

1 **Epidemiological characteristics of three SARS-CoV-2 variants of concern and implications for**
2 **future COVID-19 pandemic outcomes**

3 Wan Yang¹ and Jeffrey Shaman²

4 ¹Department of Epidemiology, ²Department of Environmental Health Sciences, Mailman School
5 of Public Health, Columbia University, New York, NY, USA

6 Correspondence to: wy2202@cumc.columbia.edu (WY); jls106@cumc.columbia.edu (JS)

7
8 **Abstract**

9 Three SARS-CoV-2 variants classified as variants of concern – B.1.1.7, B.1.351, and P.1 – have
10 spread globally. To characterize their viral and epidemiological properties in support of public
11 health planning, we develop and apply a model-inference system to estimate the changes in
12 transmissibility and immune escape for each variant, based on case and mortality data from the
13 country where each variant emerged. Accounting for under-detection of infection, disease
14 seasonality, concurrent non-pharmaceutical interventions, and mass-vaccination, we estimate
15 that B.1.1.7 has a 46.6% (95% CI: 32.3 – 54.6%) increase in transmissibility but nominal immune
16 escape from protection induced by prior wild-type infection; B.1.351 has a 32.4% (95% CI: 14.6
17 – 48.0%) increase in transmissibility and 61.3% (95% CI: 42.6 – 85.8%) immune escape; and P.1
18 has a 43.3% (95% CI: 30.3 – 65.3%) increase in transmissibility and 52.5% (95% CI: 0 – 75.8%)
19 immune escape. Model simulations indicate that B.1.351 and P.1 could supplant B.1.1.7
20 dominance and lead to increased infections. Our findings highlight the importance of
21 preventing the spread of B.1.351 and P.1, in addition to B.1.1.7, via continued preventive
22 measures, prompt mass-vaccination of all populations, continued monitoring of vaccine
23 efficacy, and possible updating of vaccine formulations to ensure high efficacy.

24
25 **Main text**

26 Multiple SARS-CoV-2 variants have been identified since summer 2020. Among these, three
27 variants – namely, B.1.1.7, B.1.351, and P.1 – have been classified as variants of concern (VOCs),
28 per evidence indicating these genotypes are substantially more transmissible, evade prior
29 immunity (either vaccine-induced or conferred by natural infection with wild-type virus),
30 increase disease severity, reduce the effectiveness of treatments or vaccines, or cause
31 diagnostic detection failures.^{1,2} Multiple lines of evidence indicate the B.1.1.7 variant is roughly
32 50% more transmissible than wild-type virus but does not produce antigenic escape.³⁻⁶ Further,
33 several studies have shown that both the B.1.351 and P.1 variants are resistant to
34 neutralization by convalescent plasma from individuals previously infected by wild-type SARS-
35 CoV-2 viruses and sera from vaccinated individuals;⁵⁻⁹ however, changes to the transmissibility
36 of these latter two variants are less well resolved. A better understanding of the
37 transmissibility and immune escape properties of VOCs – in particular, B.1.351 and P.1 at

38 present – is needed to anticipate future COVID-19 pandemic outcomes and support public
39 health planning.

40

41 In this study, we develop a model-inference system to estimate the relative change in
42 transmissibility and level of immune evasion for different SARS-CoV-2 variants, while
43 accounting for under-detection of infection, delays of reporting, disease seasonality, non-
44 pharmaceutical interventions (NPIs), and vaccination. Testing using model-generated synthetic
45 incidence and mortality data indicates this inference system is able to accurately identify shifts
46 in transmissibility and immune evasion. We then apply the validated inference system in
47 conjunction with incidence and mortality data from the UK, South Africa, and Brazil – the three
48 countries where VOCs B.1.1.7, B.1.351, and P.1 were first identified, respectively – to estimate
49 the change of transmissibility and immune evasion for the three VOCs, separately. We further
50 use these inferred findings in a multi-variant, age-structured model to simulate epidemic
51 outcomes in a municipality like New York City (NYC) where multiple variants, including B.1.1.7,
52 B.1.351, and P.1, have been detected.

53

54 **The model-inference system and validation**

55 We first tested our model-inference system using 10 model-generated synthetic datasets,
56 depicting different combinations of population susceptibility prior to the emergence of a new
57 variant, changes in transmissibility and immune evasion for the new variant, and infection-
58 detection rate. As population susceptibility, interventions, and disease seasonality can all affect
59 apparent transmissibility at a given time and in order to focus on variant-specific properties,
60 here we defined transmissibility as the average number of secondary infections per primary
61 infection, *after removing the effects of these three factors* (see Methods). We then quantified
62 the change in transmissibility as its relative increase once the new variant becomes dominant.
63 Similarly, we quantify the level of immune evasion as the increase in susceptibility after the
64 new variant becomes dominant, relative to prior population immunity from wild-type infection.

65

66 Fig. 1 shows example test results comparing model-inference system estimates with model-
67 generated ‘true’ values of transmissibility and susceptibility using an infection-detection rate of
68 20%. Across a range of epidemic dynamics, the model-inference system is able to fit both
69 weekly incidence and mortality data (Fig. 1A) and estimate the transmissibility and
70 susceptibility over time for both the initial pandemic wave and the subsequent pandemic wave
71 caused by a new variant (Fig. 1B). In addition, when aggregated over both pandemic waves, the
72 model-inference system is able to estimate the relative changes in transmissibility and immune
73 evasion due to a new variant (Fig. 1C). When the system is less well constrained (e.g., an
74 infection-detection rate of 10%), model estimates, albeit less accurate, still closely track true
75 values in most instances (Fig S1).

76

77 **Reconstructed pandemic dynamics in the three countries**

78 Following the initial emergence of SARS-CoV-2 in early 2020, the UK, South Africa, and Brazil
79 experienced very different epidemics (Fig. 2). The model-inference system is able to recreate
80 the observed incidence and mortality epidemic curves for all three countries (Fig 2 A, C, and E).
81 Further, cross-validation using independent data shows that the model estimates closely match
82 measured SARS-CoV-2 prevalence in the UK^{10,11} and serology measures of cumulative infection
83 rates in South Africa^{12,13} and Brazil,¹⁴ respectively (Fig 2 B, D, and F). These results indicate the
84 model-inference system accurately estimates the underlying transmission dynamics for all
85 three countries.

