

1
2
3
4
5 **Leveraging Serosurveillance and Postmortem Surveillance to**
6 **Quantify the Impact of COVID-19 in Africa**
7

8 Nicole E. Kogan^{1,2,*} Shae Gantt^{1,*} David Swerdlow¹ Cécile Viboud³
9 Muhammed Semakula⁴ Marc Lipsitch¹ Mauricio Santillana^{1,2,†}

10
11 ¹ Department of Epidemiology, Harvard T. H. Chan School of Public Health, Boston, MA

12 ² Machine Intelligence Group for the Betterment of Health and the Environment, Network
13 Science Institute, Northeastern University, Boston, MA

14 ³ Division of International Epidemiology and Population Studies, Fogarty International Center,
15 National Institutes of Health, Bethesda, MD

16 ⁴ Rwanda Biomedical Centre, Kigali, Rwanda

17
18 * These authors contributed equally to this manuscript

19 † **Correspondence to:** Mauricio Santillana (m.santillana@northeastern.edu)

20
21 **Keywords:** COVID-19, serology, postmortem surveillance, underreporting, cumulative
22 incidence

23
24 **Summary:** Here we present a range of estimates quantifying the extent of underascertainment of
25 COVID-19 cumulative incidence in Africa. These estimates, constructed from serology and
26 mortality data, suggest that systematic underdetection and underreporting may be contributing to
27 the seemingly low burden of COVID-19 reported in Africa.

30 **Abstract**

31 **Background**

32 The COVID-19 pandemic has had a devastating impact on global health, the magnitude of which
33 appears to differ intercontinentally: for example, reports suggest 271,900 per million people have
34 been infected in Europe versus 8,800 per million people in Africa. While Africa is the second
35 largest continent by population, its reported COVID-19 cases comprise <3% of global cases.
36 Although social, environmental, and environmental explanations have been proposed to clarify
37 this discrepancy, systematic infection underascertainment may be equally responsible.

38

39 **Methods**

40 We seek to quantify magnitude of underascertainment in COVID-19's cumulative incidence in
41 Africa. Using serosurveillance and postmortem surveillance, we constructed multiplicative
42 factors estimating ratios of true infections to reported cases in African nations since March 2020.

43

44 **Results**

45 Multiplicative factors derived from serology data – in a subset of 12 nations – suggested a range
46 of COVID-19 reporting rates, from 1 in 630 infections reported in Kenya (May 2020) to 1 in 15
47 infections reported in South Africa (November 2021). The largest multiplicative factor, 3,795,
48 corresponded to Malawi (June 2020), suggesting <0.05% of infections captured. A similar set of
49 multiplicative factors for all nations derived from postmortem data points toward the same
50 conclusion: reported COVID-19 cases are unrepresentative of true infections, suggesting a key
51 reason for low case burden in many African nations is significant underdetection and
52 underreporting.

53

54 **Conclusions**

55 While estimating COVID-19's exact burden is challenging, the multiplicative factors we present
56 provide incidence curves reflecting likely-to-worst-case ranges of infection. Our results stress the
57 need for expansive surveillance to allocate resources in areas experiencing severe discrepancies
58 between reported cases, projected infections, and deaths.

59

60 **Introduction**

61 In just over 2 years, coronavirus disease 19 (COVID-19) has led to a reported 545 million cases
62 and 6.3 million deaths globally [1]. Asia, North America, South America, and Europe have
63 respectively contributed 156 million, 103 million, 59 million, and 204 million cases to this toll
64 (roughly 33,400, 172,100, 136,400, and 271,900 cases per million population), as well as over
65 1.3 million deaths each, whereas reported COVID-19 morbidity and mortality have appeared
66 considerably lower in Africa: 12 million cases (8,800 cases per million population) and 254,900
67 deaths [1, 2].

68 This observation has led some to conclude that the pandemic has “spared” Africa and
69 parts of Asia [3]. Although several explanations have been proposed – including strong
70 government response, young age structure, sparse population density, competing comorbidities,
71 and climate effects – the World Health Organization (WHO) has instead posited that systematic
72 underascertainment explains the continental difference in reporting [3-6]. Leveraging the
73 COVID-19 prevalence calculator developed by the global health organization Resolve to Save
74 Lives, the WHO has further estimated that only 1 in 7 COVID-19 cases in Africa have been
75 detected, compared to 1 in 4 COVID-19 cases in the United States [4, 7, 8]. The WHO's

76 hypothesis has been bolstered by several recent national and sub-national studies [3, 9].
77 However, the variability in underascertainment reported by these studies motivates the need for
78 practicable methods that can triangulate between disparate data sources to provide credible
79 estimates of African nation-specific COVID-19 infections. While the exact number of infections
80 can never be known, producing a range of estimates reflective of likely- and worst-case scenarios
81 is still valuable in understanding the burden of disease. Here we propose 2 such methods that
82 approximate the magnitude of COVID-19 cumulative incidence in Africa, the first relying on
83 serosurveillance data and the second relying on postmortem surveillance data.

84

85 **Methods**

86 **Motivation for Serosurveillance**

87 National population-based serosurveys furnish cross-sectional estimates of cumulative incidence
88 of infection by SARS-CoV-2 in quantifying the proportion of a population with SARS-CoV-2
89 antibodies (“seroprevalence”). Most relevant for infection are the antibody isotypes IgA, IgG,
90 and IgM, which can be further classified based on their specificity for SARS-CoV-2 proteins
91 spike (S), receptor binding domain (RBD), or nucleocapsid (N) [10, 11]. Antibody specificity
92 can also help differentiate between antibody production due to natural infection versus
93 vaccination (coverage for which is still low in Africa), where the latter is observed to induce
94 production of anti-S and anti-RBD – but not anti-N – antibodies [11, 12, 13].

95

96 **Inferring Infection Underascertainment: Analysis of Serosurveillance Data**

97 Using SeroTracker, a global SARS-CoV-2 seroprevalence database, we identified 29 national-
98 level serosurveys collected over the course of the pandemic across 12 African nations: Cape

99 Verde, Côte d’Ivoire, Egypt, Ethiopia, Gabon, Ghana, Kenya, Malawi, Senegal, Sierra Leone,
100 South Africa, and Zambia [14].

101 Table 1 shows the sample size and seroprevalence for each such serosurvey. SeroTracker
102 not only furnishes a point estimate and a 95% confidence interval for seroprevalence at a given
103 time point during the pandemic but also provides metadata for each serosurvey. This metadata
104 includes sampling characteristics – method, location, timeframe, corresponding publication,
105 Joanna Briggs Institute-based bias risk, population group, and population age – and serosurvey
106 characteristics – manufacturer, test type, detected isotype, sensitivity, and specificity [15].

107 For each of the 12 African nations, seroprevalences were compared to the proportion of
108 that nation’s population reported to have COVID-19. The latter proportion was obtained by
109 dividing reported cumulative cases (Our World in Data) at the end of the serosurvey sampling
110 period by the corresponding nation’s population in 2020 (World Bank) or 2021 (United Nations)
111 [16, 17]. Hereafter, we refer to the ratio of seroprevalence to this proportion as a “multiplicative
112 factor.” This multiplicative factor-based framework extends the work of Angulo et al. (2021),
113 who previously formulated a procedure to estimate COVID-19 infections (both asymptomatic
114 and symptomatic), hospitalizations, and deaths in the United States using serosurveys [9].

