); Transmission observed to increase with crowding ⁵⁹ |
| A8 | Gross domestic product (GDP) per capita | World Bank | Used in PCA analysis (see below) as an indicator of socio-economic status at the national level |
| A9 | GINI index, a measure of inequality in the distribution of income | World Bank | |
| <i>(B) Comorbidities: General and nutrition related non-communicable diseases (NCDs)</i> | | | |
| B1 | NCDs overall mortality per 100 000 popn, age-standardized | WHO | Indicator of NCD burden in population; Comorbidities increase probability of severe clinical outcomes |
| B2 | Cardiovascular disease related mortality per age group (annual deaths attributable per 100,000 population) | GDB 2017 ⁷⁰ | |
| B3 | Diabetes prevalence among ages 20-79 (%) | World Bank | Increases probability of severe clinical outcomes |
| B4 | Diabetes related mortality per age group (annual deaths attributable per 100,000 population) | GDB 2017 ⁷⁰ | |

608
609

(Table S3 continued)

| ID | Variable | Source | Hypothesized association(s) with SARS-CoV-2 outcomes |
|--|---|------------------------|--|
| <i>(B) Comorbidities: General and nutrition related non-communicable diseases (NCDs)</i> | | | |
| B5 | Raised glucose prevalence, age-standardized (%) | WHO | Indicator of metabolic disease risk; Metabolic disease increases probability of severe clinical outcomes |
| B6 | Raised blood pressure prevalence, age-standardized (%) | WHO | |
| B7 | Raised cholesterol prevalence, age-standardized (%) | WHO | |
| B8 | Overweight prevalence among adults, age-standardized (%) | WHO | |
| B9 | Anemia prevalence among non-pregnant women (%) | WHO | Indicator of poor nutritional status; Poor nutritional status may increase probability of severe clinical outcomes |
| B10 | Undernourishment prevalence (%) | WHO | |
| <i>(B) Comorbidities: NCDs related to respiratory system and pollution</i> | | | |
| B11 | Annual mean PM2.5 exposure in urban areas (ug/m ³) | WHO | Exposure to air pollution increases mortality ⁴⁷ |
| B12 | Lung, tracheal, and esophageal cancer mortality per 100 000 popn, age-standardized | WHO | Indicator of prevalence and management of chronic disease and inflammation affecting the respiratory tract |
| B13 | Chronic respiratory diseases (excluding asthma) related mortality per age group (annual deaths attributable per 100,000 population) | GDB 2017 ⁷⁰ | |
| B14 | COPD mortality per 100 000 popn, age-standardized | WHO | |
| <i>(B) Comorbidities: Infectious diseases</i> | | | |
| B15 | Respiratory infections mortality per 100 000 popn, age-standardized | WHO | Indicator of prevalence and management of infectious disease affecting the respiratory tract |
| B16 | TB incidence per 100 000 popn | World Bank | Indicator of susceptibility to respiratory infections and immune suppression |
| B17 | HIV prevalence among ages 15-49 (%) | World Bank | Indicator of immunosuppressed population |
| <i>(C) Climate</i> | | | |
| C1 | Seasonal change in specific humidity (in selected urban centers) | ERA5 ⁵⁸ | Transmission rate of coronaviruses may decline with humidity |
| <i>(D) Capacity to deploy prevention measures</i> | | | |
| D1 | Proportion of urban popn with basic handwashing facilities with water and soap at home (%) | WHO | Handwashing observed to reduce infection rates for respiratory pathogens |

| | | | |
|----|---|-------------------|--|
| D2 | Proportion of the population with access to a handwashing station with soap and water in 2019 | Ref ⁷¹ | |
|----|---|-------------------|--|

610
611

(Table S3 continued)

| ID | Variable | Source | Hypothesized association(s) with SARS-CoV-2 outcomes |
|---|--|-------------------|--|
| <i>(D) Capacity to deploy prevention measures</i> | | | |
| D3 | Proportion of 1 year olds receiving full immunization coverage (%) | WHO | Proxy for coverage of routine health services |
| D4 | Reported number of completed tests reported for SARS-CoV-2 infection as of June 30, 2020 | Africa CDC | Indicator of surveillance capacity |
| <i>(E) Access to healthcare in SSA</i> | | | |
| E1 | Proportion of children with pneumonia symptoms taken to a health facility (%) | WHO | Proxy for access to medical care and care seeking |
| E2 | Subnational spatial variation in the probability of seeking treatment for fever at public facilities | Ref ⁷² | |
| E3 | Proportion of births attended by skilled staff (%) | World Bank | |
| E4 | Nurses and midwives per 100 000 popn | World Bank | Indicators of treatment capacity |
| E5 | Physicians per 100 000 popn | World Bank | |
| E6 | Hospitals per 100 000 popn | World Bank | |
| E7 | Hospital beds 100 000 popn | World Bank | |
| E8 | Health expenditure per capita in (USD) | WHO | Proxy for health system resources; A significant predictor of intensive care unit (ICU) capacity ⁷³ |
| E9 | Proportion of health expenditures that are out-of-pocket (%) | WHO | |
| <i>(F) Human mobility and travel: International</i> | | | |
| F1 | Estimated number of international passengers arriving at SSA airports from January-April 2020 | OAG | Indicator of the timing and number of introductions of SARS-CoV-2 |
| F2 | Estimated number of international passengers arriving at SSA airports from January-April 2020 by SARS-CoV-2 status at departure location | OAG | |

612
613

614

(Table S3 continued)

| ID | Variable | Source | Hypothesized association(s) with SARS-CoV-2 outcomes |
|--|---|-------------------|--|
| <i>(F) Human mobility and travel: Domestic</i> | | | |
| F3 | National population-weighted mean travel time to the nearest city (national mean of indicator F4) | Ref ⁷⁴ | |
| F4 | Population-weighted mean travel time to the nearest city (population > 50,000) for administrative level 2 units | Ref ⁷⁴ | Indicator of connectivity within countries; A proxy for the rate of human mobility |
| F5 | Relative costs of travel between centroids of administrative level 2 derived from friction surfaces obtained by integrating data on travel infrastructure (Open Street Map, land cover types, etc). | Ref ²⁸ | |

615

616

617

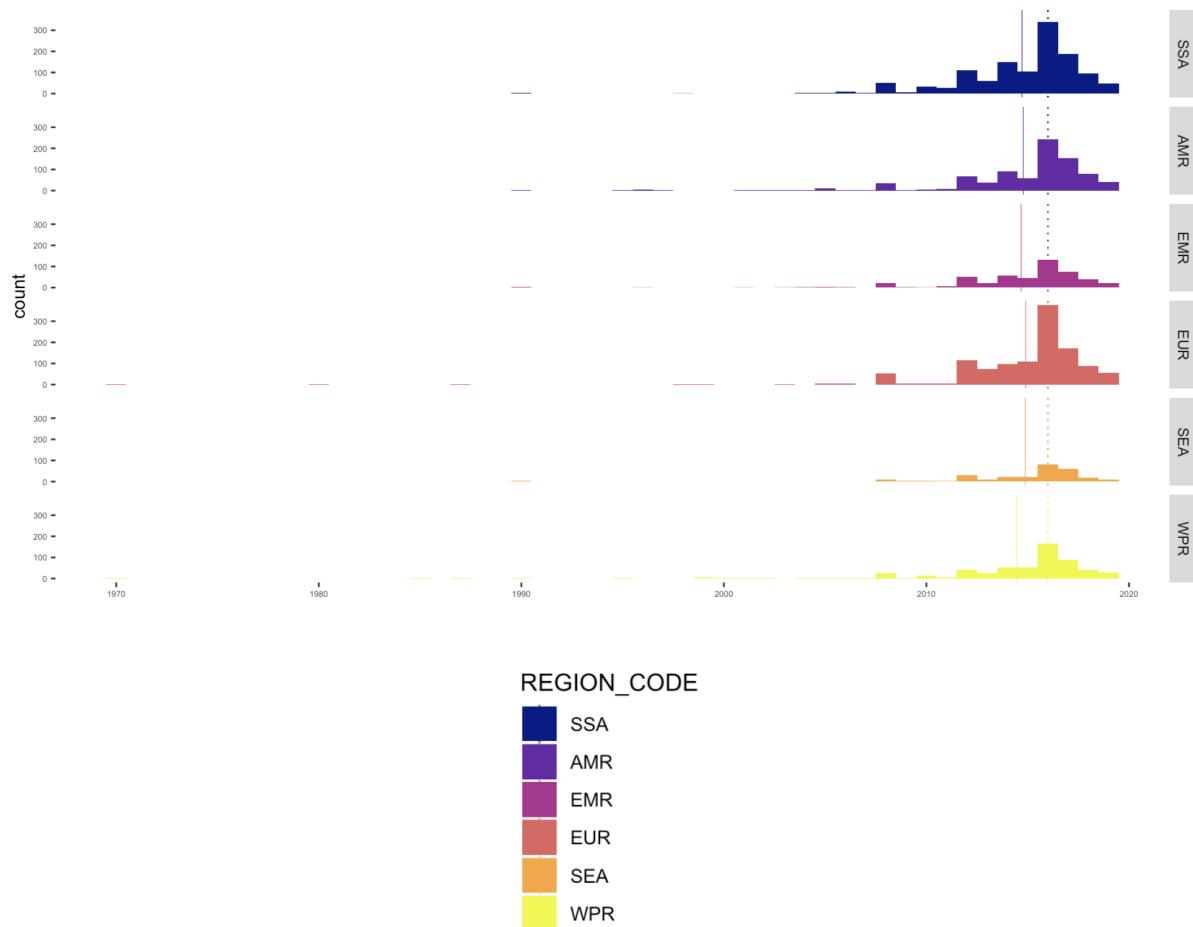
618

619 **Figure S2**

620 **Year of most recent data available for variables compared between global regions**

621 Dotted vertical line shows regional median; solid vertical line shows regional mean. Note that most data comes from
 622 2015-2019 (median = 2016, mean = 2014.62-2014.93).

623



626

627

628

629

630

631 **Figure S3**

632 **Variation among sub-Saharan African countries in determinants of SARS-CoV-2 risk by**
633 **variable**

634 A subset of variables is shown in Figure 3A-D in the main text, the remaining variables are
635 shown in supplementary file “Figure S3 compiled.pdf” and available online: **SSA-SARS-CoV-2-**
636 **tool** (<https://labmetcalf.shinyapps.io/covid19-burden-africa/>)

637

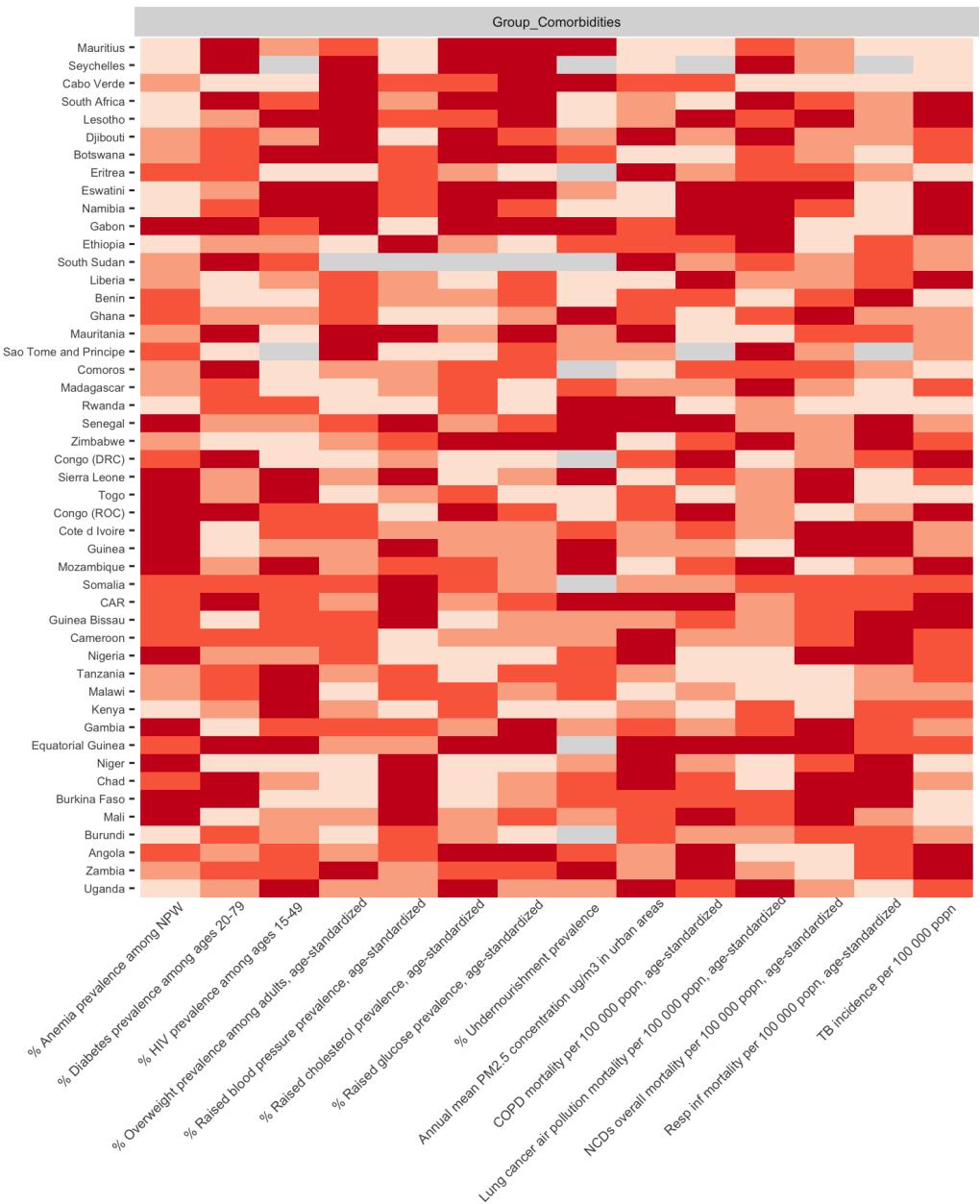
638

639 **Figure S4**

640 **Variation among sub-Saharan African countries in determinants of SARS-CoV-2 mortality**
 641 **risk by category**