86

87 In the UK, a prompt lockdown allowed the country to contain the first pandemic wave (Fig. 3).
88 The real-time effective reproduction number (R_t), which measures the average number of
89 secondary infections at a given point in time, dropped from 2.2 (95% CrI: 1.0 – 3.9) during the
90 week of 3/1/2020 to below 1 during the week of 3/22/2020, the first week of the lockdown (Fig
91 3A). By the week of 6/28/2020 (the week with the lowest incidence following the first
92 pandemic wave), 6.4% (95% CrI: 3.6 – 12.3%) of the UK population are estimated to have been
93 infected. However, with the relaxation of intervention measures during the summer,
94 transmission gradually increased again (as indicated by the estimated $R_t > 1$), leading to a large
95 surge of infections in the autumn of 2020 (Figs 2A-B and 3A). A second lockdown implemented
96 in Nov 2020 reduced transmission transiently (R_t was below 1 during the 4-week lockdown
97 period; Fig 3A). Shortly thereafter widespread transmission of the B.1.1.7 variant led to a
98 further increase of cases before this activity was curtailed by a third lockdown and mass-
99 vaccination.

100

101 In South Africa, initial transmission was low likely due to a strong public health response (a
102 lockdown was implemented from 3/26/20 to 4/16/20) and less conducive conditions for
103 transmission during southern hemisphere summer and autumn (Figs 2C, 3D, and S2B).
104 However, as the country relaxed intervention measures and entered the winter, transmission
105 increased substantially from May 2020 onwards, leading to a large pandemic wave during May
106 – Sep 2020. After accounting for under-detection of infection (Fig S3C), consistent with serology
107 data,^{12,13} the model-inference system estimates that 30.0% (95% CrI: 18.0 – 47.1%) of the
108 population had been infected by the week of 9/20/2020 (i.e., the week with the lowest
109 incidence following the first pandemic wave; Fig 2D). After two months with relatively low
110 incidence, the emergence of the B.1.351 variant led to a resurgence of infections in late 2020
111 and a larger second pandemic wave (Fig 3D). By the week of 4/11/21, another 38.6% (95% CrI:
112 24.1 – 61.1%) of the population are estimated to have been infected, including re-infections.

113

114 In Brazil, no national lockdown was implemented during the pandemic. A large first pandemic
115 wave occurred during Mar – Oct 2020 (Figs 2E and 3G). By the week of 11/1/2020 (i.e., the
116 week with the lowest incidence following the first wave), 45.7% (95% CrI: 28.4 – 69.0%) of the
117 population are estimated to have been infected (Fig 3F). This estimate includes all infections
118 and thus is much higher than the reported number of cases (3.77% of the population; see
119 estimated infection-detection rates in Fig S3E). In addition, unlike the UK and South Africa
120 where the pandemic wave rose and fell quickly, pandemic activity – based on national
121 incidence and mortality curves – remained at high levels for a much longer duration (Fig. 2E).
122 This may be due to the larger geographical area of Brazil, the aggregative nature of country-
123 level incidence and mortality data combining multiple outbreak waves from different sub-
124 regions of the country, and the lack of national restrictions to curb the pandemic. Despite this
125 large first pandemic wave, the emergence of the P.1 variant led to a second large pandemic
126 wave from Dec 2020 onwards. Similar traveling waves through the country were evident from
127 the incidence curve (Fig 2E). By the week of 4/11/21, an additional 60.7% (95% CrI: 40.5 –
128 92.0%) of the population are estimated to have been infected, including re-infections.

129

130 **Estimated increased transmissibility and immune evasion**

131 Accounting for concurrent NPIs, vaccination and seasonal transmission trends (Fig S2), model-
132 inference system estimates also enable assessment of key properties specific to the three
133 variants. Estimated transmissibility increased in conjunction with the widespread presence of
134 the new variant in each country (Fig 3B for B.1.1.7, 3E for B.1.351, and 3H for P.1). Overall,
135 estimated viral transmissibility increased by 46.6% (95% CI: 32.3 – 54.5%) for the B.1.1.7
136 variant, 32.4% (95% CI: 14.6 – 48.0%) for the B.1.351 variant, and 43.3% (95% CI: 30.3 – 65.3%)
137 for the P.1 variant, compared to the wild-type virus (Table 1). In addition, the model-inference
138 system also detects large increases of population susceptibility for the B.1.351 and P.1 variants,
139 but not the B.1.1.7 variant (Fig. 3F and 3I vs. 3C; and Table 1). Specifically, the model estimates
140 immune evasion for the B.1.351 variant among 61.3% (95% CI: 42.6 – 85.8%) of the population
141 infected with the wild-type virus during the first wave in South Africa and for the P.1 variant
142 among 52.5% (95% CI: 0 – 75.8%) of the population infected with the wild-type virus during the
143 first wave in Brazil.

144

145 **Competition among variants and potential future outcomes.**

146 As many places have now detected one or more of the three VOCs locally, it is important to
147 understand the potential pandemic outcomes given the characteristics of and competition
148 among variants, interactions with ongoing NPIs, and mass-vaccination. We thus use a multi-
149 variant, age-structured model to simulate potential pandemic outcomes for the period from
150 May 2021 to Aug 2021 under scenarios with different variant prevalence, NPIs, and vaccine

151 efficacy. We focus on NYC for which detailed data and estimates (e.g., contact patterns, variant
152 prevalence, and vaccination rates) are available.

153

154 Fig 4A and Fig S4 show example projections of infections and mortality assuming a best-case
155 scenario in which vaccine-induced immunity is as effective against all three VOCs as for the
156 wild-type virus. At the time of these simulations (i.e., end of April 2021), the B.1.1.7 variant
157 was the predominant VOC in NYC; however, given their estimated propensity for immune
158 escape, both B.1.351 and P.1 could outcompete B.1.1.7 and become predominant in the
159 coming months (Fig 4A, top panel). The relative prevalence of B.1.351 and P.1 depends largely
160 on their initial introduction and establishment in the population. These two variants would
161 arise at similar rates and co-dominate, if they are introduced and established in the population
162 simultaneously (Fig 4A, left panel). However, should either be established in the population
163 ahead of the other, it would become dominant and suppress but not preclude the rise of the
164 other variant (Fig 4A, middle and right panels). In addition, the B.1.351 variant would be slightly
165 more competitive if the vaccines are less effective against it than the P.1. variant (Fig 4B and
166 Table S1), as has been shown in laboratory studies^{7,9}.