115 A multiplicative factor of 1 would suggest complete concordance in population positivity
116 for COVID-19 between serology reports and case reports – equivalently, no apparent
117 underascertainment. In contrast, a multiplicative factor greater than 1 would suggest that
118 serology-derived estimates of infection incidence exceed reported cases, reflective of a “hidden
119 morbidity” – of apparent underascertainment. Multiplicative factors are generally expected to be
120 greater than 1 due to the large number of asymptomatic COVID-19 infections that go undetected
121 in nearly all settings, the insensitivity of tests, the failure to report positive tests, and/or the

122 absence of testing among certain populations. Applying multiplicative factors to nation-specific
123 epidemic curves produces scaled curves that estimate infection incidence while maintaining
124 local-in-time trends (Figure 1). To limit overextrapolation, these scaled curves were only
125 projected to the nearest half-year. Although there is a 1- to 2-week delay between infection and
126 antibody positivity, lagged reporting of infections counteracts the effect of this delay [18,19].

127

128 **Inferring Infection Underascertainment: Analysis of Postmortem Surveillance Data**

129 To create another set of multiplicative factors, we made use of 2 Zambian postmortem
130 surveillance studies that employed quantitative reverse transcription PCR to detect COVID-19 in
131 decedents: Mwananyanda et al.'s (2021) postmortem surveillance study, as well as its preprinted
132 (not yet peer-reviewed) follow-up from Gill et al. (2022) [20, 21]. Both studies were performed
133 in a sample from Zambia's capital, Lusaka, with the first conducted between June and October
134 2020 following the first wave of infections and the second conducted between January and June
135 2021 following the second wave of infections [22].

136 Prior to these studies, fewer than 10% of deaths attributed to or related to COVID-19 in
137 Lusaka were identified in life with antemortem testing [19]. In contrast, the 2 postmortem studies
138 found that 58 of 364 (15.9%) decedents in the first and 358 of 1,116 (32%) decedents in the
139 second showed PCR positivity, suggesting marked underascertainment of COVID-19 cases and
140 deaths. With much of Africa, including Zambia, identified as "high risk" on the Infectious
141 Disease Vulnerability Index, it is not improbable that these percentages reflect reality in other
142 African nations [23]. Under this assumption, we multiplied 15.9% and 32% by each African
143 nation's estimated all-cause mortality in 2020 and in 2021, respectively, to obtain the expected
144 number of year-end deaths with COVID-19 (Supplementary Figure 1) [16, 17]. Each of these

145 expectations was then divided by infection fatality ratios (IFRs) – defined as a comparison of
146 infections to deaths due to infection – obtained using different parametric methods by Sorensen
147 et al. (2022) and Onovo et al. at multiple time points during the pandemic [24, 25]. The result is
148 the expected number of year-end COVID-19 infections (Figure 2, Supplementary Figure 2),
149 which can be compared to reported cumulative COVID-19 cases in 2020 and in 2021 to produce
150 another set of multiplicative factors that expand the range of credible infection estimates. These
151 multiplicative factors are visualized in Figure 3.

152 Note that when discussing these data, we intentionally use the phrase “deaths with
153 COVID-19” instead of “deaths from COVID-19” because inferring causality is a challenge in
154 postmortem surveillance. Even so, the United States Centers for Disease Control and Prevention
155 have issued guidance that COVID-19 should be assumed the underlying cause of death if SARS-
156 CoV-2 is detected [26].

157

158 **Results**

159 **Serology Multiplicative Factors**

160 The available serosurveys result in multiplicative factors that are illustrative of systematic
161 underdetection and underreporting (Figure 1). Multiplicative factors ranged in magnitude from 2
162 (Cape Verde, based on data from July 2020) to 3,795 (Malawi, based on June 2020 data),
163 implying that 1 in 2 and 1 in 3,795 infections were respectively detected in these nations early in
164 the pandemic progression. More than half of these multiplicative factors were at the “hundreds”
165 order of magnitude, which facilitates discernment of 4 distinct COVID-19 “waves” previously
166 obscured by the low amplitude peaks of confirmed cases (Figure 1) [22].

167 Additionally, Kenya, for which the largest number of national-level serosurveys were
168 available (6, spanning April 2020 to September 2021), produced multiplicative factors ranging
169 from 110 in December 2020 to 630 in May 2020. This observation of declining multiplicative
170 factors as the pandemic progressed is echoed in data for Malawi, Ghana, Ethiopia, and South
171 Africa, which may reflect incremental improvements in testing capacity.

172

173 **Inter-Method Comparison: Estimates of Infection**

174 Figure 2 provides an inter-method comparison of estimated infections (per 100,000 population).
175 For many nations, the upper bound estimate of COVID-19 infection burden was produced by the
176 estimator shown in gold, which couples postmortem surveillance data from either Mwananyanda
177 et al. (2021) or Gill et al. (2022) with IFR data from Onovo et al. (2021). The large magnitude of
178 the gold estimates is less likely a byproduct of the timing of the Onovo et al. (2021) IFR
179 estimates – May 2020 – and more likely a byproduct of differences in IFR estimation
180 methodologies, as the [even earlier] April 2020 IFR estimates from Sorensen et al. (2022)
181 generated infection estimates that are generally lower by several orders of magnitude.

182 The lower bound estimate of COVID-19 infection burden was produced by the estimator
183 shown in blue, which leverages serosurveillance data. Additionally, infection estimates for most
184 nations cluster around 100,000 infections per 100,000 population. This signals that the majority
185 of individuals in most populations may have already experienced infection at least once.

186

187 **Inter-Method Comparison: Multiplicative Factors**

188 All African nations yielded multiplicative factors exceeding 1 for the serosurveillance method
189 and the postmortem surveillance method (Figure 3), which is indicative of divergence between
190 “ground truth” infections and reported cases.

191 Also notable in Figure 3 is concordance among the multiplicative factors calculated using
192 IFRs from Sorensen et al. (2022), as observed in the clustering of red-shaded markers. Of these,
193 the purple marker, which reflects IFRs from April 2020, corresponded to the smallest magnitude
194 multiplicative factor, whereas the maroon marker, which reflects IFRs from January 2021,
195 corresponded to the largest magnitude multiplicative factor.

196 Multiple North African nations (Tunisia, Morocco, Libya), island nations (Cape Verde,
197 Seychelles), Botswana, and South Africa yielded the smallest magnitude multiplicative factors
198 (<20). The similarly small magnitude across these nations is notable in that South Africa’s
199 established indicator- and event-based disease surveillance system and Cape Verde’s and
200 Seychelles’ geographic isolation contrasts with weakened public health infrastructure in the
201 Middle East and North Africa (MENA) region, as many MENA nations have grappled with a
202 paucity of resources to adequately chronicle COVID-19 [27].

203 Finally, the change in x-axis scale from subplot (a) to subplot (b) also reveals an order-of-
204 magnitude decrease in multiplicative factors from 2020 to 2021, perhaps coincident with
205 improved COVID-19 testing infrastructure and increased implementation of pharmaceutical and
206 non-pharmaceutical interventions over the pandemic’s progression.