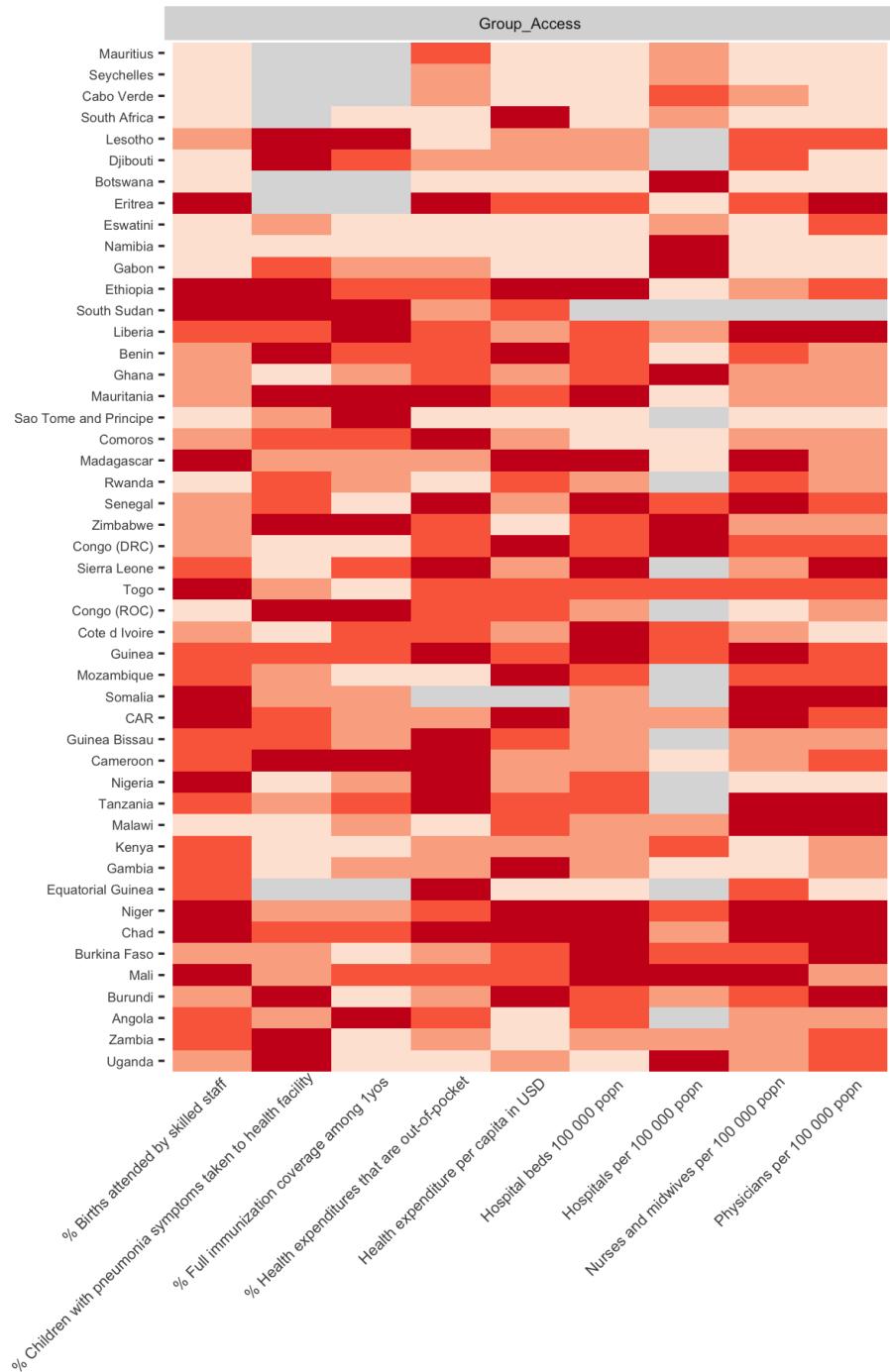
642 A subset of variables is shown in Figure 4D-E in the main text, the remaining variables are shown and available
 643 online: [SSA-SARS-CoV-2-tool](https://labmetcalf.shinyapps.io/covid19-burden-africa/) (<https://labmetcalf.shinyapps.io/covid19-burden-africa/>)

644 **A:** Select national level indicators; estimates of increased comorbidity burden (e.g., higher prevalence of raised blood
 645 pressure) shown with darker red for higher risk quartiles Countries missing data for an indicator (NA) are shown in
 646 gray. For comparison between countries, estimates are age-standardized where applicable (see **Table S3** for details)



647
 648
 649

650 **B:** Select national level indicators; estimates of reduced access to care (e.g., fewer hospitals) shown with darker red
 651 for higher risk quartiles Countries missing data for an indicator (NA) are shown in gray. For comparison between
 652 countries, estimates are age-standardized where applicable (see **Table S3** for details)



653

654 A3 | Principal component analysis (PCA) of variables considered

655

656 3.1 Selection of data and variables

657

658 The 29 national level variables from **Table S3** were selected for principal component analysis
659 (PCA). We conducted further PCA on the subset of eight indicators related to access to
660 healthcare (Category E) and the 14 national indicators variables related to comorbidities
661 (Category B).

662

663 We excluded disaggregated sub-national spatial variation data (variables A2, C1, E2, and
664 Category F), disaggregated or redundant variables derived from already included variables
665 (variables A4 and D2), and disaggregated age-specific disease data from IHME global burden
666 of disease study (variables B2, B4, and B13) from PCA analysis. COVID-19 tests per 100,000
667 population (variable D4, **Table S1**), per capita gross domestic product (GDP) (Variable A8), and
668 the GINI index of wealth inequality (Variable A9) were used to visualize patterns among sub-
669 Saharan Africa countries.

670

671 In some cases, data were missing for a country for an indicator; in these cases, missing data
672 were replaced with a zero value. This is a conservative approach as zero values (i.e., outside
673 the range of typical values seen in the data) inflate the total variance in the data set and thus, if
674 anything, deflate the percent of the variance explained by PCA. Therefore, this approach avoids
675 mistakenly attributing predictive value to principal components due to incomplete data. See
676 **Table S3** for data sources for each variable.

677

678 3.2 Principal Component Analysis

679

680 The PCA was conducted on each of the three subsets described above, using the scikitlearn
681 library⁷⁵. In order to avoid biasing the PCA due to large differences in magnitude and scale,
682 each feature was centered around the mean, and scaled to unit variance prior to the analysis.
683 Briefly, PCA applies a linear transformation to a set of n features to output a set of n orthogonal
684 principal components which are uncorrelated and each explain a percentage of the total
685 variance in the dataset⁷⁶. A link to the code for this analysis is available online at the [SSA-SARS-CoV-2-tool](#) (<https://labmetcalf.shinyapps.io/covid19-burden-africa/>).

687

688 The principal components were then analyzed for the percentage of variance explained, and
689 compared to: (i) the number of COVID-19 tests per 100,000 population as of the end of June,
690 2020 (**Table S1**), (ii) the per capita GDP, and (iii) the GINI index of wealth inequality. For the
691 GINI index, estimates from 2008-2018 were available for 45 of the 48 countries (no GINI index
692 data were available for Eritrea, Equatorial Guinea, and Somalia) (see **Data File 1** for the year
693 for each country for each metric).

694

695

696 3.3 PCA Results

697

698 The first two principal components from the analysis of 29 variables explain 32.6%, and 13.1%
699 the total variance, respectively, in the dataset. Countries with higher numbers of completed
700 SARS-CoV-2 tests reported tended to associate with an increase in principal component 1
701 (Pearson correlation coefficient, $r = 0.67$, $p = 1.1\text{e-}7$, **Figure S5A**). Similarly, high GDP
702 countries seem to associate with an increase in principal component 1 (Pearson correlation
703 coefficient, $r = 0.80$, $p = 6.02\text{e-}12$), **Figure S5B**). In contrast, countries with greater wealth
704 inequality (as measured by the GINI index) are associated with a decrease in principal
705 component 2 (Pearson correlation coefficient, $r = -0.42$, $p = .0042$, **Figure S5C**). Despite these
706 correlations, a relatively low percentage of variance is explained by each principal component:
707 for the 29 variables, 13 of the 29 principal components are required to explain 90% of the
708 variance (**Figure S5D**). When only the access to care subset of variables is considered, the first
709 two principal components explain 50.7% and 19.1% of the variance, respectively, and five of
710 eight principal components are required to explain 90% of the variance. When only the
711 comorbidities subset is considered, the first two principal components explain 27.9% and 17.8%
712 of the variance, respectively, and nine of 14 principal components are required to explain 90%
713 of the variance (**Figure S5D**).

714

715 3.4 PCA Discussion

716

717 These data suggest that inter-country variation in this dataset is not easily explained by a small
718 number of variables. Moreover, though correlations exist between principal components and
719 high-level explanatory variables (testing capacity, wealth), their magnitude is modest. These
720 results highlight that dimensionality reduction is unlikely to be an effective analysis strategy for
721 the variables considered in this study. Despite this overall finding, the PCA on the access to
722 care subset of variables highlights that the variance in these variables is more easily explained
723 by a small number of principal components, and hence may be more amenable to
724 dimensionality reduction. This finding is unsurprising as, for example, the number of hospital
725 beds per 100,000 population is likely to be directly related to the number of hospitals per
726 100,000 population (indeed $r = 0.60$, $p = 5.7\text{e-}6$ for SSA). In contrast, for comorbidities, the
727 relationship between different variables is less clear. Given the low percentages of variation
728 captured by each principal component, and the high variability between different types of
729 variables, these results motivate a holistic approach to using these data for assessing relative
730 SARS-CoV-2 risk across SSA.