167

168 Tallies of model-projected infections (Fig 4B) and deaths (Fig S4) reveal four key determinants
169 of future pandemic outcomes. First, simultaneous introduction of both the B.1.351 and P.1
170 variants would lead to larger increases of infections and mortality than sole introduction of
171 either variant (Fig 4B and Fig S4B). This result indicates the importance of limiting the
172 introduction of multiple VOCs. Second, maintaining very high vaccine efficacies against all
173 variants is critical for mitigating the risk of a large resurgence in populations with relatively high
174 vaccination coverage (e.g., compare the first three subplots in Fig 4B). Third, continued non-
175 pharmaceutical preventive measures will reduce infection resurgence as municipalities reopen
176 economies. For instance, even with high vaccine efficacy, a rapid, full reopening before a very
177 large proportion of the population are fully vaccinated could lead to approximately three times
178 as many infections as when reopening occurs more slowly (Fig 4A and Fig 4B first subplot). Four,
179 reassuringly, while projected trends for COVID-19-related mortality are in general similar to
180 those for infection (Fig S4A vs. Fig 4A), lower proportions of COVID-19-related deaths would
181 occur due to the increased transmission of B.1.351 and/or P.1 (Fig S4B vs. Fig 4B; note the
182 larger proportion of deaths due to B.1.1.7 than infections). This is due to the currently higher
183 vaccination coverage among older adults who have been prioritized for vaccination in NYC
184 similar to many other municipalities (see Table S4 for vaccination coverage by age group). This
185 finding emphasizes the importance of vaccine effectiveness against VOCs and of prioritizing
186 vulnerable populations for vaccination in order to prevent severe outcomes of COVID-19.

187

188 **DISCUSSION**

189 Despite vaccine availability, the future trajectory of the COVID-19 pandemic remains uncertain
190 due to the potential additional emergence and continued spread of multiple VOCs. To improve
191 understanding of future pandemic dynamics, here we have developed and applied a
192 comprehensive model-inference system to quantify key viral properties for three VOCs: B.1.1.7,
193 B.1.351, and P.1. Our estimates for the B.1.1.7 variant are consistent with detailed
194 epidemiological evidence (32.5 – 54.6% increase in transmissibility and minimal immune
195 evasion estimated here vs. 30 – 50% increase in secondary attack rate based on contact tracing
196 data^{3,4} and little immune evasion based on laboratory and real-world vaccination data^{5,6,15}). Our
197 estimates of the level of immune evasion for the B.1.351 and P.1 variants are also consistent
198 with antibody neutralization data suggesting both variants can evade prior immunity induced
199 by infection and vaccination, though to a larger extent for the B.1.351 variant.^{7,9} Here we
200 provide joint quantification of immune escape and the change in transmissibility for both
201 variants. Overall, the model-inference system estimates and model simulations suggest that
202 both B.1.351 and P.1 are likely more competitive than the B.1.1.7 variant due to their greater
203 propensity for immune escape. These estimates are consistent with observations from Qatar¹⁵
204 and Canada¹⁶ showing that the proportion of infections caused by B.1.351 and/or P.1 increased
205 despite earlier introduction and dominance of B.1.1.7 in these populations. Therefore, in spite
206 of the current widespread prevalence of B.1.1.7 in Europe and North America, importation of
207 B.1.351 and/or P.1 to these regions could replace B.1.1.7 dominance and lead to a further
208 increase of infections by either B.1.351, P.1 or both variants. Mass-vaccination with highly
209 effective vaccines is thus crucial for mitigating the risk of future SARS-CoV-2 resurgence,
210 particularly as economies re-open.

211
212 Our model-inference system estimates substantial immune escape for both P.1 and B.1.351.
213 For P.1, the mean estimate is bounded by a very broad confidence interval; however, for
214 B.1.351 the uncertainty is more constrained and indicates greater confidence that a substantial
215 level of immune escape occurs. These latter findings are supported by vaccine clinical trial and
216 real-world data. In particular, Shinde et al.¹³ found a similar likelihood of COVID-19, mostly due
217 to B.1.351, among trial participants who were seropositive at enrollment (i.e., after the first
218 wave) compared to those seronegative. Although potential differences in risk of exposure
219 among the two comparison groups may have somewhat biased this finding, substantial repeat
220 and breakthrough infection occurred due to B.1.351. Further, recent data from Qatar¹⁵ indicate
221 that individuals receiving the full dosing regimen of the Pfizer BNT162b2 mRNA vaccine are at
222 greater risk of breakthrough infection with the B.1.351 variant than B.1.1.7. Continued
223 monitoring of the severity of both repeat and breakthrough infections is needed to more fully
224 understand ongoing risks of VOCs to both health and economy. In addition, the Qatar study¹⁵
225 highlights the importance of full vaccination (i.e., administration of both doses for mRNA
226 vaccines), as participants gained little protection against severe illness after the first vaccine

227 dose (vs. ~50% efficacy for B.1.1.7), even though full protection against severe illness was
228 retained after two vaccine doses. It is thus likely critical that best vaccination protocols are
229 followed in order to confer protection against variants with immune escape properties and that
230 potential waning of vaccine-induced immunity over time is monitored.

231
232 In light of the spatial expansion of B.1.1.7, B.1.351, and P.1 and the potential emergence of
233 other new variants, vaccination is paramount for controlling the COVID-19 pandemic. It is
234 imperative that vaccine production, distribution, and administration proceed expeditiously,
235 particularly in resource-limited settings. Without effective global control of the pandemic,
236 continued transmission of SARS-CoV-2 will give rise to additional new variants and pose new
237 threats to all. As vaccines are distributed and administered, a continuation of non-
238 pharmaceutical preventive measures is needed to minimize infections among the unvaccinated.
239 As shown in our simulations, despite the relatively high vaccine coverage obtained to date in a
240 place like NYC, COVID-19 infections could resurge if such locations lift preventive measures
241 prematurely.