207

208 **Assessment of Vulnerabilities**

209 One explanation for underreporting and underdetection is tied to health, economic, and
210 social vulnerabilities. Lewis et al. (2022) describe analysis of these vulnerabilities as an unmet

211 need [6]. These vulnerabilities include food insecurity, poverty, and noncommunicable disease
212 comorbidity, among others [28].

213 One tool to assess these vulnerabilities is Surgo Ventures' Africa COVID-19 Community
214 Vulnerability Index (CCVI) [29]. Across and within 756 regions in 48 African nations, CCVI
215 specifically encodes 7 dimensions of vulnerability, which can together or individually contribute
216 to heterogeneous attenuation of disease signal: age, epidemiology, fragility, health system,
217 population density, socioeconomics, and transport and housing availability. These dimensions
218 are quantified on a Likert-like scale as “very low”, “low”, “moderate”, “high”, or “very high”.
219 We limit our focus to epidemiological factors (non-communicable diseases, non-COVID-19
220 infectious diseases) and health system factors (strength, capacity, and access).

221 In Supplementary Figure 3, we ranked African nations by variance in their respective
222 multiplicative factors and compared this to nation-specific health system, epidemiological, and
223 overall vulnerabilities (as quantified by CCVI). By way of example, South Africa exhibited not
224 only low variance in its multiplicative factor estimates but also low “health system”
225 vulnerability, the combination of which seems to reflect more robust surveillance infrastructure.

226 Calculating the Spearman Rank correlation between the ascending variances for 2020 and
227 the correspondingly ordered “health system” vulnerability index values yields a coefficient of
228 0.44 and a corresponding p-value of 0.001. At a prespecified α -level of 0.05, we reject the null
229 hypothesis that health system vulnerability and variability in infection estimates are uncorrelated
230 and conclude that there exists some evidence to suggest a correlation. While this comparison
231 does not establish a definitive relationship, on average, nations with lower inter-method
232 variances have a lower degree of vulnerability compared to nations with higher inter-method
233 variances.

234

235 **Discussion**

236 Although underascertainment of COVID-19 remains a global challenge, the degree to which
237 COVID-19 cumulative incidence is underdetected and underreported in Africa appears higher
238 than in other continents. By comparison, for example, multiplicative factors were previously
239 estimated to be between roughly 5 and 50 in North America [30, 31]. Both the serosurveillance-
240 and postmortem surveillance-based estimates we present here suggest that COVID-19 infections
241 are tens- to hundreds-fold greater than their reported counterparts.

242 Our process is complementary to a preprinted (not yet peer-reviewed) meta-analysis from
243 Lewis et al. (2022), which also seeks to quantify the divergence between seroprevalence-derived
244 infections and reported cases in Africa [6]. However, the meta-analysis relied predominantly on
245 local-level serosurveys for inference, where local-level serosurveys are primarily obtained in
246 urban settings. The national-level serosurveys we leverage are intended to increase
247 generalizability of results (i.e., without limiting inference to cities). Coarser resolution – as well
248 as not restricting serosurveys under study to those that strictly align with SEROPREV protocol
249 criteria – also means that our analysis includes areas un- or underrepresented in Lewis et al.’s
250 work, such as Egypt and Cape Verde [32].

251 We additionally do not limit the scope of our analysis to serosurveillance data, validating
252 our results with postmortem surveillance data. Postmortem surveillance data has the added
253 advantage of facilitating direct estimation of COVID-19-attributable deaths, where deaths have
254 also exhibited underascertainment globally [33]. As Mwananyanda et al. (2021) and Gill et al.
255 (2022) show, in Lusaka, Zambia, fewer than 10% of those who died with or from COVID-19
256 were tested for COVID-19 prior to death [19]. Supplementary Figure 1 in part captures this

257 underreporting and underdetection of deaths, comparing reported COVID-19-attributable deaths
258 in 2020 and 2021 to Zambia- and excess mortality-based deaths estimates [34]. Although several
259 nations exhibit negative excess mortalities (e.g., Seychelles and Mauritius), which would seem to
260 suggest that the COVID-19 pandemic led to a reduction in deaths, inflection points in the curves
261 at the time of first recorded COVID-19 deaths illustrates that there was instead an existing
262 negative trend in all-cause deaths that COVID-19 reversed [35].

263 The results from our multiplicative framework are concordant with other methods. For
264 example, a stochastic [multi-strain] compartmental national epidemic model for South Africa in
265 a preprinted (not yet peer-reviewed) analysis from Gozzi et al. (2022) suggests that South
266 Africa’s surveillance system could detect 1 in 16 infections between May 2021 and November
267 2021, compared to our estimates of multiplicative factors between 10 and 20 for the same
268 timeframe [36]. Additionally, using a global metapopulation epidemic model, Davis et al. (2021)
269 recovered infection attack rates around 1% for several African nations early in the pandemic’s
270 progression, further underscoring the discrepancy between recorded, reported cases and true
271 infections [31].

272 A limitation of our approach is that the magnitude of multiplicative factors derived from
273 the serosurveillance method may be associated with serosurvey-specific risk of bias.
274 Supplementary Figure 4 explores this association by stratifying the roughly 30 serosurveys by
275 their bias risk, as measured by the Joanna Briggs Institute Checklist for Prevalence Studies (JBI).
276 JBI is a validated critical appraisal tool that encompasses 9 dimensions of assessment, which
277 SeroTracker uses to classify each serosurvey [14, 15]. Ultimately, we found that although the
278 “high bias” stratum corresponded to the highest median multiplicative factor (161) and smallest
279 dispersion compared to the “low bias” stratum and “medium bias” stratum (128 and 110), the

280 overlapping nature of the box plots suggests that this difference is qualitative at best and,
281 therefore, that multiplicative factors of larger magnitude may not be any more biased than
282 multiplicative factors of smaller magnitudes. Furthermore, the propensity for bias does not
283 translate to the direction of bias: simply because a serosurvey is classified as biased does not
284 indicate that resulting seroprevalence is lower or higher than true prevalence. Therefore, we
285 cannot establish that higher multiplicative factors are indeed a result of higher bias.

286 A different source of bias in our serosurveillance method relates to origins of
287 seropositivity, specifically differentiating between seropositivity from vaccination or from
288 infection. However, this is less likely to be of concern given the low vaccination rate of 10% in
289 Africa for the 2020 – 2021 period of analysis [13]. Even so, about one third of tests chronicled
290 by SeroTracker measure IgG (about half are anti-S), one third measure both IgM and IgG, and
291 one third measure total antibody [14]. Additionally, we do not explicitly account for waning
292 immunity and resulting reinfection, which may have contributed to certain nations exhibiting
293 estimates of infection that exceed the total population (i.e., any points in Figure 2 and
294 Supplementary Figure 2 exceeding 100,000 cases per 100,000 population). Relatedly, we cannot
295 rule out underestimation of “ground truth” disease burden due to seroreversion, the phenomenon
296 in which antibody levels decline over time to a level below the cutoff for seropositivity, resulting
297 in an underestimation of cumulative incidence [37]. Seroreversion and reinfection would each
298 lead to underestimates of cumulative incidence, and thus underestimates of the multiplicative
299 factor between reported cases and true infections. This further underscores the narrative that
300 reported cases do not adequately capture the burden of infection.