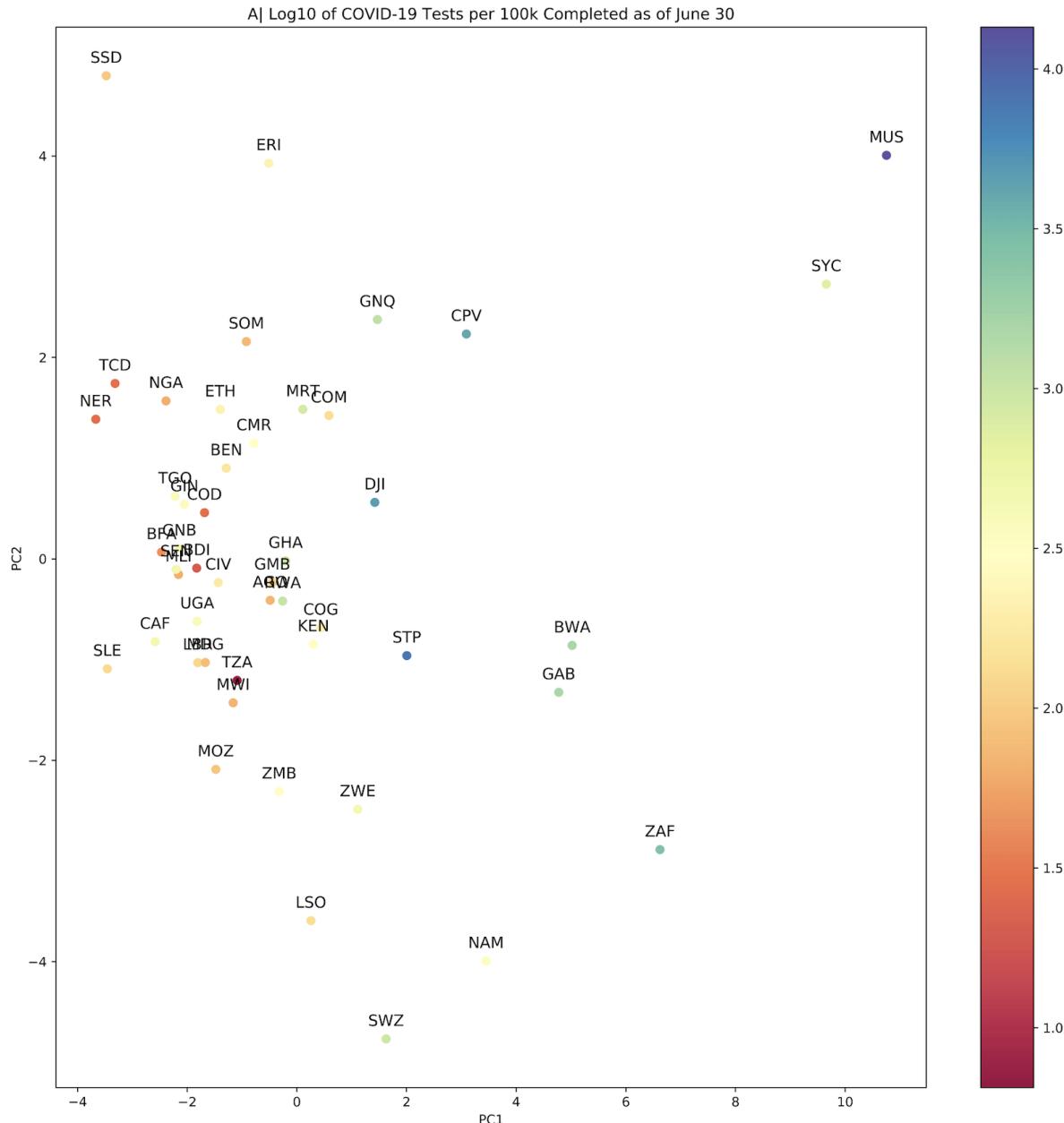
731

732

733 Figure S5

734 Principal Component Analysis of all variables and category specific subsets of variables

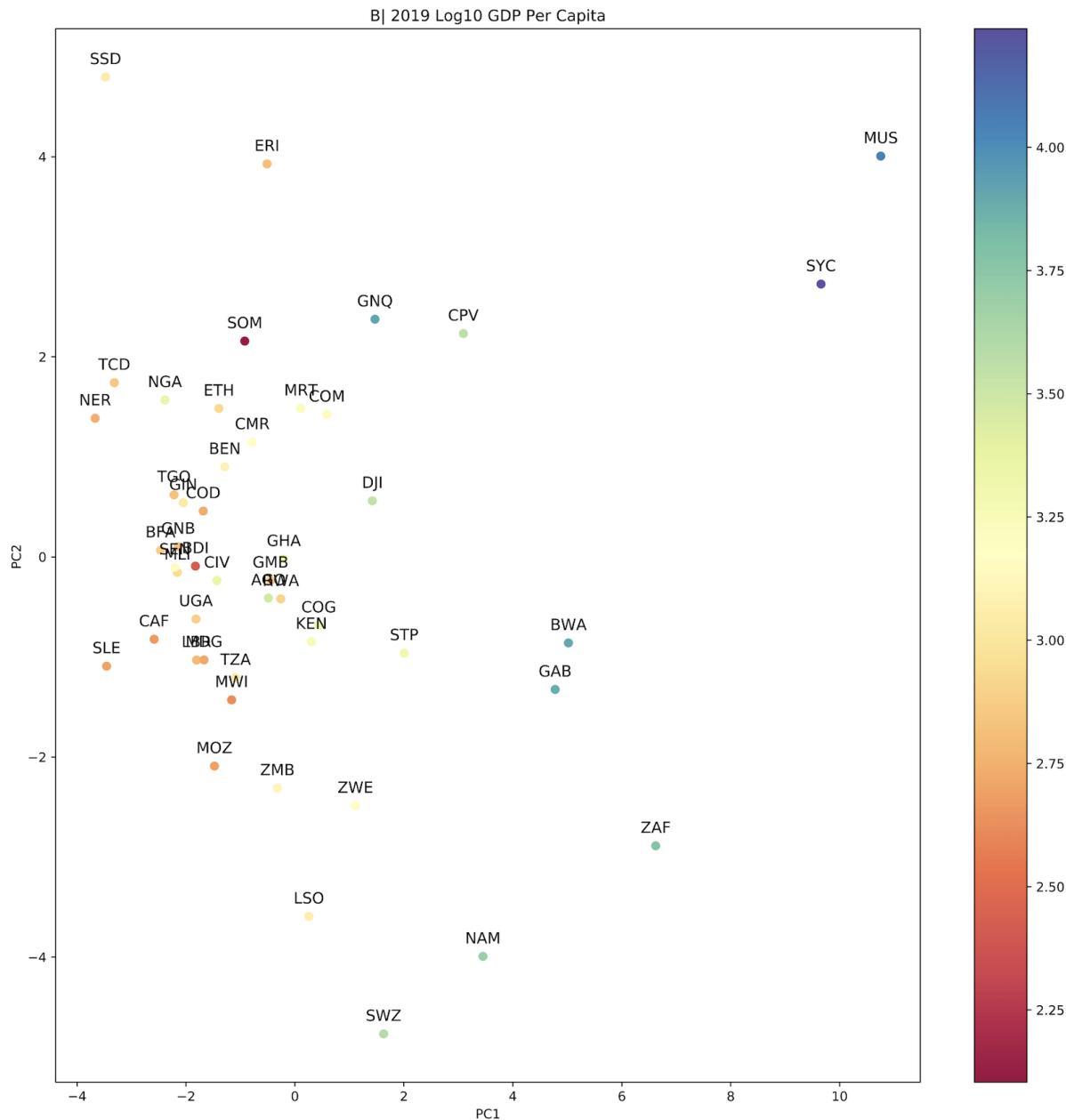
A: Principal Component 1 and 2, countries colored by Log10 scaled tests per 100,000 population (as of June 30, 2020)



737

738

739 (Figure S5 continued)

Figure S5**Principal Component Analysis of all variables and category specific subsets of variables**742 **B:** Principal Component 1 and 2, countries colored by Log10 scaled GDP per capita

743

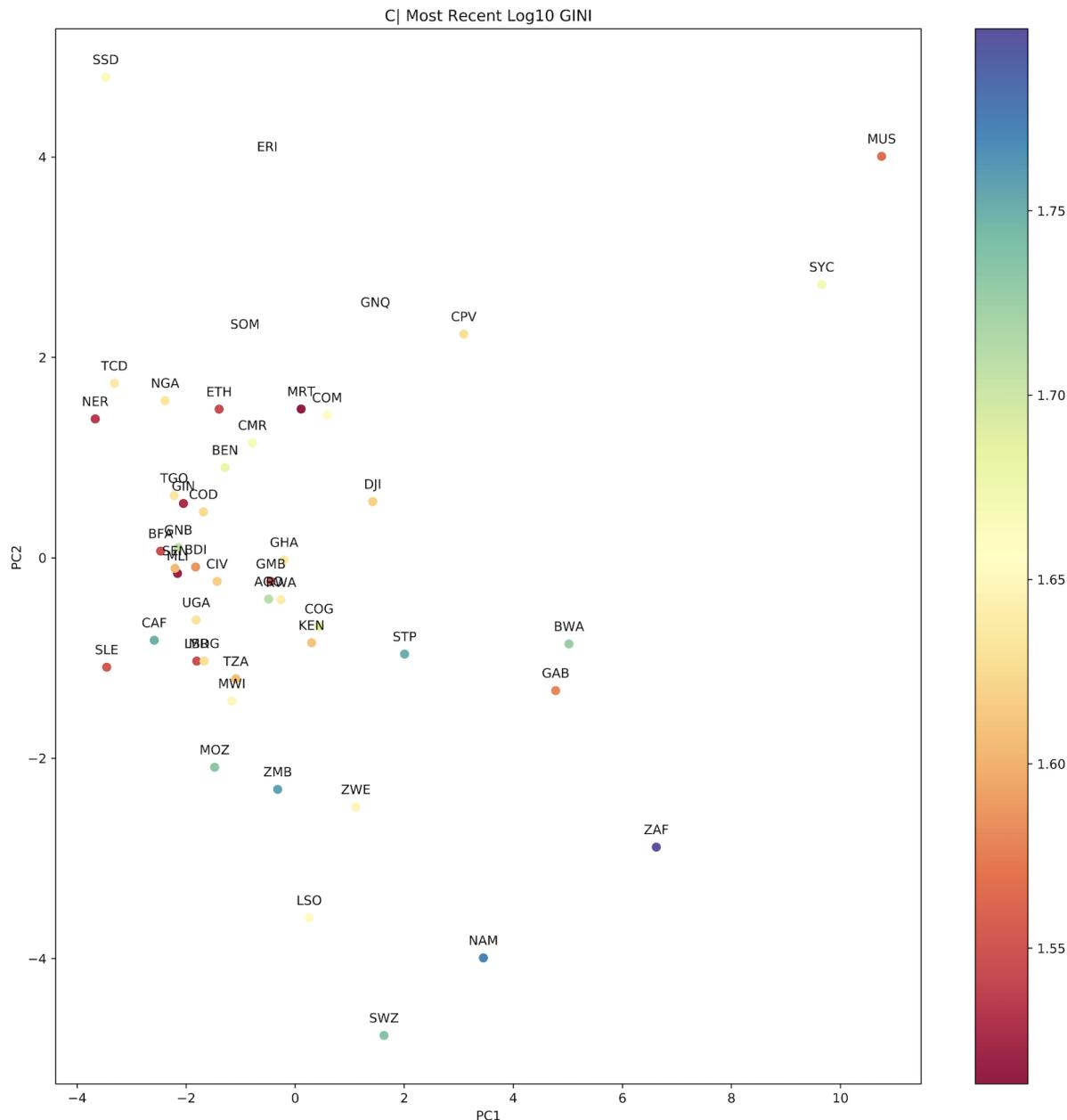
744

745

746 (Figure S5 continued)

Figure S5**Principal Component Analysis of all variables and category specific subsets of variables**

749 C: Principal Component 1 and 2, countries colored by the GINI index (a measure of wealth disparity)



750

751

752

753 (Figure S5 continued)

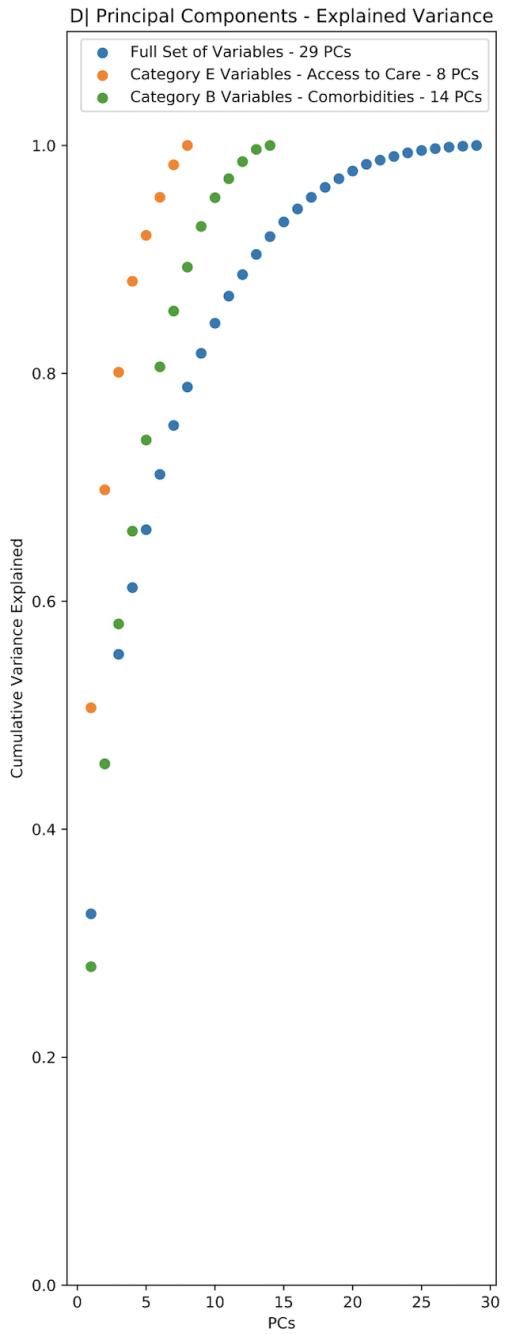
754 **Figure S5**

755 **Principal Component Analysis of all variables and category specific subsets of variables**

756 D: Scree plot showing the cumulative proportion of variance explained by principal component for analysis done

757 using all variables (blue, 29 variables), comorbidity indicators (green, 14 variables, Section B in **Table S3**), and

758 access to care indicators (orange, 8 variables, Section E in **Table S3**)



759

760

761 **A4 | Evaluating the burden emerging from the severity of infection
762 outcome**

764 *4.1 Data sourcing: Empirical estimates of IFR*

766 Estimates of the infection fatality ratio (*IFR*) that account for asymptomatic cases,
 767 underreporting, and delays in reporting are few, however, it is evident that *IFR* increases
 768 substantially with age⁷⁷. We use age-stratified estimates of *IFR* from three studies (two
 769 published^{2,4}, one preprint³) that accounted for these factors in their estimation (**Table S4**).
 770

771 **Table S4: Sources of age-stratified IFR estimates**

| Study | Population | Methods |
|----------------------------------|---|---|
| Salje et al. 2020 ² | Deaths and hospitalizations due to COVID-19 in French public and private hospitals across the country between 13 March - 11 May | Combined data from France with data from Diamond Princess Cruise ship to estimate age-stratified IFR, case severity, and hospitalization probabilities accounting for asymptomatic cases and underascertainment. |
| Verity et al. 2020 ⁴ | Deaths due to COVID-19 in Hubei province, China | Combined data from Hubei with data from PCR testing of repatriated citizens under quarantine to estimate age-stratified IFR accounting for asymptomatic cases and underascertainment. |
| Rinaldi et al. 2020 ³ | Deaths due to COVID-19 reported in Lombardia, Italy | Analyzed deaths in the Lombardia region, one of the hardest hit regions in Italy, and used seroprevalence surveys of the region to estimate that 30% of the population was infected to estimate age-stratified IFR. |

772
 773 To apply these estimates to other age-stratified data with different bin ranges and generate
 774 continuous predictions of *IFR* with age, we fit the relationship between the midpoint of the age
 775 bracket and the *IFR* estimate using a generalized additive model (GAM) using the ‘mgcv’
 776 package⁷⁸ in R version 4.0.2⁷⁹. We use a beta distribution as the link function for IFR estimates
 777 (data distributed on [0, 1]). For the upper age bracket (80+ years), we take the upper range to
 778 be 100 years and the midpoint to be 90.
 779

780 We assume a given level of cumulative infection (here 20% in each age class, i.e., a constant
 781 rate of infection among age classes) and then apply *IFRs* by age to the population structure of
 782 each country to generate estimates of burden. Age structure estimates were taken from the
 783 UNPOP (see **Table S3**) country level estimates of population in 1 year age groups (0 - 100
 784 years of age) to generate estimates of burden.
 785
 786
 787

788 4.2 Data sourcing: Comorbidities over age from IHME
789

790 Applying these *IFR* estimates to the demographic structure of SSA countries provides a
791 baseline expectation for mortality, but depends on the assumption that mortality patterns in sub-
792 Saharan Africa will be similar to those from where the *IFR* estimates were sourced (France,
793 China, and Italy). Comorbidities have been shown to be an important determinant of the severity
794 of infection outcomes (i.e., *IFR*); to assess the relative risk of comorbidities across age in SSA,
795 estimates of comorbidity severity by age (in terms of annual deaths attributable) were obtained
796 from the Institute for Health Metrics and Evaluation (IHME) Global Burden of Disease (GBD)
797 study in 2017⁸⁰. Data were accessed through the GBD results tool for cardiovascular disease,
798 chronic respiratory disease (not including asthma), and diabetes, reflecting three categories of
799 comorbidity with demonstrated associations with risk (**Table S2**). We make the assumption that
800 higher mortality rates due to these NCDs, especially among younger age groups, is indicative of
801 increased severity and lesser access to sufficient care for these diseases - suggesting an
802 elevated risk for their interaction with SARS-CoV-2 as comorbidities. While there are significant
803 uncertainties in these data, they provide the best estimates of age specific risks and have been
804 used previously to estimate populations at risk²².

805
806 The comorbidity by age curves for SSA countries were compared to those for the three
807 countries from which SARS-CoV-2 *IFR* by age estimates were sourced. Attributable mortality
808 due to all three NCD categories is higher at age 50 in all 48 SSA countries when compared to
809 estimates from France and Italy and for 42 of 48 SSA countries when compared to China
810 (**Figure S6**).

811 Given the potential for populations in SSA to experience a differing burden of SARS-CoV-2 due
812 to their increased severity of comorbidities in younger age groups, we explore the effects of
813 shifting *IFRs* estimated by the GAM of *IFR* estimates from France, Italy, and China younger by
814 2, 5, and 10 years (**Figure 3** in main text).