242
243 Due to a lack of sub-regional data, we used aggregated country-level data to estimate the
244 properties for the three VOCs. Our model-inference system also did not account for differences
245 of disease severity and infection-detection rate among age groups, which may vary
246 substantially.¹⁷ This model simplification may introduce uncertainty and bias to our estimates,
247 particularly for Brazil where country-level data may mask more intense transmission in
248 subpopulations and may have led to underestimation of the transmissibility of P.1.
249 Nevertheless, validation using independent data, including for Brazil, indicates that the model-
250 inference system is able to closely capture pandemic dynamics and accurately estimate
251 cumulative infection rates (Fig 2). Further, because the model-inference system simultaneously
252 accounts for population susceptibility, disease seasonality, NPIs, and vaccination, it is able to
253 specifically estimate changes related to a given new variant and closely matches available
254 epidemiological data (e.g., for B.1.1.7). The model simulations using these estimated
255 characteristics further illustrate the relative competitiveness of the three VOCs and delineate
256 key determinants of future infection outcomes. Overall, our findings point to the importance of
257 preventing the spread of the B.1.351 and P.1 variants, in addition to B.1.1.7, via continued NPIs,
258 prompt mass-vaccination of all populations, continued monitoring of vaccine efficacy, and
259 potentially updating vaccine formulations to ensure high efficacy.

260

261 **METHODS**

262 **Data sources and processing**

263 The model-inference system uses reported COVID-19 case and mortality data to capture
264 transmission dynamics, weather data to estimate disease seasonality, mobility data to

265 represent concurrent NPIs, and vaccination data to account for changes in population
266 susceptibility due to vaccination, for each of the three countries (i.e., the UK, South Africa, and
267 Brazil). Country-level daily COVID-19 case and mortality data came from COVID-19 Data
268 Repository of the Center for Systems Science and Engineering (CSSE) at Johns Hopkins
269 University;^{18,19} we aggregated the data to weekly intervals until the week of 4/11/2021 but
270 excluded initial weeks with low case rates (<2 per million population). Hourly surface station
271 temperature and relative humidity came from the Integrated Surface Dataset (ISD) maintained
272 by the National Oceanic and Atmospheric Administration (NOAA) and are accessible using the
273 “stationaRy” R package.^{20,21} We computed specific humidity using temperature and relative
274 humidity per the Clausius-Clapeyron equation.²² We then aggregated these data for all
275 weather stations in each country with measurements since 2000 and calculated the average for
276 each week of the year during 2000-2020. To compute the seasonal trend, we used a method
277 developed by Yuan et al.^{23,24} which estimates the relative reproduction number based on
278 temperature and specific humidity (see details in Supplemental Information). Daily mobility
279 data were derived from Google Community Mobility Reports;²⁵ we aggregated all business-
280 related categories (i.e., retail and recreational, transit stations, and workplaces) in all locations
281 in each country to weekly intervals. Daily vaccination data (for 1st and 2nd dose if applied) for
282 the UK were sourced from Public Health England;²⁶ and data for South Africa and Brazil were
283 obtained from Our World in Data.^{27,28}

284

285 **Model-inference system**

286 Contact tracing data capturing chains of transmission can be used to compute the secondary
287 attack rate and quantify changes in transmissibility due to a given new variant. Surveillance
288 data and laboratory viral characterization can be used to document and quantify levels of
289 immune evasion. Yet such detailed data are often not available, particularly for resource-
290 limited settings. Given these circumstances, mathematical modeling that assimilates epidemic
291 surveillance data provides an attractive alternate means for estimating key epidemiological
292 properties of novel variants, including the transmission rate. However, joint estimation of the
293 transmission rate and population susceptibility, which is related to immune evasion, is
294 challenging, as both quantities can change for a new variant. This problem arises
295 mathematically because the product of these two quantities, rather than either individually,
296 determines disease incidence – i.e., at any point in time, given the incidence, transmissibility
297 and susceptibility are not individually identifiable. Nevertheless, we posit that transmissibility
298 and susceptibility affect epidemic dynamics differentially *over time*, and, as such, data at
299 multiple time points can enable joint estimation. We thus design a model-inference system to
300 estimate the most plausible joint changes in these two quantities using commonly available
301 incidence and mortality time series.

302

303 The model-inference system is comprised of an epidemic model for simulating the transmission
304 dynamics of SARS-CoV-2 and a statistical inference method for estimating the model state
305 variables and parameters. The epidemic model is a susceptible-exposed-infectious-recovered-
306 susceptible (SEIRS) construct that further accounts for two-dose vaccination. In addition, to
307 include the effects of public health interventions and disease seasonality, it further adjusts the
308 transmission rate each week using mobility data and the estimated seasonal trend based on
309 climate conditions (see Eqn S3 in the Supplemental Information). The system then combines
310 the model-simulated (prior) estimates and observed case and mortality data to compute
311 posterior estimates using the ensemble Kalman adjustment filter (EAKF).²⁹ We also apply a
312 technique termed space re-probing³⁰ that accommodates possible large changes mid-pandemic
313 to transmissibility and population susceptibility. Further, due to the challenge identifying these
314 two quantities individually, we ran the model-inference system, repeatedly and in turn, in order
315 to test 14 major combinations of changes in transmissibility and susceptibility (see details in
316 Supplemental Information). Briefly, depending on the hypothesized change, we restricted the
317 EAKF update to a given related set of parameters or variables. For instance, for the hypothesis
318 that the new variant changes the transmissibility but does not evade immunity, the system only
319 allows major adjustment to the transmission rate and the infectious period; for the hypothesis
320 that the new variant induces both changes, the system allows major adjustment to the
321 transmission rate, the infectious period, and population susceptibility. The system then selects
322 the run with the best performance based on the accuracy of model-fit, one-step ahead
323 prediction, and magnitude of changes to key state variables to identify the most plausible
324 combination of changes in transmissibility and level of immune evasion (see Supplemental
325 Information). To approximate the distribution of the system (including all model state variables
326 and parameters), we employed an ensemble of model replica ($n = 500$ here) and updated the
327 ensemble posterior each week using the EAKF. In addition, to account for model stochasticity,
328 we repeated each model-inference simulation 100 times for each dataset, each with initial
329 parameters and variables randomly drawn from prior distributions (Table S2). Consequently,
330 model estimates are aggregated from 50,000 model runs in total.