301 Another important feature of serosurveys is the populations from which they sample.
302 These populations often include healthcare workers, blood donors, and the elderly, who may be

303 differentially exposed to SARS-CoV-2. Our analysis does not consider these characteristics of
304 the sample frame and, therefore, may overestimate cumulative incidence of COVID-19 if the
305 sampled individuals indeed have higher cumulative incidence than the underlying population. In
306 Malawi, for example, the thousands order of magnitude for the June 2020 multiplicative factor
307 could be in part attributable to an especially unrepresentative sample of healthcare workers and
308 caregivers. Given the scarcity of reported infections stratified on serum donor profiles, one
309 possibility for mitigating this bias would entail the extrapolation of seroprevalence to the general
310 population using weights that account for demographic differences. The United States Centers
311 for Disease Control and Prevention recently published such a reweighting protocol specific to
312 blood donors [38]. Another possibility would be to assess intra-nation variability of serology
313 estimates, as nations like Cameroon and South Africa have started releasing local-level
314 serosurveys to the public, and to employ these estimates in a sensitivity analysis for
315 multiplicative factor heterogeneity [39].

316 An additional consideration for contextualizing our results is that our methods combine
317 multiple timeframes to create time-invariant multiplicative factors. For example, the
318 serosurveillance studies from SeroTracker cover multiple dates between June 2020 and
319 November 2021, whereas postmortem surveillance studies reflect either the time period of June
320 2020 to October 2020 or January 2021 to July 2021. It follows that multiplicative factors
321 estimated from early pandemic stages may not be as representative of disease dynamics seen in
322 later pandemic stages from SARS-CoV-2 strain-specific effects (e.g., Delta and Omicron).
323 However, the cross-sectional nature of serosurveys suggests that cumulative estimates of
324 infection, such as those presented in our analysis, may be more reliable than conversions to
325 weekly estimates of infection.

326 Finally, in calculating the estimated number of infections using serosurveillance and
327 postmortem surveillance data, we employed 2020 and 2021 national population size estimates
328 from the World Bank and the United Nations [16, 17]. However, these estimates were the result
329 of projections prior to the COVID-19 pandemic, which would have the effect of overestimating
330 population size due to unaccounted for deaths. To quantify the impact of this compounding
331 uncertainty, we performed a sensitivity analysis for 2021 infections and deaths by subtracting the
332 number of COVID-19-attributable deaths in 2020 from the projected population size in 2021.
333 The percent difference between unadjusted infections and adjusted infections was less than 0.5%
334 for all African nations, a result summarized in Supplementary Table 1, which suggests that the
335 projected population sizes without adjustment are sufficient for analysis.

336 Asymptomatic infections and variable rates of testing for COVID-19 in Africa, as well as
337 globally, have made it particularly challenging to ascertain the true burden of the pandemic from
338 data, as reflected in the range of multiplicative factors from our analysis (1 in 2 cases reported
339 for Cape Verde, 1 in 3,795 cases reported in Malawi) [40, 41]. While previous studies have
340 sought to estimate this systematic underascertainment in resource-limited settings, they have
341 relied on a single data stream to make inferences. The innovation of our approach lies in
342 leveraging not only serosurveillance data (antibody testing) but also postmortem surveillance
343 data (PCR testing) to characterize COVID-19 in Africa. Concordance between these methods,
344 when taken with other sources of epidemiological data, such as syndromic surveillance and
345 digital surveillance [19, 42] can be leveraged in 2 ways: (1) to help inform allocation of
346 resources to areas exhibiting noticeable divergence in reported infections and recalibrated
347 infections; and (2) to help inform initiatives that can bolster surveillance infrastructure.

348

349 **Funding**

350 This work was supported by institutional research funds from Pfizer Inc. (NEK, SG, ML, and
351 MS). The content is solely the responsibility of the authors and does not necessarily represent the
352 official views of Pfizer Inc. NEK was also supported by the National Institute of Allergy and
353 Infectious Diseases of the National Institutes of Health under award number 2T32AI007535.

354

355 **Conflict of Interest**

356 The authors declare that they have no conflicts of interest.

357

358 **Acknowledgements**

359 We thank Bethany Hedt-Gauthier (Harvard T.H. Chan School of Public Health) for her thorough
360 review of the manuscript and for her facilitation of international collaboration in this work. We
361 also thank Amit Srivastava (Pfizer Inc.) for his invaluable industry perspective.

362

363 **References**

- 364 1. Center for Systems Science and Engineering at Johns Hopkins University. Coronavirus
365 Resource Center. Available at: <https://coronavirus.jhu.edu/map.html>. Accessed 27 June 2022.
- 366 2. Our World in Data. Coronavirus (COVID-19). Available at: <https://ourworldindata.org/>.
367 Accessed 27 June 2022.

- 368 3. Rahmandad H, Lim TY, Sterman J. Behavioral dynamics of COVID-19: Estimating
369 underreporting, multiple waves, and adherence fatigue across 92 nations. *Syst Dyn Rev*,
370 **2021**;37(1):5-31. doi:10.1002/sdr.1673
- 371 4. WHO Africa. Six in seven COVID-19 infections go undetected in Africa. Available at:
372 <https://www.afro.who.int/news/six-seven-covid-19-infections-go-undetected-africa>. Accessed 15
373 December 2021.
- 374 5. Wamai RG, Hirsch JL, Van Damme W, et al. What could explain the lower COVID-19
375 burden in Africa despite considerable circulation of the SARS-CoV-2 virus? *Int J Environ Res*
376 *Public Health*, **2021**;18(16):8638. doi:10.3390/ijerph18168638
- 377 6. Lewis HC, Ware H, Whelan M, et al. SARS-CoV-2 infection in Africa: A systematic review
378 and meta-analysis of standardized seroprevalence studies, from January 2020 to December 2021.
379 medRxiv 2022.02.14.22270934 [Preprint]. February 14, 2022 [cited 25 May 2022]. Available
380 from: <https://doi.org/10.1101/2022.02.14.22270934>.
- 381 7. Centers for Disease Control and Prevention. Estimating COVID-19 Burden. Available at:
382 <https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/burden.html>. Accessed 15 January
383 2022.
- 384 8. Resolve to Save Lives. COVID-19 Calculator Used in WHO Analysis. Available at:
385 <https://resolvetosavelives.org/timeline/covid-19-calculator-used-in-who-analysis>. Accessed 15
386 January 2022.
- 387 9. Angulo FJ, Finelli L, Swerdlow DL. Estimation of US SARS-CoV-2 infections, symptomatic
388 infections, hospitalizations, and deaths using seroprevalence surveys. *JAMA Netw Open*,
389 **2021**;4(1):e2033706. doi:10.1001/jamanetworkopen.2020.33706