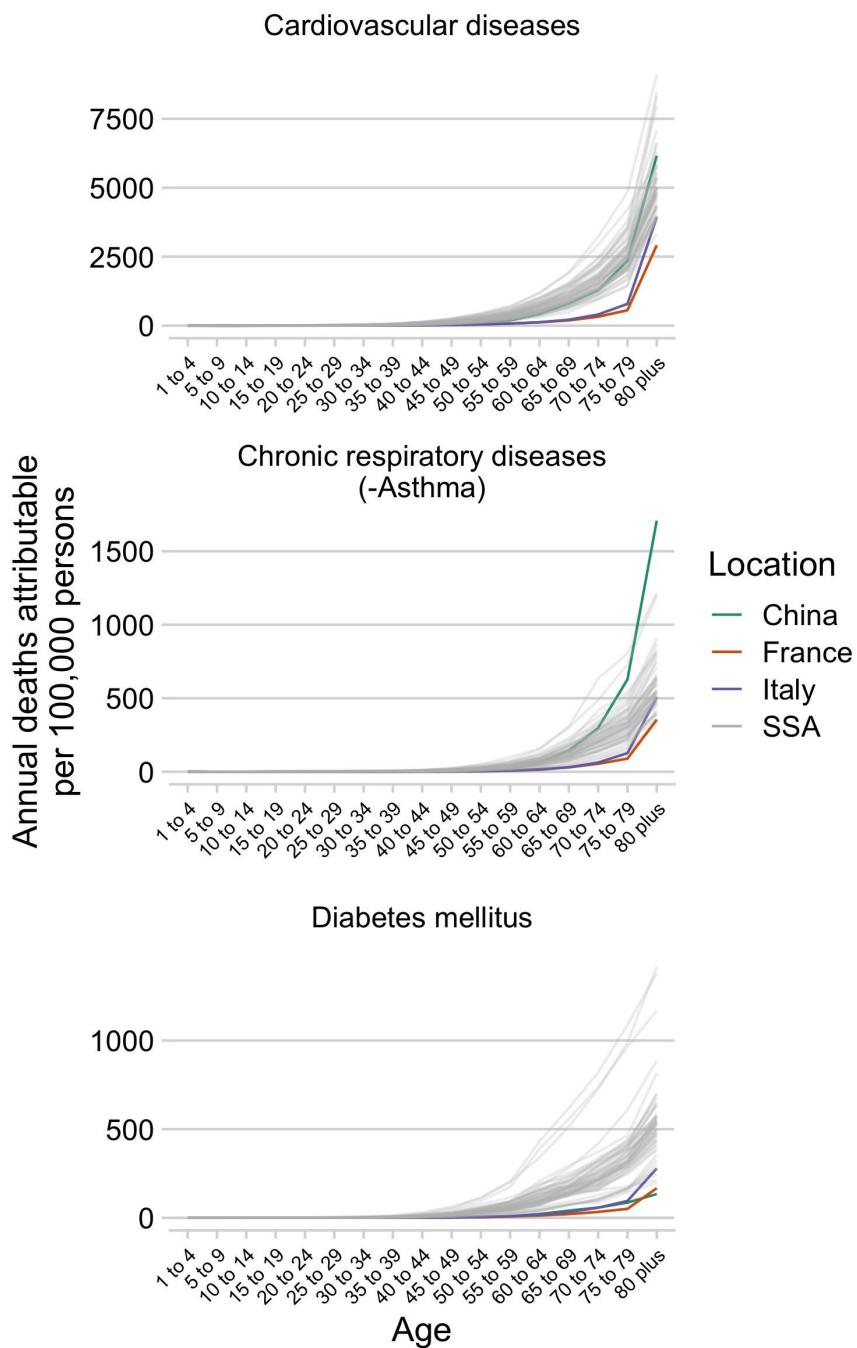
815
816

817

818 **Figure S6**

819 **Comorbidity burden by age in sub-Saharan Africa**

820 Estimated mortality per age group for sub-Saharan African countries (gray lines) compared to China, France, and
 821 Italy (the countries from which estimates of SARS-CoV-2 infection fatality ratios (*IFRs*) by age are available) for three
 822 NCD categories (cardiovascular diseases, chronic respiratory diseases excluding asthma, and diabetes).



A5 | International Air Travel to SSA

The number of passenger seats on flights arriving to international airports were grouped by country and month for January 2020 to April 2020 (**Table S5**) - the months when the introduction of SARS-CoV-2 to SSA countries was likely to have first occurred. The first confirmed case reported from a SSA country, per the Johns Hopkins Coronavirus Research Center was in Nigeria on February 28, 2020. By March 31, 2020, 43 of 48 SSA countries had reported SARS-CoV-2 infections and international travel was largely restricted by April. Lesotho was the last SSA country to report a confirmed SARS-CoV-2 infection (on May 13, 2020); however, given difficulties in surveillance, the first reported detections were likely delayed relative to the first importations of the virus.

The probability of importation of the virus is defined by the number of travelers from each source location each date and the probability that a traveler from that source location on that date was infectious. Due to limitations in surveillance, especially early in the SARS-CoV-2 pandemic, empirical data on infection rates among travelers is largely lacking. To account for differences in the status of the SARS-CoV-2 pandemic across source locations, and thus differences in the importation risk for travelers from those locations, we coarsely stratified travelers arriving each day into four categories based on the status of their source countries:

- i. Travelers from countries with zero reported cases (i.e., although undetected transmission was possibly occurring, SARS-CoV had not yet been confirmed in the source country by that date)
- ii. Those traveling from countries with more than one reported case (i.e., SARS-CoV-2 had been confirmed to be present in that source country by that date),
- iii. Those traveling from countries with more than 100 reported cases (indicating community transmission was likely beginning), and
- iv. Those traveling from countries with more than 1000 reported cases (indicating widespread transmission)

For determining reported case counts at source locations for travelers, no cases were reported outside of China until January 13, 2020 (the date of the first reported case in Thailand). Over January 13 to January 21, cases were then reported in Japan, South Korea, Taiwan, Hong Kong, and the United States (<https://covid19.who.int/>). Subsequently, counts per country were tabulated daily by the Johns Hopkins Coronavirus Resource Center⁸¹ beginning January 22 (<https://coronavirus.jhu.edu/map.html>); we use that data from January 22 onwards and the WHO reports prior to January 22.

The number of travelers within each category arriving per month is shown in **Table S5**. This approach makes the conservative assumption that the probability a traveler is infected reflects the general countrywide infection rate of the source country at the time of travel (i.e., travelers are not more likely to be exposed than non-travelers in that source location) and does not account for complex travel itineraries (i.e., a traveler from a high risk source location transiting through a low risk source location would be grouped with other travelers from the low risk

source location). Consequently, the risk for viral importation is likely systematically underestimated. However, as the relative risk for viral importation will still scale with the number of travelers, comparisons among SSA countries can be informative (e.g., SSA countries with more travelers from countries with confirmed SARS-CoV-2 transmission are at higher risk for viral importation).

Table S5

Arrivals to SSA airports by the number of passenger seats and status of the SARS-CoV-2 pandemic at the origin at the time of travel

(see csv file: “Table S5 International Airtravel to SSA.csv”)

Data Fields:

1. country: Name of country
2. n_airports: Number of airports with flight data
3. month: January, February, March, April 2020; or total for all 4 months
4. total_n_seats: Total number of passenger seats on arriving aircraft
5. From source with cases > 0: Number of passengers arriving from source locations with 1 or more reported SARS-CoV-2 infection by the date of travel
6. From source with cases > 100: Number of passengers arriving from source locations more than 100 reported SARS-CoV-2 infection by the date of travel
7. From source with cases > 1000: Number of passengers arriving from source locations with more than 1000 reported SARS-CoV-2 infection by the date of travel

A6 | Subnational connectivity among countries in sub-Saharan Africa

6.1. Indicators of subnational connectivity

To allow comparison of the relative connectivity across countries, we use the friction surface estimates provided by Weiss et al.²⁸ as a relative measure of the rate of human movement between subregions of a country. For connectivity within subregions of a country (e.g., transport from a city to the rural periphery), we use as an indicator the population weighted mean travel time to the nearest urban center (i.e., population density > 1,500 per square kilometer or a density of built-up areas > 50% coincident with population > 50,000) within administrative-2 units⁷⁴. For some countries, estimates at administrative-2 units were unavailable (Comoros, Cape Verde, Lesotho, Mauritius, Mayotte, and Seychelles); estimates at the administrative-1 unit level were used for these cases (these were all island nations, with the exception of Lesotho).

6.2. Metapopulation model methods

Once SARS-CoV-2 has been introduced into a country, the degree of spread of the infection within the country will be governed by subnational mobility: the pathogen is more likely to be introduced into a location where individuals arrive more frequently than one where incoming travellers are less frequent. Large-scale consistent measures of mobility remain rare. However, recently, estimates of accessibility have been produced at a global scale²⁸. Although this is unlikely to perfectly reflect mobility within countries, especially as interventions and travel restrictions are put in place, it provides a starting point for evaluating the role of human mobility in shaping the outbreak pace across SSA. We use the inverse of a measure of the cost of travel between the centroids of administrative level 2 spatial units to describe mobility between locations (estimated by applying the `costDistance` function in the `gdistance` package in R to the friction surfaces supplied in ref²⁸). With this, we develop a metapopulation model for each country to develop an overview of the possible range of trajectories of unchecked spread of SARS-CoV-2.

We assume that the pathogen first arrives into each country in the administrative 2 level unit with the largest population (e.g., the largest city) and the population in each administrative 2 level (of size N_j) is entirely susceptible at the time of arrival. We then track spread within and between each of the administrative 2 level units of each country. Within each administrative 2 level unit, dynamics are governed by a discrete time Susceptible (S), Infected (I) and Recovered (R) model with a time-step of ~ 1 week, which is broadly consistent with the serial interval of SARS-CoV-2. Within the spatial unit indexed j , with total size N_j , the number of infected individuals in the next time-step is defined by:

$$I_{j,t+1} = \beta I_{j,t}^\alpha S_{j,t}/N_j + \iota_{j,t}$$

where β captures the magnitude of transmission over the course of one discrete time-step, and since the discrete time-step chosen is set to approximate the serial interval of the virus, this will reflect the R_0 of SARS-CoV-2, and is thus set to 2.5; the exponent $\alpha = 0.97$ is used to capture the effects of discretization⁸², and $\iota_{j,t}$ captures the introduction of new infections into site j at time t . Susceptible and recovered individuals are updated according to:

$$S_{j,t+1} = S_{j,t} + wR_{j,t} - I_{j,t+1} + b$$

$$R_{j,t+1} = (1 - w)R_{j,t} + I_{j,t}$$

where b reflects the introduction of new susceptible individuals resulting from the birth rate, set to reflect the most recent estimates for that country from the World Bank Data Bank (<https://data.worldbank.org/indicator/SP.DYN.CBRT.IN>), and w reflects the rate of waning of immunity. The population is initiated with $S_{j,1} = N_j R_{j,1} = 0$ and $I_{j,1} = 0$ except for the spatial unit corresponding to the largest population size N_j for each country, as this is assumed to be the location of introduction; for this spatial unit, we set $I_{j,1} = 1$.

We make the simplifying assumption that mobility linking locations i and j , denoted $c_{i,j}$, scales with the inverse of the cost of travel between sites i and j evaluated according to the friction surface provided in²⁸. The introduction of an infected individual into location j is then defined by a draw from a Bernoulli distribution following:

$$\iota_{j,t} \sim \text{Bern}(1 - \exp(-\sum_1^L c_{i,j} I_{i,t} / N_i))$$

where L is the total number of administrative 2 units in that country, and the rate of introduction is the product of connectivity between the focal location and each other location multiplied by the proportion of population in each other location that is infected.

Some countries show rapid spread between administrative units within the country (e.g., a country with parameters that broadly reflect those available for Malawi, **Figure S7**), while in others (e.g., reflecting Madagascar), connectivity may be so low that the outbreak may be over in the administrative unit of the largest size (where it was introduced) before introductions successfully reach other poorly connected administrative units. Where duration of immunity is sufficiently long, the result may be a hump shaped relationship between the proportion of the population that is infected after 5 years and the time to the first local extinction of the pathogen (**Figure S7**, right top). In countries with lower connectivity (e.g., that might resemble Madagascar), local outbreaks can go extinct rapidly before travelling very far; in other countries (e.g., that might resemble Gabon), the pathogen goes extinct rapidly because it travels rapidly and rapidly depletes susceptible individuals everywhere. The U-shaped pattern diminishes as the rate of waning of immunity increases, replaced by a monotonic negative relationship. With sufficiently rapid waning of immunity, local extinction ceases to occur in the absence of control efforts.

The impact of the pattern of travel between centroids is echoed by the pattern of travel within administrative districts: countries where the pathogen does not reach a large fraction of the administrative 2 units within the country in 5 years are also those where within administrative unit travel is low (**Figure S7**, right bottom).

These simulations provide a window onto qualitative patterns expected for subnational spread of the pandemic virus, but there is no clear way of calibrating the absolute rate of travel between regions of relevance for SARS-CoV-2, further complicated by remaining uncertainties around rates of waning of immunity. Thus, the time-scales of these simulations should be considered in relative, rather than absolute terms. Variation in lockdown effectiveness, or other changes in mobility for a given country may also compromise relative comparisons, as might large volumes of land-border crossings in some settings, which we have not accounted for here. Variability in case reporting complicates clarifying this (**Figure S8**) but we highlight countries with less connectivity (i.e., less synchronous outbreaks expected) relative to the median among SSA countries and with older populations (i.e., a greater proportion in higher risk age groups) (**Figure S9**).

While faster waning of immunity will act to increase the rate of spread of the infection, resulting in a higher proportion infected after one year, control efforts will generally act to slow the rate of spread of the infection (**Figure S10**). As different countries are likely to have differently effective control efforts (**Figure S10B-E**), this precludes making country specific predictions as to the relative impact of control efforts on delay.

Figure S7

Pace of the outbreak

Each grey line on the left hand panels indicates the total infected across all administrative units in a metapopulation simulation with parameters reflecting the country indicated by the plot title, assuming interventions are constant and that immunity does not wane. Increases after the first peak indicate the pathogen reaching a new administrative 2 unit. In Malawi-like settings (higher connectivity), more administrative units are reached rapidly, whereas in Madagascar-like settings (lower connectivity), a lower proportion of the administrative units are reached by a given time, as fewer introductions occur before the outbreak has burned out in the administrative 2 unit with the largest population. More generally, rapid disappearance of the outbreak (top right hand plot, y axis shows time to extinction) could either indicate rapid spread with a high proportion of the countries' population reached (top right hand plot, x axis) or slow spread, with many administrative units unreached, and therefore remaining susceptible (grey horizontal bars indicate quartiles across 100 simulations); noting that shorter duration of immunity reduces the probability of extinction. The pattern of between-administrative unit travel also echoes travel time within administrative units (lower panel, right hand side, x axis indicates fraction of administrative units unreached, and upper panel indicates travel time in hours to the nearest city of 50,000 or more people); grey horizontal bars again reflect quartiles across 100 simulations.