331

332 **Estimation of variant-specific changes in transmissibility and level of immune evasion**

333 The model decouples the impact of concurrent NPIs and disease seasonality from the
334 transmission rate and infectious period (Eqn S3); as such, estimates for the latter two
335 parameters are variant-specific. We thus compute transmissibility as the product of the
336 transmission rate and infectious period. To reduce uncertainty, we average transmissibility
337 estimates over the first pandemic wave (excluding the first three weeks when model estimates
338 are less accurate) for the wild-type SARS-CoV-2 virus; similarly, we average the transmissibility
339 estimates over the period when the new variant is dominant. We identify this latter time period
340 based on the transmissibility estimates: 1) For the UK (B.1.1.7) and South Africa (B.1.351), the

341 estimated transmissibility increased and plateaued within 10 weeks (see Fig 3); we thus used
342 the period from the week with the maximal transmissibility during the 10 weeks following its
343 initial increase to the end of our study period (i.e., the week of 4/11/2021). However, for the
344 UK, we excluded the 3rd lockdown period when estimated transmissibility is lower, potentially
345 due to better awareness of B.1.1.7 and preventive measures taken at the time not fully
346 captured by the model. Of note, contact tracing data also indicate a lower increase of the
347 secondary attack rate around that time: 25-40% during 11/30/20 – 1/10/21 (among 1,364,301
348 cases for this expanded analysis)⁴ vs. 30-50% during 11/30 – 12/20/20 (among 386,805 cases).³
349 2) For Brazil (P.1), the estimated transmissibility increased more gradually (see Fig 3), we thus
350 instead used either the weeks identified per 1) or the last 8 weeks of our study period,
351 whichever with a longer time period, to ensure the robustness of estimation. We then
352 compute the average change in transmissibility due to a new variant as the ratio of the two
353 averaged estimates (i.e., after: before the rise of the new variant).

354
355 To quantify immune evasion, we record all time points inducing major EAKF adjustments to
356 posterior estimates of susceptibility, compute the change in immunity as $\Delta Imm = S_{t+1} - S_t + i_t$
357 (with S_t as the susceptibility at time- t and i_t as the new infections occurring at time- t), and sum
358 over all ΔImm estimates to compute the total change in immunity due to the new variant. We
359 then compute the level of immune evasion as the ratio of the total change in immunity during
360 the second wave to the model-estimated population immunity at the end of the first wave. This
361 ratio provides an estimate of the fraction of individuals previously infected who are susceptible
362 to re-infection with the new variant.

363
364 For both quantities, we report the mean and 95% CI based on the mean estimates from 100
365 repeated model-inference runs.

366 367 **Model validation using model-generated synthetic data**

368 To test the accuracy of the model-inference system, we generated 10 synthetic datasets using a
369 two-variant SEIRS model (see Eqn. S4) and different scenarios of changing transmissibility and
370 immune evasion (Table S3). In each scenario, a new variant was introduced at week 21 of the
371 simulation. We then combined the incidence and mortality due to both variants and added
372 noise drawn from a Poisson distribution to represent observational error. We then applied the
373 model-inference system to estimate the model variables and parameters for each synthetic
374 dataset, per the procedure described above for real data. For comparison with model-
375 inference system estimates, we computed the true values of population susceptibility and
376 transmissibility over time as the weighted average of the two variants based on the relative
377 prevalence at each time point (i.e., each week).

378

379 **Model validation using independent data**

380 To compare model estimates with independent observations not assimilated into the model-
381 inference system, we identified four relevant datasets: 1) the REACT study, which measures the
382 prevalence of SARS-CoV-2 using PCR-testing of volunteers from the general public living in the
383 UK. At the time of this study, the REACT study has conducted 10 rounds of testing during
384 5/1/2020 – 3/30/2021 (n = 1,572,951 tests in total)^{10,11}; Fig 2B plots our estimates of prevalence
385 of SARS-CoV-2 each week, overlaying all 10 measures from the REACT study for corresponding
386 time periods; 2) a serosurvey of workers in Cape Town, South Africa, conducted
387 during 8/17 – 9/4/2020 (n = 405 participants);¹² 3) serology tests among participants enrolled
388 during 8/17 – 11/25/2020 in the Novavax NVX-CoV2373 vaccine phase 2a-b trial in South Africa
389 (n=1324).¹³ Given this long enrollment period, we used the centered 2-week window (9/29 –
390 10/13/20) to match with our model estimates. and 4) two nationwide random household
391 serosurveys conducted in Brazil during 5/14 – 5/21/2020 (n = 25,025 participants) and 6/4 –
392 6/7/2020 (n = 31,165 participants).¹⁴ To account for the delay in antibody generation, we
393 shifted the timing of each serosurvey 14 days when comparing to model-inference system
394 estimates of cumulative infection rates in Fig 2 D and E.

395

396 **Model simulations testing the relative competitiveness of VOCs and projecting future** 397 **transmission dynamics.**

398 Here we modified a multi-variant model previously developed for influenza virus³¹ to include
399 age structure and interactions (Eqn. S5). The multi-variant model accounts for: 1) competition
400 between each pair of SARS-CoV-2 variants (e.g. wild-type and B.1.1.7) via cross-protective
401 immunity; 2) variant-specific transmissibility and population susceptibility, based on estimates
402 derived in this study; 3) variant-specific vaccine efficacy under different scenarios (see Table
403 S5); 4) age-specific differences in vaccination coverage at the start of simulation and
404 vaccination uptake rates for the simulation period (see Table S4); 5) seasonality; and 6) changes
405 in NPIs under different scenarios (see Table S5). We used data from NYC for baseline
406 vaccination coverage³² and initial prevalence of different variants,³³ as well as key model
407 estimates (e.g., transmission rates and infection fatality risk by age group; see Table S5)^{17,34,35}.
408 As in previous work,^{17,34,35} we included 8 age-groups (i.e. <1, 1-4, 5-14, 15-24, 25-44, 45-64, 65-
409 74, and 75+ year-olds) in the model to account for age-specific differences. To focus on the
410 three VOCs, we only included the B.1.1.7, B.1.351, and P.1 variants and attributed all other
411 variants as “wild-type” virus, even though at the start of the simulations, the B.1.526 variant
412 made up approximately one third of sequenced infections (N.b., the B.1.526 variant likely
413 emerged locally in NYC; we estimated a ~20% increase in transmissibility and nominal immune
414 evasion for this variant; based on these estimates the impact of this variant is expected to be
415 relatively minor). We did not account for potential differences in infection fatality risk by
416 variant, as such information is not available; therefore, the simulated mortality under different

417 scenarios only reflect the relative infection rate by age group, for which we apply age-specific
418 infection-fatality risk (see Table S5). In addition, due to uncertainty vis-à-vis the severity and
419 infection fatality risk among breakthrough infections (i.e., those who have been vaccinated), we
420 only show mortality-related simulations for the “Same VE” scenario which assumes no
421 reduction in VE.