- 390 10. Centers for Disease Control and Prevention. SARS-CoV-2 infection-induced and vaccine-
391 induced immunity. Available at: [https://www.cdc.gov/coronavirus/2019-ncov/science/science-](https://www.cdc.gov/coronavirus/2019-ncov/science/science-briefs/vaccine-induced-immunity.html)
392 [briefs/vaccine-induced-immunity.html](https://www.cdc.gov/coronavirus/2019-ncov/science/science-briefs/vaccine-induced-immunity.html). Accessed 12 March 2022.
- 393 11. Batra M, Tian R, Zhang C, et al. Role of IgG against N-protein of SARS-CoV2 in COVID19
394 clinical outcomes. *Sci Rep*, **2021**;11(1):3455. doi:10.1038/s41598-021-83108-0
- 395 12. WHO Africa. Africa needs to ramp up COVID-19 vaccination six-fold. Available at:
396 <https://www.afro.who.int/news/africa-needs-ramp-covid-19-vaccination-six-fold>. Accessed 4
397 February 2022.
- 398 13. Africa CDC. COVID-19 Vaccination. Available at: [https://africacdc.org/covid-19-](https://africacdc.org/covid-19-vaccination/)
399 [vaccination/](https://africacdc.org/covid-19-vaccination/). Accessed 30 December 2022.
- 400 14. Arora RK, Joseph A, Van Wyk J, et al. SeroTracker: A global SARS-CoV-2 seroprevalence
401 dashboard. *Lancet Inf Dis*, **2021**;21(4):E75-76. doi: 10.1016/S1473-3099(20)30631-9
- 402 15. Joanna Briggs Institute. Checklist for Prevalence Studies: Systematic Review Appraisal
403 Tools. Available at: [https://jbi.global/sites/default/files/2019-05/JBI_Critical_Appraisal-](https://jbi.global/sites/default/files/2019-05/JBI_Critical_Appraisal-Checklist_for_Prevalence_Studies2017_0.pdf)
404 [Checklist_for_Prevalence_Studies2017_0.pdf](https://jbi.global/sites/default/files/2019-05/JBI_Critical_Appraisal-Checklist_for_Prevalence_Studies2017_0.pdf). Accessed 1 March 2022.
- 405 16. The World Bank. World Population Prospects. Available at:
406 <https://data.worldbank.org/indicator/SP.POP.TOTL>. Accessed 24 April 2022.
- 407 17. United Nations Department of Economic and Social Affairs. World Population Prospects
408 2019. Available at: <https://population.un.org/wpp/Download/Standard/Population/>. Accessed 17
409 June 2022.
- 410 18. Imai K, Tabata S, Ikeda M, et al. Clinical evaluation of an immunochromatographic IgM/IgG
411 antibody assay and chest computed tomography for the diagnosis of COVID-19. *J Clin Virol*,
412 **2020**;128:104393. doi:10.1016/j.jcv.2020.104393

- 413 19. Kogan NE, Clemente L, Liataud P, et al. An early warning approach to monitor COVID-19
414 activity with multiple digital traces in near real time. *Sci Adv*, **2021**;7(10):eabd6989. Published
415 2021 Mar 5. doi:10.1126/sciadv.abd6989
- 416 20. Mwananyanda L, Gill CJ, MacLeod W, et al. Covid-19 deaths in Africa: Prospective
417 systematic postmortem surveillance study. *BMJ*, **2021**;372:n334. doi:10.1136/bmj.n334
- 418 21. Gill CJ, Mwananyanda L, MacLwod W, et al. Sustained high prevalence of COVID-19
419 deaths from a systematic post-mortem study in Lusaka, Zambia: One year later. medRxiv
420 2022.03.08.22272087 [Preprint]. March 8, 2022 [cited 25 May 2022]. Available from:
421 <https://doi.org/10.1101/2022.03.08.22272087>.
- 422 22. Salyer SJ, Maeda J, Sembuche S, et al. The first and second waves of the COVID-19
423 pandemic in Africa: A cross-sectional study. *Lancet*, **2021**;397(10281):1265-1275.
424 doi:10.1016/S0140-6736(21)00632-2
- 425 23. Moore, M, Gelfeld B, Okunogbe AT, Paul C. Identifying Future Disease Hot Spots:
426 Infectious Disease Vulnerability Index. Available at:
427 https://www.rand.org/pubs/research_reports/RR1605.html. Accessed 12 March 2022.
- 428 24. COVID-19 Forecasting Team (led by Sorensen RJ). Variation in the COVID-19 infection-
429 fatality ratio by age, time, and geography during the pre-vaccine era: A systematic analysis.
430 *Lancet*. **2022**;399(10334):1469-1488. doi:10.1016/S0140-6736(21)02867-1
- 431 25. Onovo AA, Kalaiwo A, Obanubi C, et al. Estimates of the COVID-19 infection fatality rate
432 for 48 African countries: A model-based analysis. *BioMed*, **2021**;1(1);63-79.
433 doi:10.3390/biomed1010005

- 434 26. Centers for Disease Control and Prevention. Guidance for Certifying Deaths due to
435 Coronavirus Disease 2019 (COVID-19). Available at:
436 <https://www.cdc.gov/nchs/data/nvss/vsrg/vsrg03-508.pdf>. Accessed 15 March 2022.
- 437 27. Wehbe S, Fahme SA, Rizk A, et al. COVID-19 in the Middle East and North Africa region:
438 An urgent call for reliable, disaggregated and openly shared data. *BMJ Global Health*,
439 **2021**;6:e005175. doi:10.1136/bmjgh-2021-005175
- 440 28. Ulimwengu JM, Domgho, LM, Collins J. Assessing the vulnerability of West and Central
441 African countries to COVID-19. Available at:
442 <https://ebrary.ifpri.org/digital/collection/p15738coll2/id/134755>. Accessed 24 April 2022.
- 443 29. Surgo Ventures. Precision for COVID: Africa CCVI. Available at:
444 <https://precisionforcovid.org/africa>. Accessed 15 December 2021.
- 445 30. Lu FS, Nguyen AT, Link NB, et al. Estimating the cumulative incidence of COVID-19 in the
446 United States using influenza surveillance, virologic testing, and mortality data: Four
447 complementary approaches. *PLoS Comput Biol*, **2021**;17(6):e1008994.
448 doi:10.1371/journal.pcbi.1008994
- 449 31. Davis JT, Chinazzi M, Perra N, et al. Cryptic transmission of SARS-CoV-2 and the first
450 COVID-19 wave. *Nature*, **2021**;600:127-132. doi:10.1038/s41586-021-04130-w
- 451 32. WHO IRIS. Population-Based Age-Stratified Seroepidemiological Investigation Protocol for
452 COVID-19 Virus Infection. Available at: <https://apps.who.int/iris/handle/10665/331656>.
453 Accessed 24 April 2022.
- 454 33. COVID-19 Excess Mortality Collaborators. Estimating excess mortality due to the COVID-
455 19 pandemic: A systematic analysis of COVID-19-related mortality. *Lancet*.
456 **2022**;399(10334):1513-1536. doi:10.1016/S0140-6736(21)02796-3