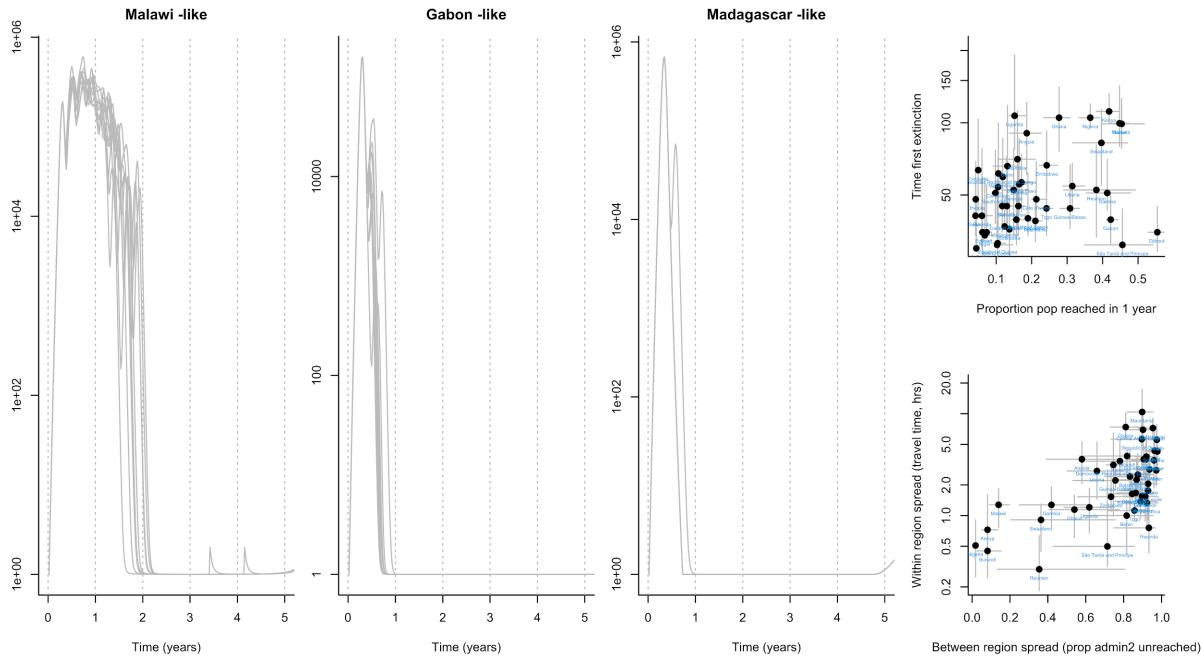


Figure S8

Cases and testing vs. the pace of the outbreak

The total number of confirmed cases reported by country (x axis, left, as reported for June 28th by Africa CDC) and the test positivity (x axis, right, defined as the total number of confirmed cases divided by the number of tests run, as reported by Africa CDC, likewise) show no significant relationship with the proportion of the population estimated to be infected after one year using the metapopulation simulation described in A6, assuming no waning of immunity:

respectively, $\rho = -0.04, p > 0.5, df = 41$ and $\rho = 0.02, p > 0.5, df = 41$

All else equal, a positive relationship is expected; however, both uncertainty in case numbers, and uncertainty associated with the simulation might both drive the absence of a signal.

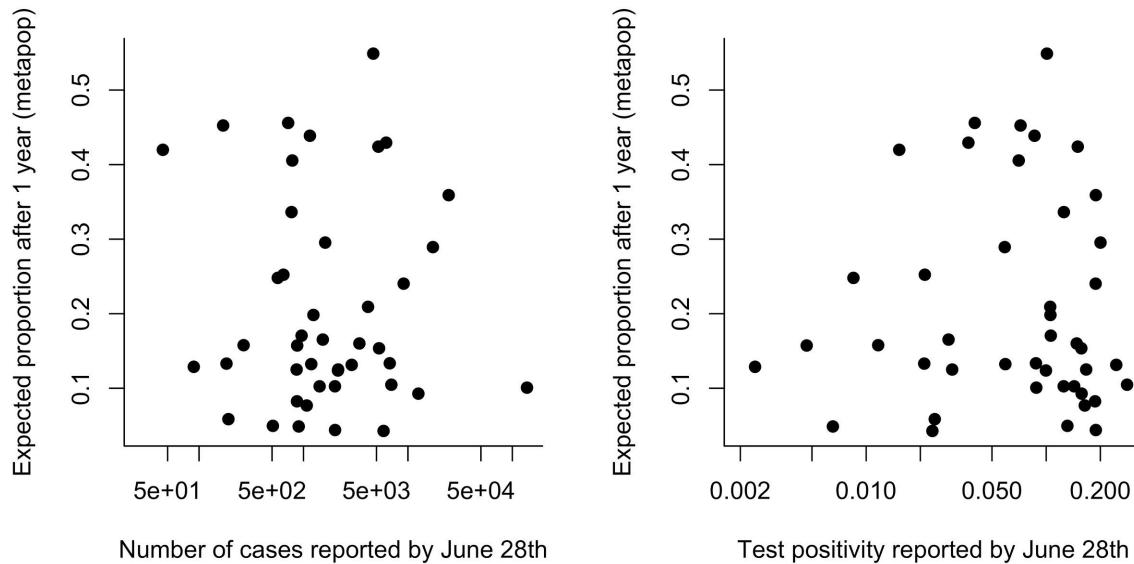
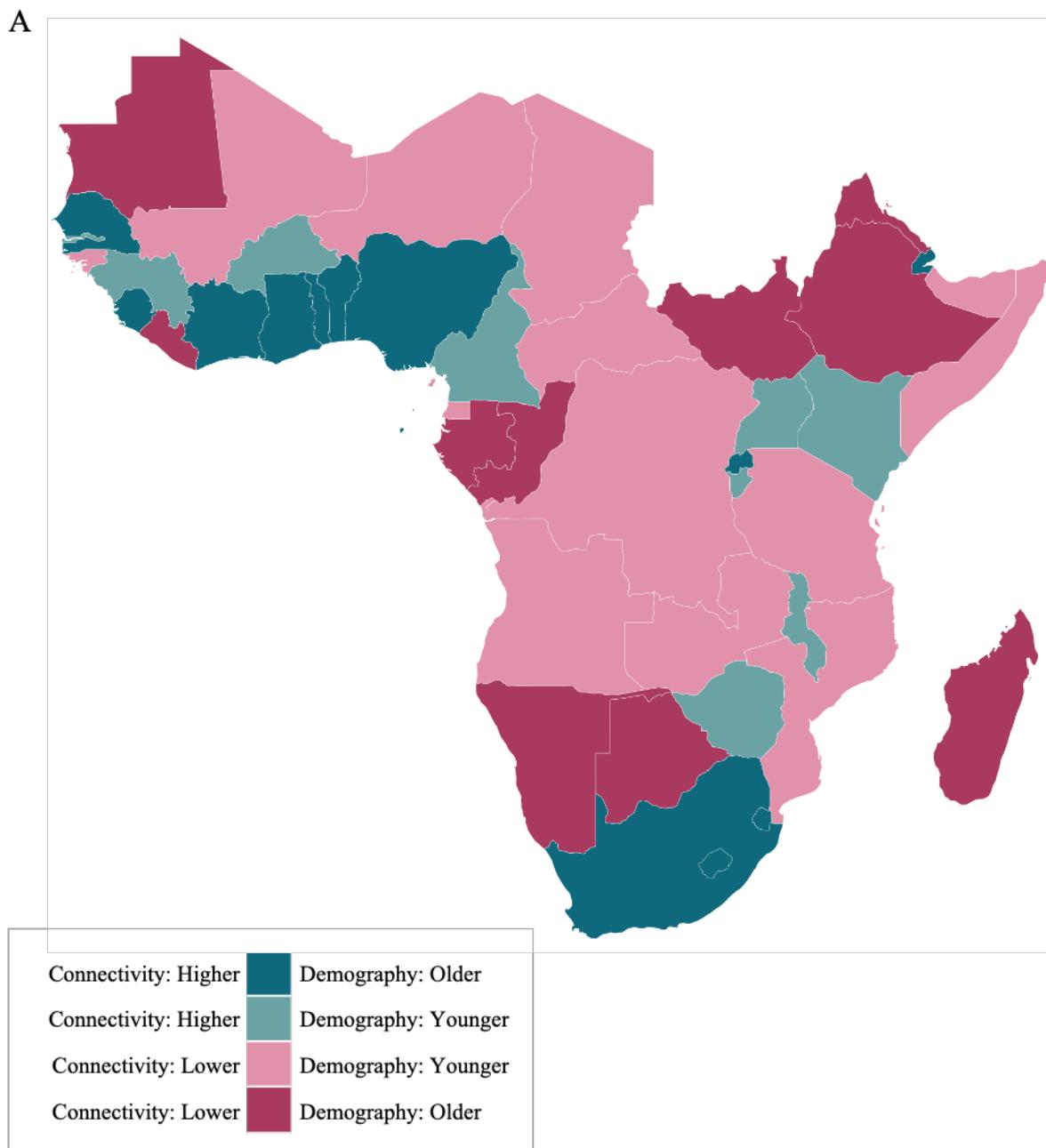


Figure S9

Bivariate example of expected pace versus expected burden at the national level in SARS-CoV-2 outbreaks in sub-Saharan Africa

Countries are colored by with respect to indicators of their expected epidemic pace (using as an example subnational connectivity in terms of travel time to nearest city) and potential burden (using as an example the proportion of the population over age 50). **A:** In pink, countries with less connectivity (i.e., less synchronous outbreaks) relative to the median among SSA countries; in blue, countries with more connectivity; darker colors show countries with older populations (i.e., a greater proportion in higher risk age groups).



(Figure S9 continued)

B: Dotted lines show the median; in the upper right, in dark pink, countries are highlighted due to their increased potential risk for an outbreak to be prolonged (see metapopulation model methods) and high burden (see burden estimation methods).

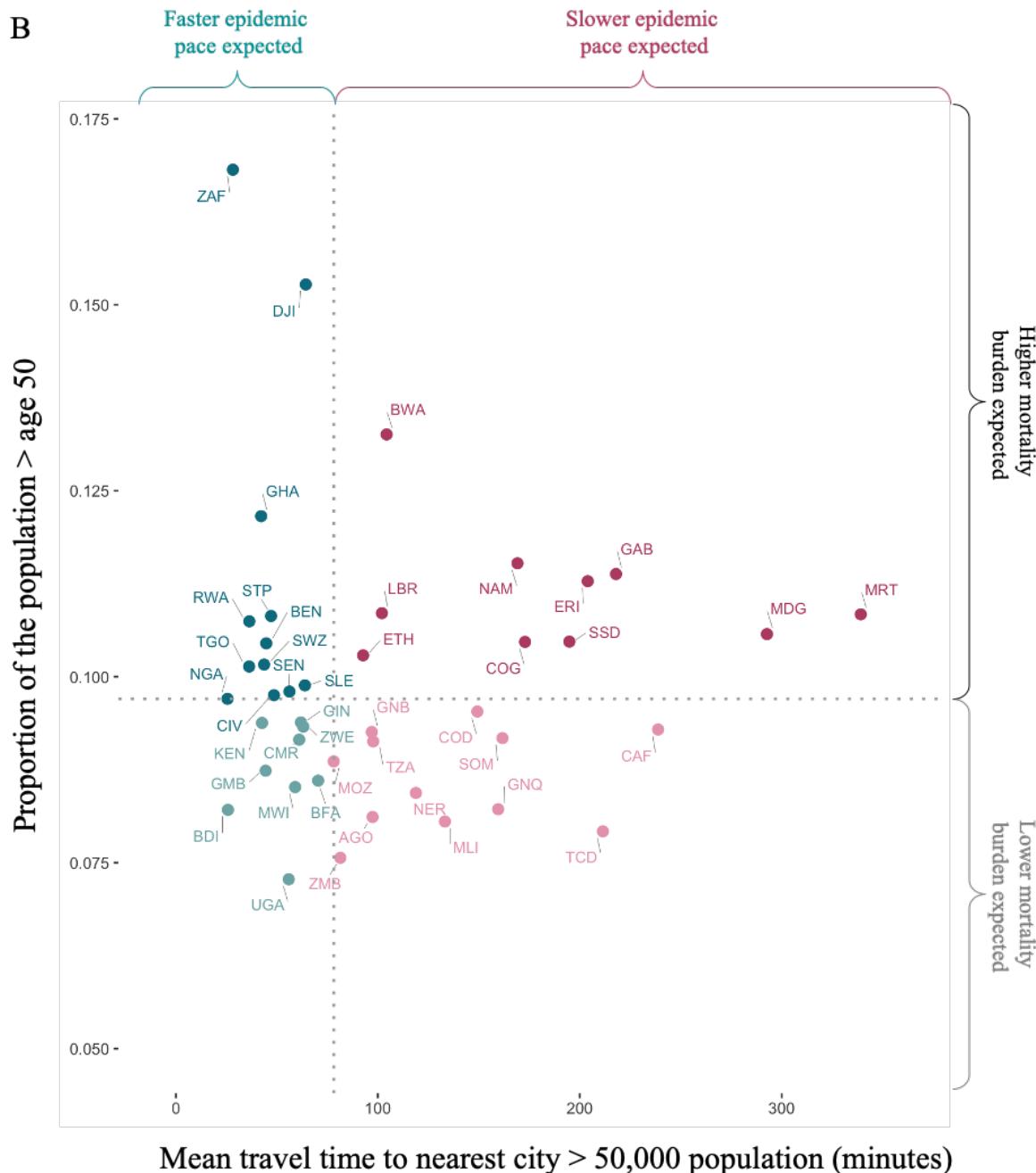


Figure S10

A. Impact of waning of immunity and the introduction of control efforts (Non-Pharmaceutical Interventions, NPIs) on spatial spread

The left panel indicates the proportion of the population infected after one year in the absence (x-axis) or presence (y-axis) of waning of immunity (duration of immunity taken to be ~40 weeks, i.e., $w=1/40$, reflecting estimates for other coronaviruses HCoV-OC43 and HCoV-HKU1) across countries in SSA; grey horizontal lines indicate quartiles across 100 simulations. All points are above the 0,1 line indicating that waning of immunity accelerates spatial spread, but overall rankings remain. The central panel indicates the proportion of the population infected after one year in the absence (x axis) or presence (y axis) of control efforts with 12 weeks of a 20% reduction in transmission as an exemplar. All points are below the 0,1 line indicating a lower proportion infected as a result of control efforts.