422

423 **Data Availability:** All data used in this study are publicly available as described in the “Data
424 sources and processing” section.

425

426 **Code availability:** All source code and data necessary for the replication of our results and
427 figures will be made publicly available.

428

429 **Acknowledgements:** This study was supported by the National Institute of Allergy and
430 Infectious Diseases (AI145883), the National Science Foundation Rapid Response Research
431 Program (RAPID; DMS2027369) and a gift from the Morris-Singer Foundation.

432

433 **Author contributions:** WY designed the study (main), conducted the model analyses,
434 interpreted results, and wrote the first draft. JS designed the study (supporting), interpreted
435 results, and critically revised the manuscript.

436

437 **Competing interests:** JS and Columbia University disclose partial ownership of SK Analytics. JS
438 discloses consulting for BNI.

439 **References:**

- 440 1 Centers for Disease Control and Prevention. *SARS-CoV-2 Variant Classifications and*
441 *Definitions*, <[https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/variant-](https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/variant-surveillance/variant-info.html#Concern)
442 [surveillance/variant-info.html#Concern](https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/variant-surveillance/variant-info.html#Concern)> (2021).
- 443 2 National Collaborating Centre for Infectious Diseases. *Updates on COVID-19 Variants of*
444 *Concern*, <<https://nccid.ca/covid-19-variants/>> (2021).
- 445 3 Public Health England. Investigation of novel SARS-CoV-2 variant, Variant of Concern
446 202012/01, Technical briefing 3. (2020).
- 447 4 Public Health England. Investigation of novel SARS-CoV-2 variant, Variant of Concern
448 202012/01, Technical briefing 5. (2021).
- 449 5 Wang, P. *et al.* Antibody resistance of SARS-CoV-2 variants B.1.351 and B.1.1.7. *Nature*,
450 doi:10.1038/s41586-021-03398-2 (2021).
- 451 6 Planas, D. *et al.* Sensitivity of infectious SARS-CoV-2 B.1.1.7 and B.1.351 variants to
452 neutralizing antibodies. *Nat Med*, doi:10.1038/s41591-021-01318-5 (2021).
- 453 7 Hoffmann, M. *et al.* SARS-CoV-2 variants B.1.351 and P.1 escape from neutralizing
454 antibodies. *Cell* **184**, 2384-2393 e2312, doi:10.1016/j.cell.2021.03.036 (2021).

- 499 26 Public Health England. *Coronavirus (COVID-19) in the UK*,
500 [<https://coronavirus.data.gov.uk/details/download>](https://coronavirus.data.gov.uk/details/download) (2020).
- 501 27 *Data on COVID-19 (coronavirus) vaccinations by Our World in Data*,
502 [<https://github.com/owid/covid-19-data/tree/master/public/data/vaccinations>](https://github.com/owid/covid-19-data/tree/master/public/data/vaccinations) (2020).
- 503 28 Mathieu, E. *et al.* A global database of COVID-19 vaccinations. *Nature Human Behaviour*,
504 doi:10.1038/s41562-021-01122-8 (2021).
- 505 29 Anderson, J. L. An Ensemble Adjustment Kalman Filter for Data Assimilation. *Mon.*
506 *Weather Rev.* **129**, 2884-2903, doi:10.1175/1520-0493(2001)129<2884:aeakff>2.0.co;2
507 (2001).
- 508 30 Yang, W. & Shaman, J. A simple modification for improving inference of non-linear
509 dynamical systems. *arXiv*, 1403.6804 (2014).
- 510 31 Yang, W., Lau, E. H. Y. & Cowling, B. J. Dynamic interactions of influenza viruses in Hong
511 Kong during 1998-2018. *PLOS Computational Biology* **16**, e1007989,
512 doi:10.1371/journal.pcbi.1007989 (2020).
- 513 32 New York City Department of Health and Mental Hygiene. *COVID-19: Data - COVID-19*
514 *Vaccines*, [<https://www1.nyc.gov/site/doh/covid/covid-19-data-vaccines.page>](https://www1.nyc.gov/site/doh/covid/covid-19-data-vaccines.page) (2021).
- 515 33 New York City Department of Health and Mental Hygiene. *COVID-19: Data - COVID-19*
516 *Virus Variants*, [<https://www1.nyc.gov/site/doh/covid/covid-19-data-archive.page>](https://www1.nyc.gov/site/doh/covid/covid-19-data-archive.page)
517 (2021).
- 518 34 Yang, W., Kandula, S. & Shaman, J. *Projections of COVID19 Epidemic Outcomes and*
519 *Healthcare Demands for New York City (NYC)*, [<https://github.com/wan-
520 yang/COLUMBIA-COVID19-PROJECTIONS-FOR-NYC>](https://github.com/wan-yang/COLUMBIA-COVID19-PROJECTIONS-FOR-NYC) (2020).
- 521 35 Yang, W., Shaff, J. & Shaman, J. Effectiveness of non-pharmaceutical interventions to
522 contain COVID-19: a case study of the 2020 spring pandemic wave in New York City. *J R*
523 *Soc Interface* **18**, 20200822, doi:10.1098/rsif.2020.0822 (2021).
- 524

Table and figure captions

Table 1. Estimated changes in transmissibility and level of immune evasion. Numbers show model estimated mean (95% CI) from 100 model-inference runs.

Fig 1. Model-inference system validation using model-generated synthetic data with an infection-detection rate of 20%. (A) 5 sets of synthetic data (dots) and model-fits to each dataset; lines show mean estimates and surrounding areas show 50% (dark) and 95% (light) CrIs. (B) weekly model estimated transmissibility and population susceptibility; lines show mean estimates and surrounding areas show 50% (dark) and 95% (light) CrIs, compared to the true values (dots). (C) overall estimates of the change in transmissibility and immune evasion (boxes and whiskers show model estimated median, interquartile range, and 95% CI from 100 model-inference simulations) compared to the true values (dots).