- 457 34. Our World in Data. Excess mortality during the Coronavirus pandemic (COVID-19).
458 Available at: <https://ourworldindata.org/excess-mortality-covid>. Accessed 27 June 2022.
- 459 35. Our World in Data. Estimated cumulative excess deaths during COVID (from the
460 Economist), Seychelles. Available at: [https://ourworldindata.org/grapher/excess-deaths-](https://ourworldindata.org/grapher/excess-deaths-cumulative-economist-single-entity?country=~SYC)
461 [cumulative-economist-single-entity?country=~SYC](https://ourworldindata.org/grapher/excess-deaths-cumulative-economist-single-entity?country=~SYC). Accessed 27 June 2022.
- 462 36. Gozzi N, Chinazzi M, Davis JT, et al. Preliminary modeling estimates of the relative
463 transmissibility and immune escape of the Omicron SARS-CoV-2 variant of concern in South
464 Africa. medRxiv 2022.01.04.22268721 [Preprint]. January 05, 2022 [cited 27 June 2022].
465 Available from: <https://doi.org/10.1101/2022.01.04.22268721>
- 466 37. COVID-19 Community Research Partnership Study Group. Duration of SARS-CoV-2 sero-
467 positivity in a large longitudinal sero-surveillance cohort: The COVID-19 Community Research
468 Partnership. BMC Infect Dis, **2021**;21(1):889. doi:10.1186/s12879-021-06517-6
- 469 38. Stone M, Di Germanio C, Wright DJ, et al. Use of US blood donors for national
470 serosurveillance of Severe Acute Respiratory Syndrome Coronavirus 2 antibodies: Basis for an
471 expanded national donor serosurveillance program. Clin Infect Dis, **2022**;74(5):871-881.
472 doi:10.1093/cid/ciab537
- 473 39. Nwosu K, Fokam J, Wanda F, et al. SARS-CoV-2 antibody seroprevalence and associated
474 risk factors in an urban district in Cameroon. Nat Commun, **2021**;12(1):5851.
475 doi:10.1038/s41467-021-25946-0
- 476 40. Chitungo I, Dzobo M, Hlongwa M, Dzinamarira T. COVID-19: Unpacking the low number
477 of cases in Africa. Public Health Pract (Oxf), **2020**;1:100038. doi:10.1016/j.puhip.2020.100038

478 41. Chookajorn T, Kochakarn T, Wilasang C, Kotanan N, Modchang C. Southeast Asia is an
479 emerging hotspot for COVID-19. *Nat Med*, **2021**;27(9):1495-1496. doi:10.1038/s41591-021-
480 01471-x

481 42. Fulcher IR, Boley EJ, Gopaluni A, et al. Syndromic surveillance using monthly aggregate
482 health systems information data: Methods with application to COVID-19 in Liberia. *Int J*
483 *Epidemiol*, **2021**;50(4):1091-1102. doi:10.1093/ije/dyab094