Similarly, all else equal, NPIs have the effect of prolonging outbreaks (i.e., more weeks until the first extinction) (right hand panel, all points above the 0,1 line). This effect may vanish if waning occurs sufficiently rapidly (e.g., a duration of immunity of 40 weeks yields no local extinction).

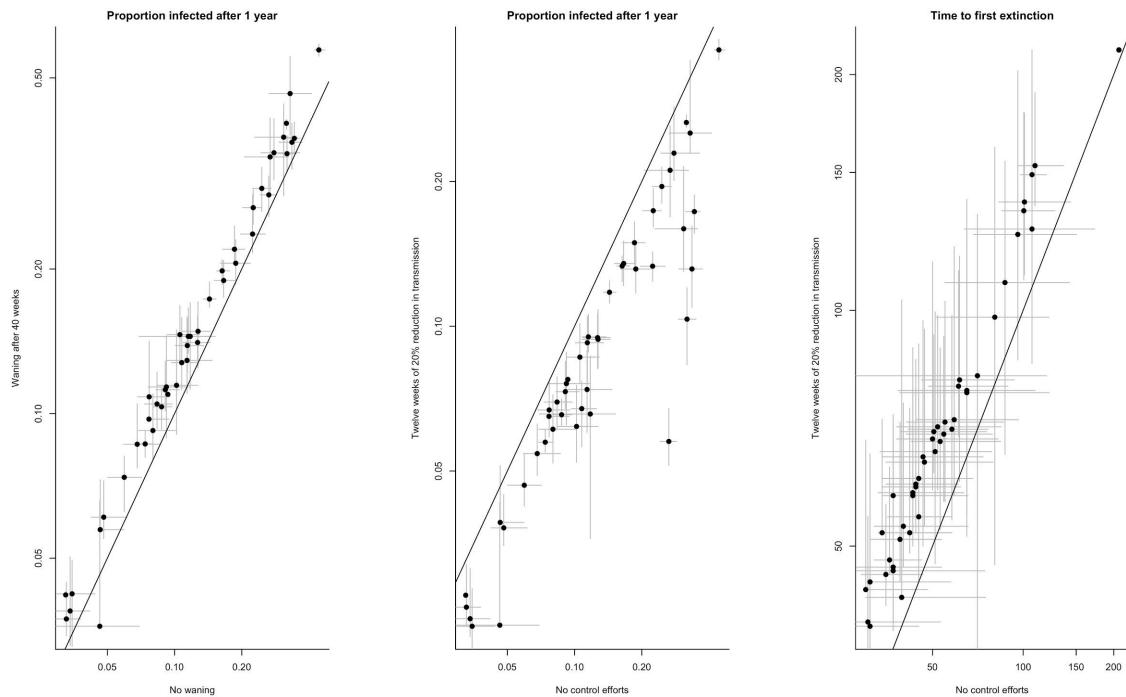


Figure S10**B. Time course of the range of policies deployed across different countries.**

A composite score of government response (left), interventions for containment (middle) and economic support provided (right) each scored from 0-100, provided by the University of Oxford Blavatnik School of Government⁸³, showing SSA countries (black lines) relative to other countries (grey lines)

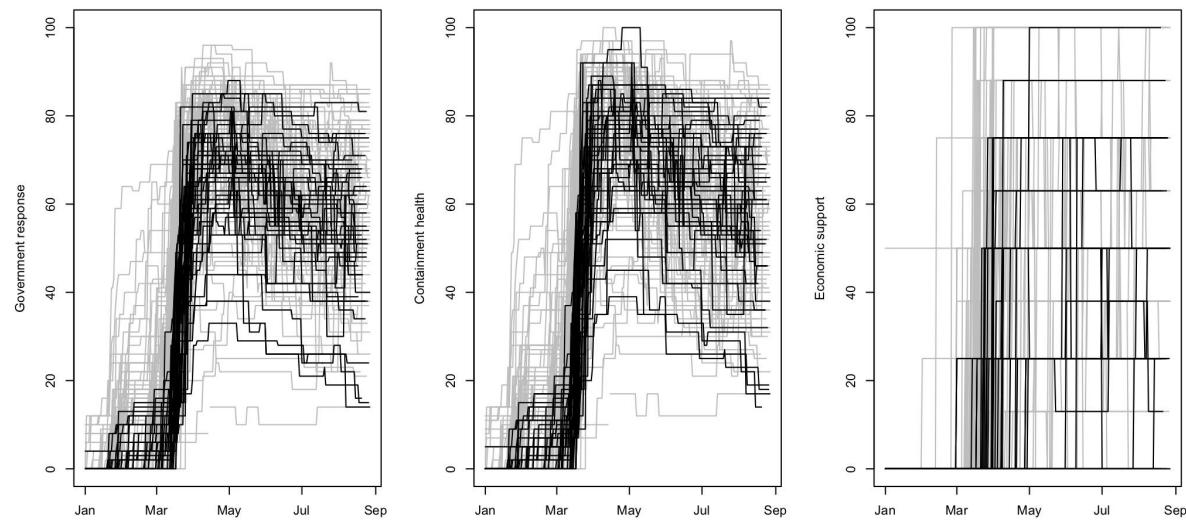


Figure S10

C. Comparison of policies implemented in SSA and google derived measures of mobility

The black line indicates a score of the magnitude of policies directed towards health containment for each country (plot title) on a scale from 0-100⁸³ with other SSA countries for which data on mobility was available (n = 24 of the 48 SSA countries) shown for comparison in grey; the red line indicates the percent reductions in mobility to work relative to baseline⁸⁴ for that country (similar patterns seen for other mobility measures). The vertical blue line shows the day on which 10 cases were exceeded based on the Johns Hopkins dashboard data⁸⁵.

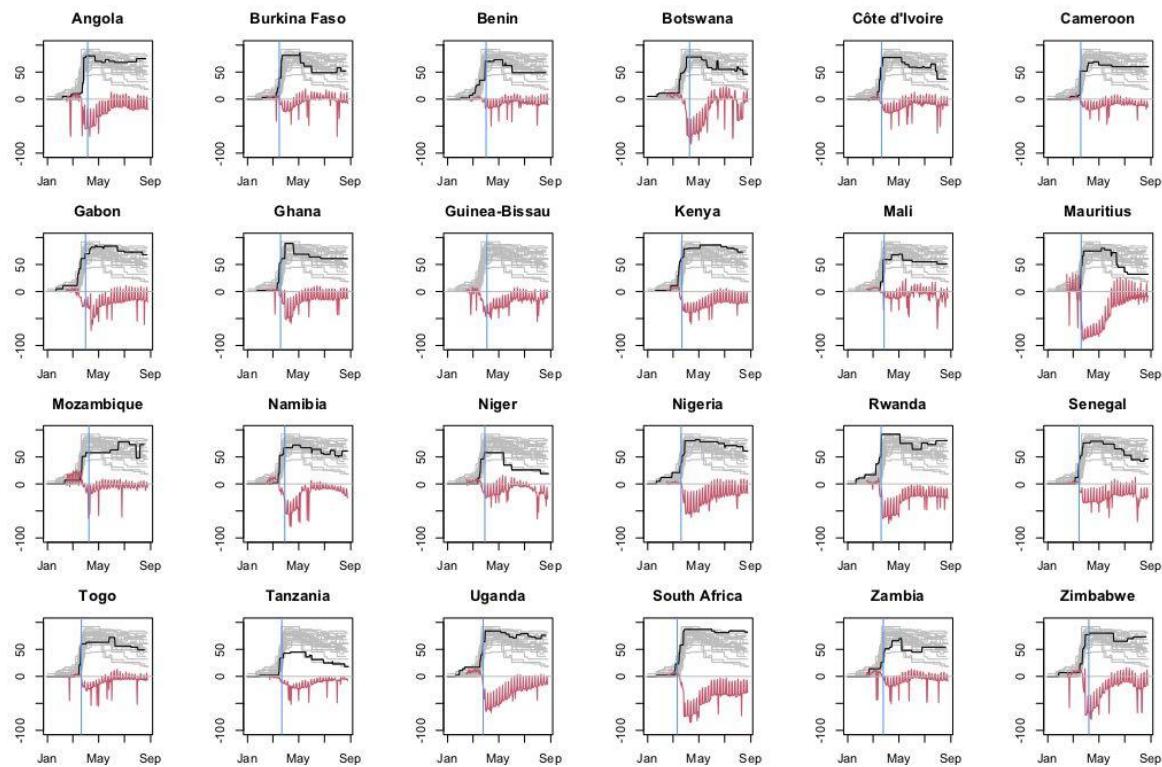


Figure S10

D. Indication of the impact of intervention policies in sub-Saharan Africa

Correlation between magnitude of policies directed towards health containment (black line in **Figure S10C**) and percent reductions in mobility (red line in **Figure S10C**), with countries ordered by cumulative SARS-CoV-2 cases reported to date. Policy effectiveness is indicated by a negative correlation (noting that health containment contains a range of different interventions; the range of policies directed, e.g., specifically at work have much lower variance). Ability to implement control does not necessarily align with fewer reported cases (largest negative correlations are not concentrated among the countries with the fewest cases shown at the bottom of the figure), but this relationship is difficult to interpret given likely heterogeneity in case reporting as well as varied targets of health containment strategies. Similar results are obtained with different lags between policy introduction and mobility.

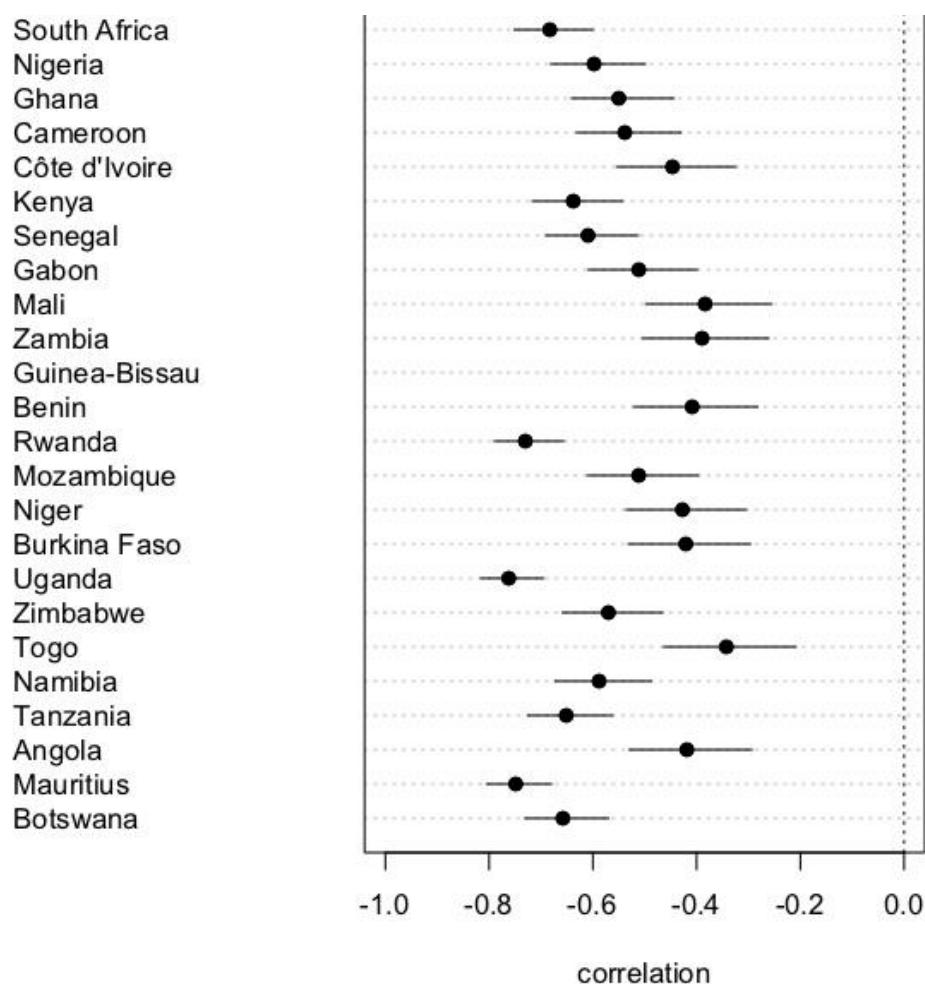


Figure S10

E. Comparison of reductions in transmission with another directly transmitted infection

Monthly measles incidence (y-axis) between 2011 and 2019⁸⁶ is shown in gray, and the first 6 months of 2020 (months on the x-axis) shown in red for countries for which data is available in SSA ($n = 34$ of 48 SSA countries). China and Germany (which have been relatively successful in controlling the virus) shown for comparison at the bottom right. Although multi-annual features might drive measles incidence (e.g., dynamics in Madagascar are largely dominated by a honeymoon outbreak that occurred in 2018-2019⁸⁷) for countries that slowed the SARS-CoV-2 pandemic, signatures of reduction in measles can be identified (e.g., Germany and China; similar patterns are seen in Viet Nam). Although pathogens more similar to SARS-CoV-2 (e.g., influenza) might provide a better focus, measles data is used as an exemplar as it is available broadly in our focal region.



A7 | Modeling epidemic trajectories in scenarios where transmission rate depends on climate

7.1 Climate data sourcing: Variation in humidity in SSA

Specific humidity data for selected urban centers comes from ERA5 using an average climatology (1981-2017)⁵⁸; we do not consider year-to-year climate variations. Selected cities ($n = 56$) were chosen to represent the major urban areas in SSA. The largest city in each SSA country was included as well as any additional cities that were among the 25 largest cities or busiest airports in SSA.