Fig 2. Model-inference system fit to data for the three countries and validation using independent datasets. The left column shows the model fit to reported weekly case and mortality data for the UK (A), South Africa (C), and Brazil (D). Dots show the weekly number of

cases (in blue) and deaths (in red) per 1 million persons; boxes and whiskers show corresponding model estimates (mean, 50% and 95% Crls). Grey shaded boxes indicate the timing of lockdowns or key periods of restricted activity; horizontal arrows indicate the timing of variant identification and vaccination rollout. The right column compares available, independent measurements to corresponding model estimates. Red dots and error bars show measured prevalence over 10 periods of time from the REACT study for the UK (B), cumulative infection rates from two serology studies in South Africa (D), and cumulative infection rates from two nationwide household serosurveys in Brazil (F). Blue lines and surrounding areas show model estimated mean, 50% (dark) and 95% (light) Crls.

Fig 3. Key model-inference system estimates. Left column shows estimated real-time reproduction number R_t , middle column shows estimated transmissibility, and right column shows estimated population susceptibility for each week during the study period, for the three countries. For comparison, estimated weekly infection rates are superimposed in each plot (right y-axis). Blue lines and surrounding areas show the estimated mean, 50% (dark) and 95% (light) Crls. Boxes and whiskers show the estimated weekly infection rates (mean, 50% and 95% Crls). Grey shaded areas indicate the timing of lockdowns or key periods of restricted activity; horizontal arrows indicate the timing of variant identification and vaccination rollout. *Note that the transmissibility estimates (middle column) have removed the effects of changing population susceptibility, NPIs, and disease seasonality; thus, the trends are more stable than the reproduction number (R_t ; left column) and reflect changes in variant-specific properties.*

Fig 4. Model projections of infection, under different scenarios of VOC co-circulation, NPIs, and vaccine efficacy (VE). (A) Example projected epidemic trajectories for each variant assuming the VE is as high as for the wild-type virus. Lines and surrounding areas show model projected median and interquartile range (color-coded for each variant as indicated by the legend). (B) Tallies over the entire simulation period (May – Aug 2021) for different scenarios of seeding, VE (as indicated in the subtitles), and NPIs (as indicated by the x-axis labels). All numbers are scaled for 1 million people. For the projected percentages for each variant and uncertainty bounds, see Table S1.

Table

Table 1. Estimated changes in transmissibility and level of immune evasion. Numbers show model estimated mean (95% CI) from 100 model-inference runs.

Location	Variant	Changes in	
		Transmissibility (%)	Immune evasion (%)
United Kingdom	B.1.1.7	46.6 (32.3, 54.6)	3.9 (0, 36.2)
South Africa	B.1.351	32.4 (14.6, 48)	61.3 (42.6, 85.8)
Brazil	P.1	43.3 (30.3, 65.3)	52.5 (0, 75.8)

Fig 1. Model-inference system validation using model-generated synthetic data with an infection-detection rate of 20%. (A) 5 sets of synthetic data (dots) and model-fits to each dataset; lines show mean estimates and surrounding areas show 50% (dark) and 95% (light) CrIs. (B) weekly model estimated transmissibility and population susceptibility; lines show mean estimates and surrounding areas show 50% (dark) and 95% (light) CrIs, compared to the true values (dots). (C) overall estimates of the change in transmissibility and immune evasion (boxes and whiskers show model estimated median, interquartile range, and 95% CI from 100 model-inference simulations) compared to the true values (dots).