484

485

486

487

488

489

490

491

492

493

494

495

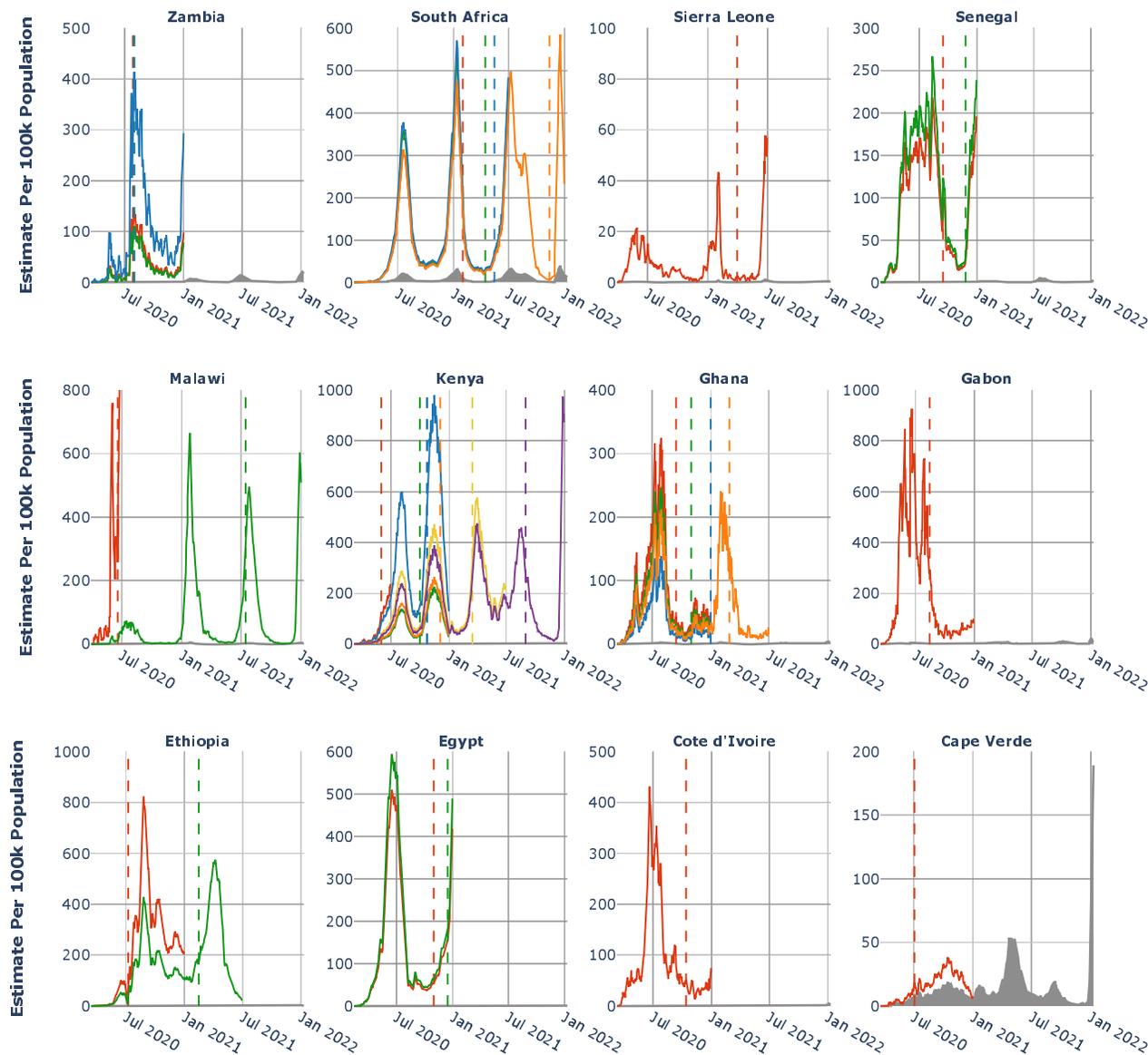
496 Figures

497 **Table 1: Summary of Serosurveillance Study Metadata from SeroTracker***

Nation	# Tested	Seroprevalence (95% CI)	Sampling Period (MM.DD.YY)	Population	Ages
Cape Verde	5,348	0.4%	06.26.20-07.04.20	Household and community	Multiple groups
Côte d'Ivoire	1,687	25.1%	07.15.20-10.13.20	Essential non-HCW	18-64 years
Egypt	1,260 82	34.8% 46.3%	10.01.20-10.31.20 10.15.20-12.15.20	Household and community HCW and caregivers	Multiple groups 18-64 years
Ethiopia	16,932 1,997	3.5% (3.2%-3.8%) 39.8% (37.4%-41.7%)	06.24.20-07.08.20 12.15.20-02.15.21	Household and community HCW and caregivers	Multiple groups 18-64 years
Gabon	319	55.0%	03.15.20-08.15.20	COVID-19 patient contacts	Multiple groups
Ghana	1,305 395 393 636	19.0% 15.2% 9.7% 22.8%	07.27.20-09.14.20 10.01.20-10.31.20 12.01.20-12.31.20 02.01.21-02.28.21	Multiple general populations	Multiple groups
Kenya	830	42.3% (38.4%-46.3%)	09.30.20-10.23.20	Essential non-HCW	18-64 years
	684	20.8% (17.5%-24.4%)	07.30.20-12.04.20	HCW and caregivers	18-64 years
	9,922	7.9% (6.7%-9.0%)	04.30.20-09.30.20	Blood donors	18-64 years
	3,018	48.5% (45.2%-52.1%)	01.03.20-03.15.21	Blood donors	Multiple groups
	176	83.0%	08.01.21-08.31.21	Blood donors	18-64 years
	176	2.3%	05.01.20-05.31.20	Blood donors	18-64 years
Malawi	500	12.3% (8.2%-16.5%)	05.22.20-06.19.20	HCW and caregivers	18-64 years
	5,085	27.6%	01.15.20-07.15.21	Blood donors	Multiple groups
Senegal	303	21.2% (16.7%-26.1%)	06.15.20-09.15.20	Non-COVID-19 patients	Multiple groups
	1,422	28.4% (26.1%-30.8%)	10.25.20-11.26.20	Household and community	
Sierra Leone	1,893	2.6% (1.9%-3.4%)	03.01.21-03.31.21	Household and community	Multiple groups
South Africa	927	35.8%	11.23.20-01.31.21	Household and community	Multiple groups
	16,762	47.4% (46.2%-48.6%)	01.15.21-05.15.21	Blood donors	
	7,577	45.2% (43.7%-46.7%)	11.15.20-04.15.21	Multiple general populations	
	3,395	71.1% (68.8%-73.5%)	11.08.21-11.12.21	Blood donors	
Zambia	2,704	2.1% (1.1% - 3.1%)	07.04.20-07.27.20	Household and community	Multiple groups
	575	2.2% (0.5% - 3.9%)	07.02.20-07.31.20	HCW and caregivers	
	1,657	8.2% (5.1% - 11.4%)	07.02.20-07.31.20	Residual sera	

499 * *Abbreviations: HCW, healthcare workers; “multiple groups”, at least a 30% overlap within*
500 *another age group*

501 **Figure 1: Confirmed COVID-19 Cases and Serology-Based Infection Estimates for 12**
 502 **African nations**



Country, Date (Multiplicative Factor)

- Confirmed Cases
- Zambia, 2020-07-27 (85)
- Zambia, 2020-07-31 (68)
- Zambia, 2020-07-31 (253)
- South Africa, 2021-01-31 (15)
- South Africa, 2021-04-15 (17)
- South Africa, 2021-05-15 (18)
- South Africa, 2021-11-12 (15)
- Sierra Leone, 2021-03-31 (53)
- Senegal, 2020-09-15 (244)
- Senegal, 2020-11-26 (298)
- Malawi, 2020-06-19 (3795)
- Malawi, 2021-07-15 (131)
- Kenya, 2020-05-31 (630)
- Kenya, 2020-09-30 (110)
- Kenya, 2020-10-23 (475)
- Kenya, 2020-12-04 (128)
- Kenya, 2021-03-15 (234)
- Kenya, 2021-08-31 (193)
- Ghana, 2020-09-14 (129)
- Ghana, 2020-10-31 (98)
- Ghana, 2020-12-31 (55)
- Ghana, 2021-02-28 (86)
- Ethiopia, 2020-07-08 (594)
- Ethiopia, 2021-02-14 (317)
- Egypt, 2020-10-31 (331)
- Egypt, 2020-12-15 (386)
- Cote d'Ivoire, 2020-10-13 ()
- Cape Verde, 2020-07-04 (2)
- Gabon, 2020-08-15 (149)

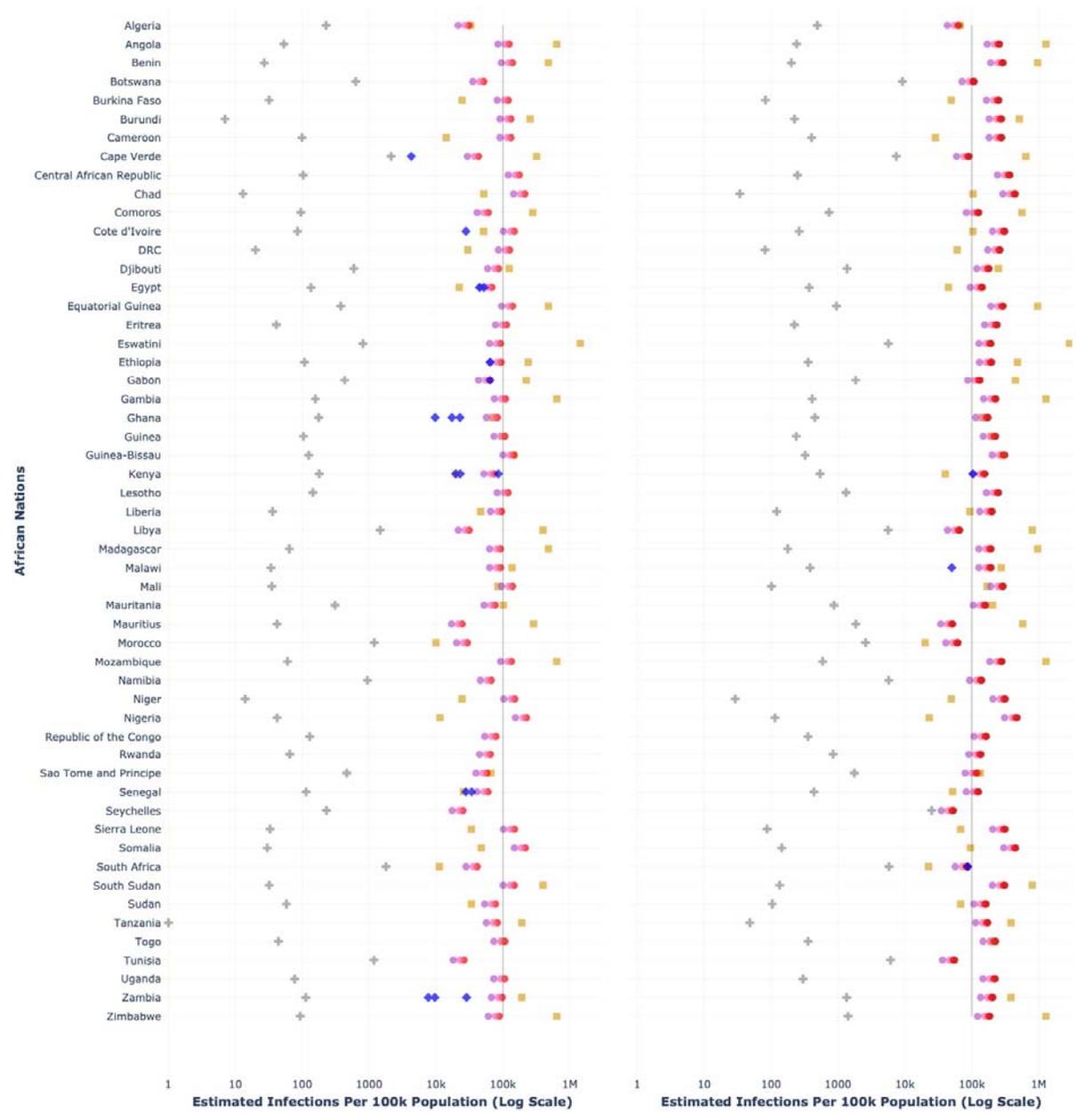
512

513 *Reported COVID-19 cases (gray) versus seroprevalence-derived COVID-19 infections (red,*
 514 *green, blue, orange, yellow, purple) per 100,000 population for 12 African nations. Each non-*