7.2 Methods for climate driven modelling of SARS-CoV-2

We use a climate-driven SIRS (Susceptible-Infected-Recovered-Susceptible) model to estimate epidemic trajectories (i.e., the time of peak incidence) in different cities in 2020, assuming no control measures are in place or a 10% or 20% reduction in R_0 beginning 2 weeks after the total reported cases for a country exceed 10 cases^{27,88}. The model is given by:

$$\frac{dS}{dt} = \frac{N - S - L}{L} - \frac{\beta(t)IS}{N}$$

$$\frac{dI}{dt} = \frac{\beta(t)IS}{N} - \frac{I}{D}$$

where S is the susceptible population, I is the infected population and N is the total population. D is the mean infectious period, set at 5 days following ref^{27,60}.

To investigate the effects on epidemic trajectories of a climate-dependency of SARS-CoV-2 on cities with the climate patterns of the selected cities in SSA, we use parameters from the most climate-dependent scenario in ref²⁷, based on the endemic betacoronavirus HKU1 in the USA. In this scenario L , the duration of immunity, is found to be 66.25 weeks (i.e., greater than 1 year and such that waning immunity does not affect timing of the epidemic peak). We initially select a range where R_0 declines from $R_{0\max} = 2.5$ to $R_{0\min} = 1.5$ (i.e., transmission declines 40% at high humidity) as this exceeds the range observed for influenza and other coronaviruses for which there is data available (from the USA). $R_{0\max} = 2.5$ is chosen as 2.5 is oft cited as the approximate R_0 for SARS-CoV-2. Thus, we initially assume that the climate dependence of SARS-CoV-2 in SSA will not greatly exceed that of other known coronaviruses from the US context. Then we explore the effects of different degrees of climate dependency (i.e., wider ranges between $R_{0\max} = 2.5$ to $R_{0\min} = 1.5$ and scenarios where $R_{0\min}$ approaches 1) (Figure S11).

Transmission is governed by $\beta(t)$ which is related to the basic reproduction number R_0 by $R_0(t) = \beta(t)D$. The basic reproduction number varies based on the climate and is related to specific humidity according to the equation:

$$R_0 = \exp(a * q(t) + \log(R_{0max} - R_{0min})) + R_{0min}$$

where $q(t)$ is specific humidity⁵⁸ and a is set at -227.5 based on estimated HKU1 parameters²⁷. We assume the time of introduction for cities to be the date at which the total reported cases for a country exceed 10 cases.

7.3 Sensitivity analysis

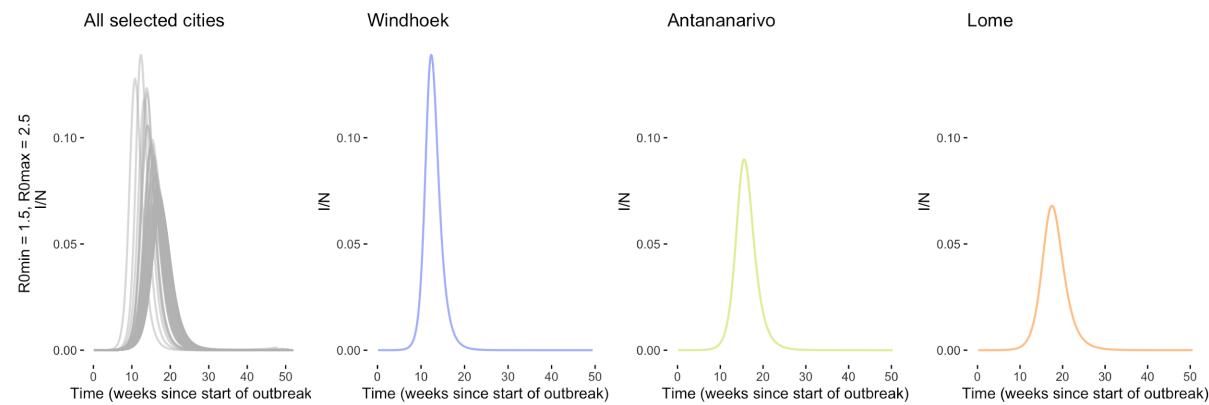
Selecting an R_{0min} value of 1, such that epidemic growth stops at high humidities is likely implausible as simulations indicate no outbreaks would occur in cities such as Antananarivo (countered by the observation that SARS-CoV-2 outbreaks did in fact occur) (**Figure S11B**) (see **Table S1** for reported case counts at the country level). Expanding the range between R_{0min} and R_{0max} by increasing R_{0max} results in epidemic peaks being reached earlier after outbreak onset, but does not increase the difference in timing between cities with different climates (**Figure S11C**; e.g., the difference in timing between peaks in Windhoek and Lome is similar in panels A and C). Finally, we explored scenarios where R_{0min} was between 1.0 and 1.5. When $R_{0min} > 1.1$, epidemic peaks are seen in each SSA city with the difference in timing of the peak growing larger when smaller values of R_{0min} are selected (**Figure S11D**). However, the difference in timing, even when small values of R_{0min} are selected, is a maximum of 25 weeks (i.e., a shorter time period than the interval between present time (October 2020) and the beginning of outbreaks (approx. March 2020)) and rapidly reduces to only a few weeks as R_{0min} approaches 1.5.

Figure S11

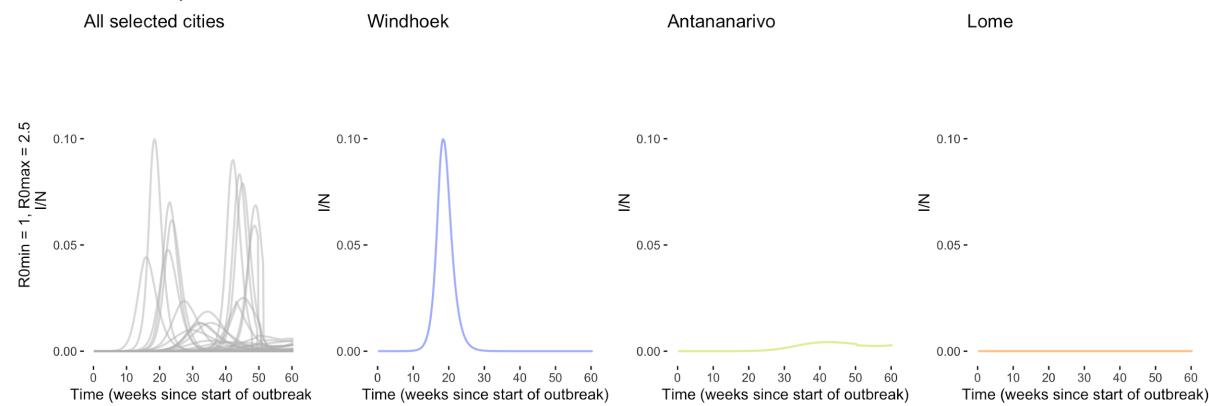
Transmission climate-dependency and sensitivity to $R_{0\min}$ and $R_{0\max}$ value selection

Transmission (R_0) declines with increasing specific humidity from $R_{0\max}$ to $R_{0\min}$. Three exemplar cities with low, intermediate, and high average specific humidity are shown across rows (Windhoek, Antananarivo, and Lome, respectively)

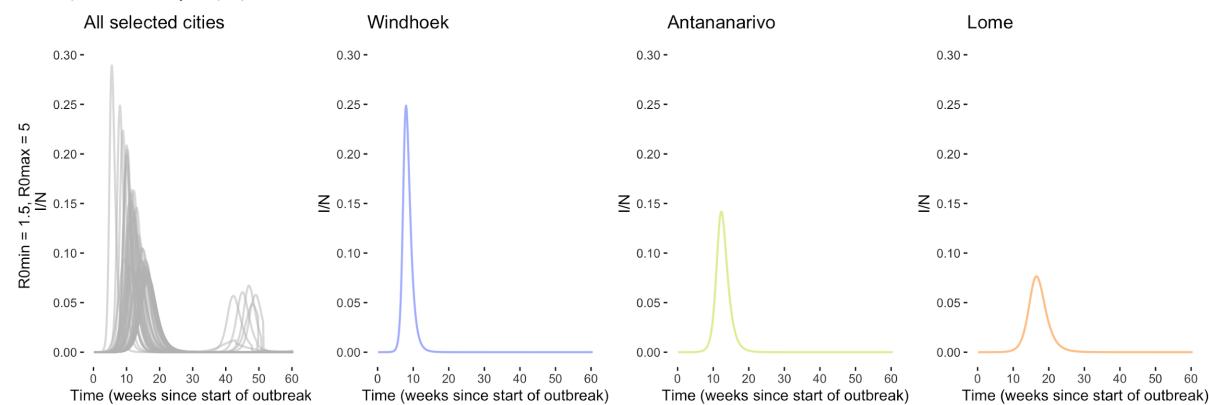
A. $R_{0\min} = 1.5$, $R_{0\max} = 2.5$



B. $R_{0\min} = 1.0$, $R_{0\max} = 2.5$



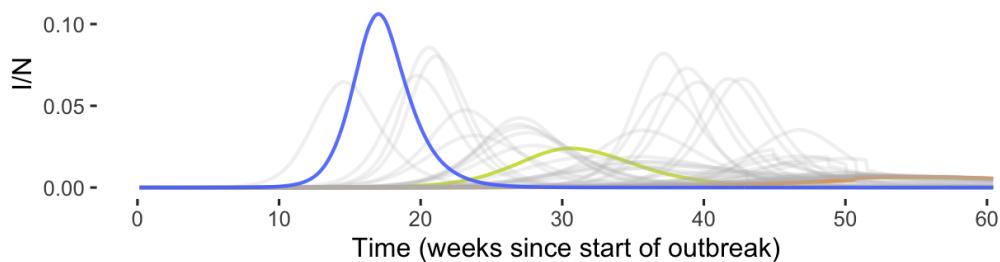
C. $R_{0\min} = 1.5$, $R_{0\max} = 5$



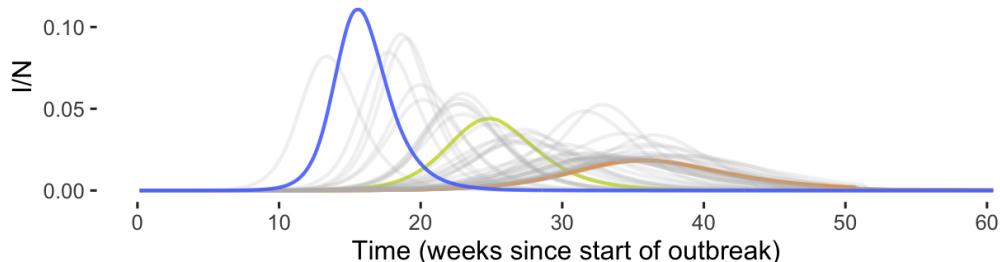
D. Variation in peak size and timing when $1.0 < R_{0min} < 1.5$.

Three exemplar cities with low, intermediate, and high average specific humidity are shown across rows (Windhoek (blue), Antananarivo (green), and Lome (orange), respectively)

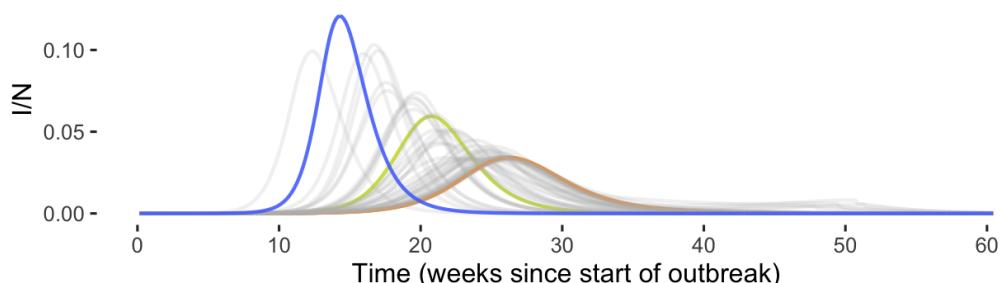
$$R_{0min} = 1.1, R_{0max} = 2.5$$



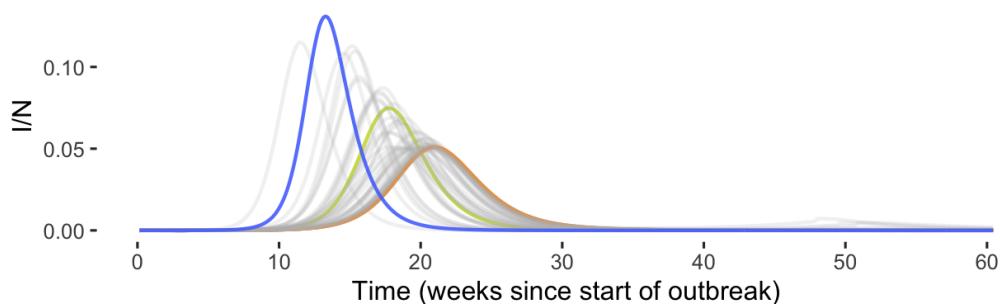
$$R_{0min} = 1.2, R_{0max} = 2.5$$



$$R_{0min} = 1.3, R_{0max} = 2.5$$



$$R_{0min} = 1.4, R_{0max} = 2.5$$



References

1. Mecenas, P., Bastos, R., Vallinoto, A. & Normando, D. Effects of temperature and humidity on the spread of COVID-19: A systematic review. doi:10.1101/2020.04.14.20064923.
2. Salje, H. *et al.* Estimating the burden of SARS-CoV-2 in France. *Science* (2020) doi:10.1126/science.abc3517.
3. Rinaldi, G. & Paradisi, M. An empirical estimate of the infection fatality rate of COVID-19 from the first Italian outbreak. doi:10.1101/2020.04.18.20070912.
4. Verity, R. *et al.* Estimates of the severity of coronavirus disease 2019: a model-based analysis. *Lancet Infect. Dis.* **20**, 669–677 (2020).
5. Africa CDC - COVID-19 Daily Updates. *Africa CDC* <https://africacdc.org/covid-19/>.
6. Mortality Analyses. *Johns Hopkins Coronavirus Resource Center* <https://coronavirus.jhu.edu/data/mortality>.
7. Uyoga, S. *et al.* Seroprevalence of anti-SARS-CoV-2 IgG antibodies in Kenyan blood donors. doi:10.1101/2020.07.27.20162693.
8. Majiya, H. *et al.* Seroprevalence of COVID-19 in Niger State. doi:10.1101/2020.08.04.20168112.
9. Chibwana, M. G. *et al.* High SARS-CoV-2 seroprevalence in Health Care Workers but relatively low numbers of deaths in urban Malawi. *medRxiv* (2020) doi:10.1101/2020.07.30.20164970.
10. Mbow, M. *et al.* COVID-19 in Africa: Dampening the storm? *Science* **369**, 624–626 (2020).
11. Skrip, L. A. *et al.* Seeding COVID-19 across sub-Saharan Africa: an analysis of reported importation events across 40 countries. doi:10.1101/2020.04.01.20050203.
12. Deng, X. *et al.* Case fatality risk of novel coronavirus diseases 2019 in China. doi:10.1101/2020.03.04.20031005.
13. Collaborative, T. O. *et al.* OpenSAFELY: factors associated with COVID-19-related hospital death in the linked electronic health records of 17 million adult NHS patients.