515 *gray curve is created through scaling the corresponding gray curve by a multiplicative factor,*
516 *parenthetically indicated in the legend, with scaling performed to the nearest half-year to reduce*
517 *overextrapolation of results. Color order corresponds to chronological temporal order (red:*
518 *earliest, purple: latest) of serosurvey. Vertical dotted lines represent the last date of each*
519 *serosurvey sampling period.*
520

521 **Figure 2: Estimated COVID-19 Infections Using Serosurveillance and Postmortem**
 522 **Surveillance in Africa**

523
 524 **(a) Estimates Through December 2020 (b) Estimates Through December 2021**



525
 526
 527
 528
 529
 530
 531

Data Source
 + Reported Cases (JHU)
 ■ Deaths (Zambia-based) / May '20 IFR (Onovo et al.)
 ■ Deaths (Zambia-based) / April '20 IFR (Sorensen et al.)
 ■ Deaths (Zambia-based) / July '20 IFR (Sorensen et al.)
 ■ Deaths (Zambia-based) / October '20 IFR (Sorensen et al.)
 ■ Deaths (Zambia-based) / January '21 IFR (Sorensen et al.)
 ◆ Serology (SeroTracker)

532 *Comparison of reported COVID-19 cases, cumulative seroprevalence-derived COVID-19*
 533 *infections (available for 12 nations), and cumulative postmortem-derived COVID-19 infections*

534 *(Onovo et al. available for 44 nations, Sorensen et al. available for 54 nations) per 100,000*
535 *population in Africa. The solid black line in each subplot represents 100,000 infections per*
536 *100,000 population, implying that an entire population has been infected. (a) represents*
537 *infection estimates through December 31, 2020, and it draws from serosurveys taken between*
538 *July 1, 2020 and December 31, 2020 as well as from postmortem surveys detailed in*
539 *Mwananyanda et al. (2021) to derive Zambia-based COVID-19 deaths. (b) represents infection*
540 *estimates through December 31, 2021, and it draws from serosurveys taken between July 1, 2021*
541 *and December 31, 2021 as well as from postmortem surveys detailed in Gill et al. (2022) to*
542 *derive Zambia-based COVID-19 deaths.*
543

544

545

546

547

548

549

550

551

552

553

554

555

556

557

558

559

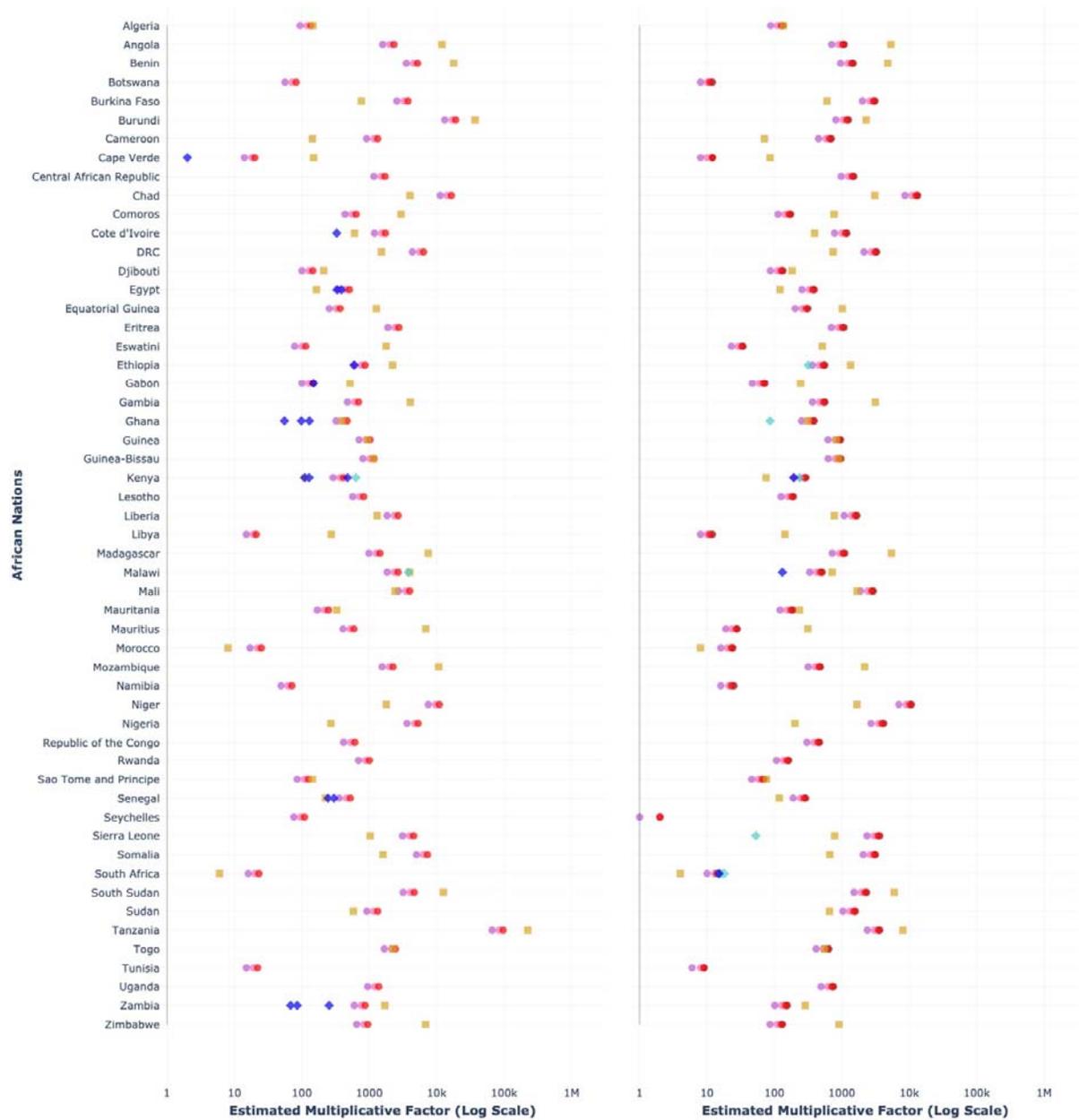
560

561

562 **Figure 3: Multiplicative Factors Needed to Estimate True Infections from Confirmed Cases**
 563 **of COVID-19 in Africa**

564
 565

(a) Estimates Through December 2020 (b) Estimates Through December 2021



566
 567
 568
 569
 570
 571
 572
 573

574 *Comparison of multiplicative factors summarizing ratio of seroprevalence-derived COVID-19*
575 *infections (available for 12 nations) and postmortem-derived COVID-19 infections (Onovo et al.*
576 *available for 44 nations, Sorensen et al. available for 54 nations) to reported cases in Africa. (a)*
577 *represents ratio estimates through December 31, 2020, and it leverages postmortem surveys*
578 *detailed in Mwananyanda et al. (2021) to derive Zambia-based COVID-19 deaths. (b) represents*
579 *ratio estimates through December 31, 2021, and it leverages postmortem surveys detailed in Gill*
580 *et al. (2022) to derive Zambia-based COVID-19 deaths.*
581