- doi:10.1101/2020.05.06.20092999.
14. COVID-19 significantly impacts health services for noncommunicable diseases.
<https://www.who.int/news-room/detail/01-06-2020-covid-19-significantly-impacts-health-services-for-noncommunicable-diseases>.
 15. Maintaining essential health services: operational guidance for the COVID-19 context.
<https://www.who.int/publications/i/item/covid-19-operational-guidance-for-maintaining-essential-health-services-during-an-outbreak>.
 16. Santoli, J. M. Effects of the COVID-19 Pandemic on Routine Pediatric Vaccine Ordering and Administration — United States, 2020. *MMWR Morb. Mortal. Wkly. Rep.* **69**, (2020).
 17. Roberton, T. *et al.* Early estimates of the indirect effects of the COVID-19 pandemic on maternal and child mortality in low-income and middle-income countries: a modelling study. *Lancet Glob Health* **8**, e901–e908 (2020).
 18. Walker, P. G. T. *et al.* The impact of COVID-19 and strategies for mitigation and suppression in low- and middle-income countries. *Science* (2020)
doi:10.1126/science.abc0035.
 19. Pei, S., Kandula, S. & Shaman, J. Differential Effects of Intervention Timing on COVID-19 Spread in the United States. *medRxiv* (2020) doi:10.1101/2020.05.15.20103655.
 20. Lai, S. *et al.* Assessing the effect of global travel and contact reductions to mitigate the COVID-19 pandemic and resurgence. doi:10.1101/2020.06.17.20133843.
 21. Nepomuceno, M. R. *et al.* Besides population age structure, health and other demographic factors can contribute to understanding the COVID-19 burden. *Proceedings of the National Academy of Sciences of the United States of America* vol. 117 13881–13883 (2020).
 22. Clark, A. *et al.* Global, regional, and national estimates of the population at increased risk of severe COVID-19 due to underlying health conditions in 2020: a modelling study. *Lancet Glob Health* (2020) doi:10.1016/S2214-109X(20)30264-3.
 23. Ghisolfi, S. *et al.* Predicted COVID-19 fatality rates based on age, sex, comorbidities, and

- health system capacity. doi:10.1101/2020.06.05.20123489.
24. [No title]. <https://www.imperial.ac.uk/media/imperial-college/medicine/mrc-gida/2020-05-12-COVID19-Report-22.pdf>.
 25. Kapata, N. *et al.* Is Africa prepared for tackling the COVID-19 (SARS-CoV-2) epidemic. Lessons from past outbreaks, ongoing pan-African public health efforts, and implications for the future. *International journal of infectious diseases: IJID: official publication of the International Society for Infectious Diseases* vol. 93 233–236 (2020).
 26. Nchega, J. B. *et al.* From Easing Lockdowns to Scaling-Up Community-Based COVID-19 Screening, Testing, and Contact Tracing in Africa - Shared Approaches, Innovations, and Challenges to Minimize Morbidity and Mortality. *Clin. Infect. Dis.* (2020) doi:10.1093/cid/ciaa695.
 27. Baker, R. E., Yang, W., Vecchi, G. A., Metcalf, C. J. E. & Grenfell, B. T. Susceptible supply limits the role of climate in the early SARS-CoV-2 pandemic. *Science* (2020) doi:10.1126/science.abc2535.
 28. Weiss, D. J. *et al.* A global map of travel time to cities to assess inequalities in accessibility in 2015. *Nature* **553**, 333–336 (2018).
 29. Tatem, A. J. WorldPop, open data for spatial demography. *Scientific Data* vol. 4 (2017).
 30. [No title]. <https://washdata.org/sites/default/files/documents/reports/2019-05/JMP-2018-core-questions-for-household-surveys.pdf>.
 31. Korevaar, H. M. *et al.* Quantifying the impact of US state non-pharmaceutical interventions on COVID-19 transmission. doi:10.1101/2020.06.30.20142877.
 32. O'Driscoll, M. *et al.* Age-specific mortality and immunity patterns of SARS-CoV-2 infection in 45 countries. *medRxiv* 2020.08.24.20180851 (2020).
 33. Mikkelsen, L. *et al.* A global assessment of civil registration and vital statistics systems: monitoring data quality and progress. *Lancet* **386**, 1395–1406 (2015).
 34. Gilbert, M. *et al.* Preparedness and vulnerability of African countries against importations of

- COVID-19: a modelling study. *Lancet* **395**, 871–877 (2020).
35. Haider, N. *et al.* Passengers' destinations from China: low risk of Novel Coronavirus (2019-nCoV) transmission into Africa and South America. *Epidemiol. Infect.* **148**, e41 (2020).
 36. Takahashi, S. *et al.* Reduced vaccination and the risk of measles and other childhood infections post-Ebola. *Science* vol. 347 1240–1242 (2015).
 37. Cash, R. & Patel, V. Has COVID-19 subverted global health? *Lancet* **395**, 1687–1688 (2020).
 38. Adams, J. G. & Walls, R. M. Supporting the Health Care Workforce During the COVID-19 Global Epidemic. *JAMA* (2020) doi:10.1001/jama.2020.3972.
 39. Kilmarx, P. H. *et al.* Ebola virus disease in health care workers--Sierra Leone, 2014. *MMWR Morb. Mortal. Wkly. Rep.* **63**, 1168–1171 (2014).
 40. Covid 19 Mauritius Dashboard. *Google Data Studio*
<http://datastudio.google.com/reporting/510dbd29-25cd-4fbb-a47a-68effeda6cf5/page/0z6JB?feature=opengraph>.
 41. Rwanda: WHO Coronavirus Disease (COVID-19) Dashboard. <https://covid19.who.int>.
 42. Mina, M. J. *et al.* A Global Immunological Observatory to meet a time of pandemics. *Elife* **9**, (2020).
 43. World Population Prospects - Population Division - United Nations.
<https://population.un.org/wpp/Download/Standard/Population/>.
 44. The Novel Coronavirus Pneumonia Emergency Response Epidemiology Team. The Epidemiological Characteristics of an Outbreak of 2019 Novel Coronavirus Diseases (COVID-19) — China, 2020. *CCDCW* **2**, 113–122 (2020).
 45. Team, C. C.-19 R. *et al.* Severe Outcomes Among Patients with Coronavirus Disease 2019 (COVID-19) — United States, February 12–March 16, 2020. *MMWR. Morbidity and Mortality Weekly Report* vol. 69 343–346 (2020).
 46. Simonnet, A. *et al.* High prevalence of obesity in severe acute respiratory syndrome

- coronavirus-2 (SARS-CoV-2) requiring invasive mechanical ventilation. *Obesity* (2020) doi:10.1002/oby.22831.
47. Conticini, E., Frediani, B. & Caro, D. Can atmospheric pollution be considered a co-factor in extremely high level of SARS-CoV-2 lethality in Northern Italy? *Environ. Pollut.* **261**, 114465 (2020).
48. Mapping COVID-19 Risk Factors – Africa Center for Strategic Studies.
<https://africacenter.org/spotlight/mapping-risk-factors-spread-covid-19-africa/>.
49. Bringing Greater Precision to the COVID-19 Response. <https://precisionforcovid.org/africa>.
50. Griffith, G. *et al.* Collider bias undermines our understanding of COVID-19 disease risk and severity. doi:10.1101/2020.05.04.20090506.
51. Shankar, A. H. Nutritional Modulation of Malaria Morbidity and Mortality. *The Journal of Infectious Diseases* vol. 182 S37–S53 (2000).
52. Cooper, T. J., Woodward, B. L., Alom, S. & Harky, A. Coronavirus disease 2019 (COVID-19) outcomes in HIV/AIDS patients: a systematic review. *HIV Medicine* vol. 21 567–577 (2020).
53. Chanda-Kapata, P., Kapata, N. & Zumla, A. COVID-19 and malaria: A symptom screening challenge for malaria endemic countries. *Int. J. Infect. Dis.* **94**, 151–153 (2020).
54. Lesosky, M. & Myer, L. Modelling the impact of COVID-19 on HIV. *The Lancet. HIV* vol. 7 e596–e598 (2020).
55. Sherrard-Smith, E. *et al.* The potential public health consequences of COVID-19 on malaria in Africa. *Nat. Med.* **26**, 1411–1416 (2020).
56. The DHS Program - Quality information to plan, monitor and improve population, health, and nutrition programs. <https://dhsprogram.com>.
57. Chin, T. *et al.* U.S. county-level characteristics to inform equitable COVID-19 response. *medRxiv* (2020) doi:10.1101/2020.04.08.20058248.
58. Hersbach, H. *et al.* The ERA5 global reanalysis. *Q.J.R. Meteorol. Soc.* **64**, 29 (2020).

59. Report 23 - State-level tracking of COVID-19 in the United States. *Imperial College London* <http://www.imperial.ac.uk/medicine/departments/school-public-health/infectious-disease-epidemiology/mrc-global-infectious-disease-analysis/covid-19/report-23-united-states/>.
60. Kissler, S. M., Tedijanto, C., Goldstein, E., Grad, Y. H. & Lipsitch, M. Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. *Science* **368**, 860–868 (2020).
61. Jiwani, S. S. & Antiporta, D. A. Inequalities in access to water and soap matter for the COVID-19 response in sub-Saharan Africa. *Int. J. Equity Health* **19**, 82 (2020).
62. Stoler, J., Jepson, W. E. & Wutich, A. Beyond handwashing: Water insecurity undermines COVID-19 response in developing areas. *J. Glob. Health* **10**, 010355 (2020).
63. Makoni, M. Keeping COVID-19 at bay in Africa. *Lancet Respir Med* **8**, 553–554 (2020).
64. Tatem, A. J. et al. Ranking of elimination feasibility between malaria-endemic countries. *Lancet* **376**, 1579–1591 (2010).
65. Metcalf, C. J. E. et al. Transport networks and inequities in vaccination: remoteness shapes measles vaccine coverage and prospects for elimination across Africa. *Epidemiol. Infect.* **143**, 1457–1466 (2015).
66. Adepoju, P. Africa's struggle with inadequate COVID-19 testing. *The Lancet Microbe* vol. 1 e12 (2020).
67. Pindolia, D. K. et al. The demographics of human and malaria movement and migration patterns in East Africa. *Malar. J.* **12**, 397 (2013).
68. Qiu, Y., Chen, X. & Shi, W. Impacts of social and economic factors on the transmission of coronavirus disease (COVID-19) in China. doi:10.1101/2020.03.13.20035238.
69. Li, H. et al. Age-Dependent Risks of Incidence and Mortality of COVID-19 in Hubei Province and Other Parts of China. *Front. Med.* **7**, 190 (2020).
70. Global Burden of Disease Study 2017 (GBD 2017) Population Estimates 1950-2017 | GHDx. <http://ghdx.healthdata.org/record/ihme-data/gbd-2017-population-estimates-1950->

2017.

71. Brauer, M., Zhao, J. T., Bennett, F. B. & Stanaway, J. D. Global access to handwashing: implications for COVID-19 control in low-income countries.
doi:10.1101/2020.04.07.20057117.
72. Alegana, V. A. *et al.* National and sub-national variation in patterns of febrile case management in sub-Saharan Africa. *Nat. Commun.* **9**, 4994 (2018).
73. Murthy, S., Leligdowicz, A. & Adhikari, N. K. J. Intensive care unit capacity in low-income countries: a systematic review. *PLoS One* **10**, e0116949 (2015).
74. Linard, C., Gilbert, M., Snow, R. W., Noor, A. M. & Tatem, A. J. Population distribution, settlement patterns and accessibility across Africa in 2010. *PLoS One* **7**, e31743 (2012).
75. [No title]. <https://dl.acm.org/doi/10.5555/1953048.2078195>.
76. Jolliffe, I. T. & Cadima, J. Principal component analysis: a review and recent developments. *Philos. Trans. A Math. Phys. Eng. Sci.* **374**, 20150202 (2016).
77. Meyerowitz-Katz, G. & Merone, L. A systematic review and meta-analysis of published research data on COVID-19 infection-fatality rates. doi:10.1101/2020.05.03.20089854.
78. Wood, S. N. *Generalized Additive Models: An Introduction with R, Second Edition*. (CRC Press, 2017).
79. R: The R Project for Statistical Computing. <https://www.R-project.org/>.
80. Global Burden of Disease Study 2017 (GBD 2017) Population Estimates 1950-2017 | GHDx. <http://ghdx.healthdata.org/record/ihme-data/gbd-2017-population-estimates-1950-2017>.
81. Dong, E., Du, H. & Gardner, L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect. Dis.* **20**, 533–534 (2020).
82. Bjornstad, O. N., Finkenstadt, B. F. & Grenfell, B. T. Dynamics of Measles Epidemics: Estimating Scaling of Transmission Rates Using a Time Series SIR Model. *Ecological Monographs* vol. 72 169 (2002).

83. Coronavirus Government Response Tracker. <https://www.bsg.ox.ac.uk/research/research-projects/coronavirus-government-response-tracker>.
84. Google. COVID-19 Community Mobility Reports.
85. COVID-19 Map - Johns Hopkins Coronavirus Resource Center.
<https://coronavirus.jhu.edu/map.html>.
86. WHO | Measles and Rubella Surveillance Data. (2020).
87. Nimpa, M. M. *et al.* Measles outbreak in 2018-2019, Madagascar: epidemiology and public health implications. *Pan Afr. Med. J.* **35**, 84 (2020).
88. Shaman, J., Pitzer, V. E., Viboud, C., Grenfell, B. T. & Lipsitch, M. Absolute humidity and the seasonal onset of influenza in the continental United States. *PLoS Biol.* **8**, e1000316 (2010).