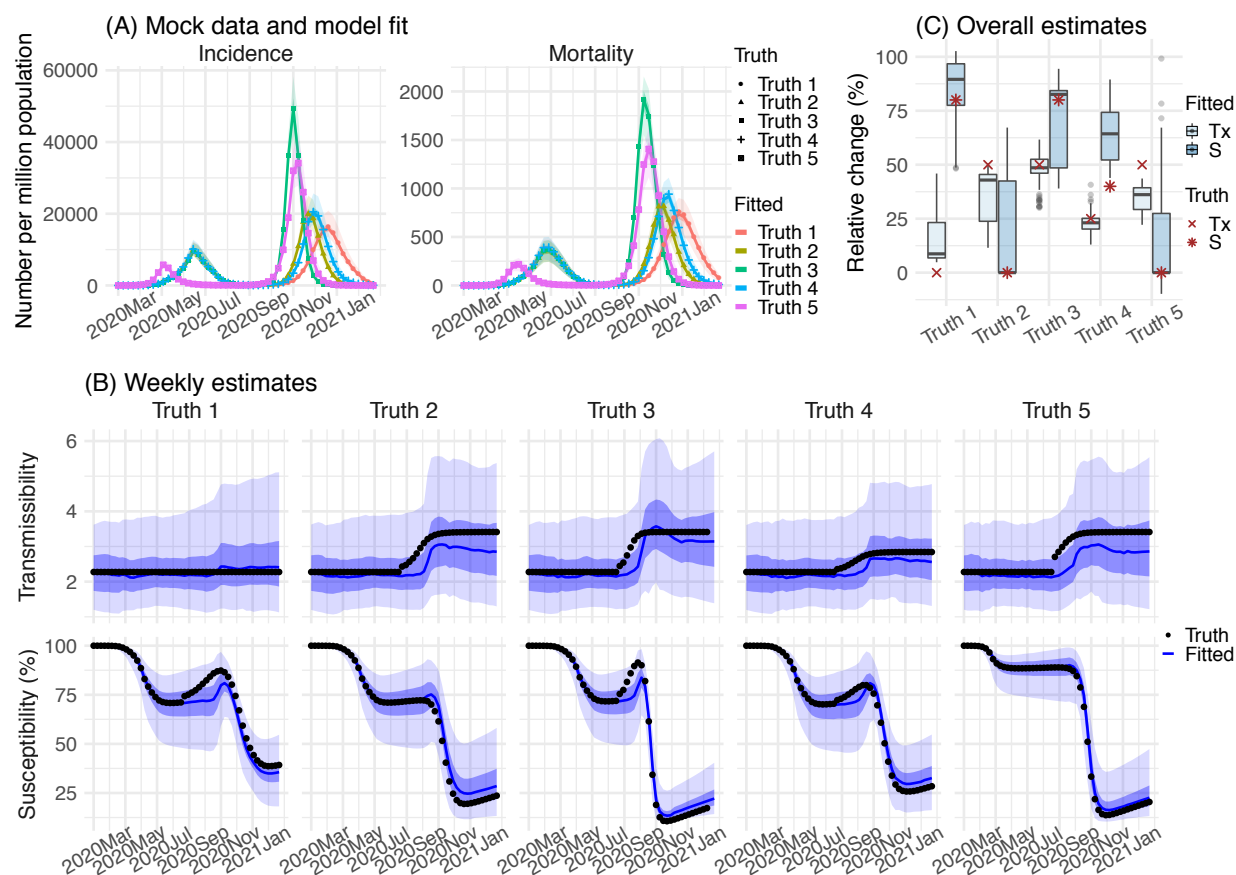


Fig 2. Model-inference system fit to data for the three countries and validation using independent datasets. The left column shows the model fit to reported weekly case and mortality data for the UK (A), South Africa (C), and Brazil (E). Dots show the weekly number of cases (in blue) and deaths (in red) per 1 million persons; boxes and whiskers show corresponding model estimates (mean, 50% and 95% CrIs). Grey shaded boxes indicate the timing of lockdowns or key periods of restricted activity; horizontal arrows indicate the timing of variant identification and vaccination rollout. The right column compares available, independent measurements to corresponding model estimates. Red dots and error bars show measured prevalence over 10 periods of time from the REACT study for the UK (B), cumulative infection rates from two serology studies in South Africa (D), and cumulative infection rates from two nationwide household serosurveys in Brazil (F). Blue lines and surrounding areas show model estimated mean, 50% (dark) and 95% (light) CrIs.

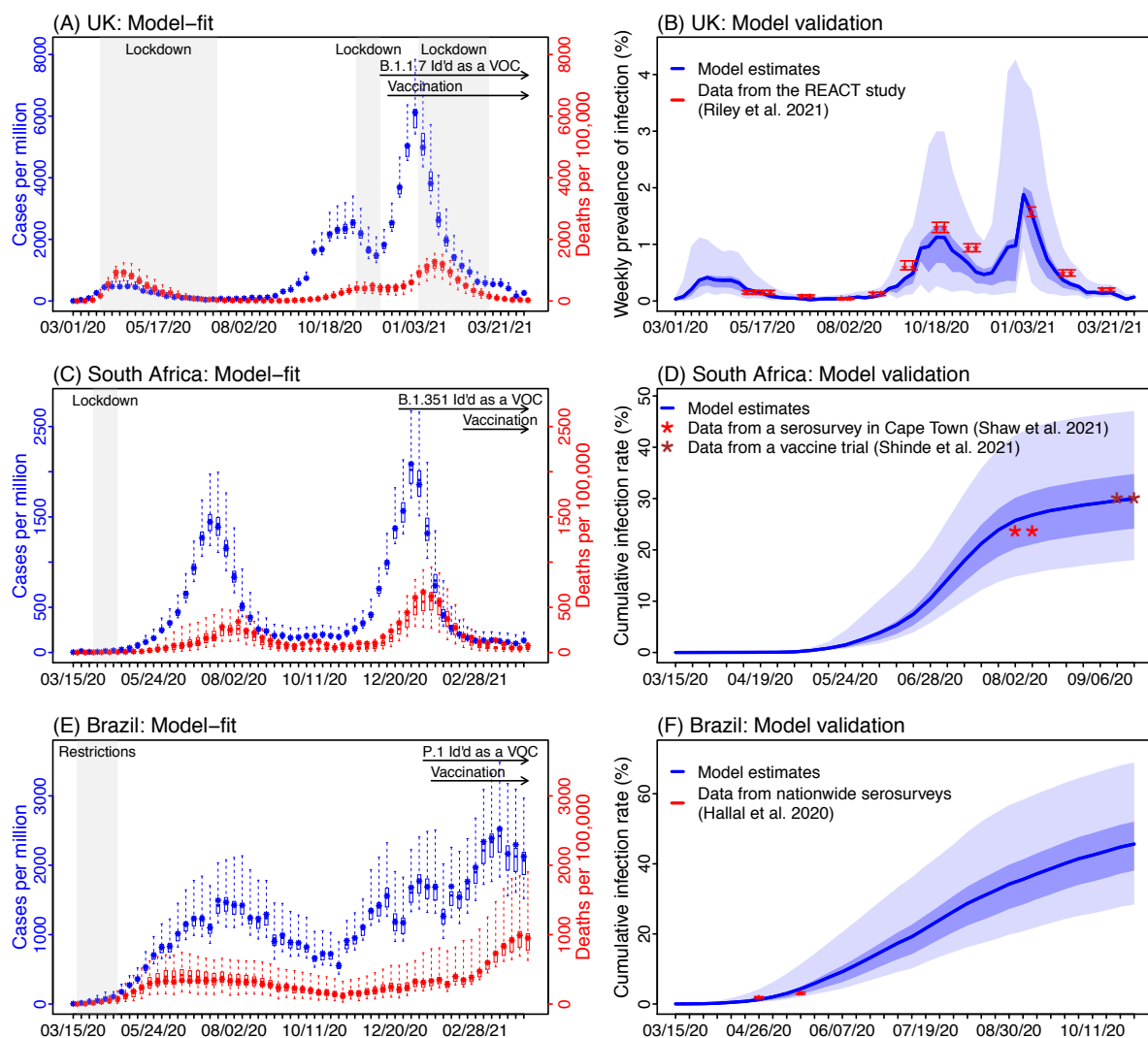


Fig 3. Key model-inference system estimates. Left column shows estimated real-time reproduction number R_t , middle column shows estimated transmissibility, and right column shows estimated population susceptibility for each week during the study period, for the three countries. For comparison, estimated weekly infection rates are superimposed in each plot (right y-axis). Blue lines and surrounding areas show the estimated mean, 50% (dark) and 95% (light) CrIs. Boxes and whiskers show the estimated weekly infection rates (mean, 50% and 95% CrIs). Grey shaded areas indicate the timing of lockdowns or key periods of restricted activity; horizontal arrows indicate the timing of variant identification and vaccination rollout. *Note that the transmissibility estimates (middle column) have removed the effects of changing population susceptibility, NPIs, and disease seasonality; thus, the trends are more stable than the reproduction number (R_t ; left column) and reflect changes in variant-specific properties.*

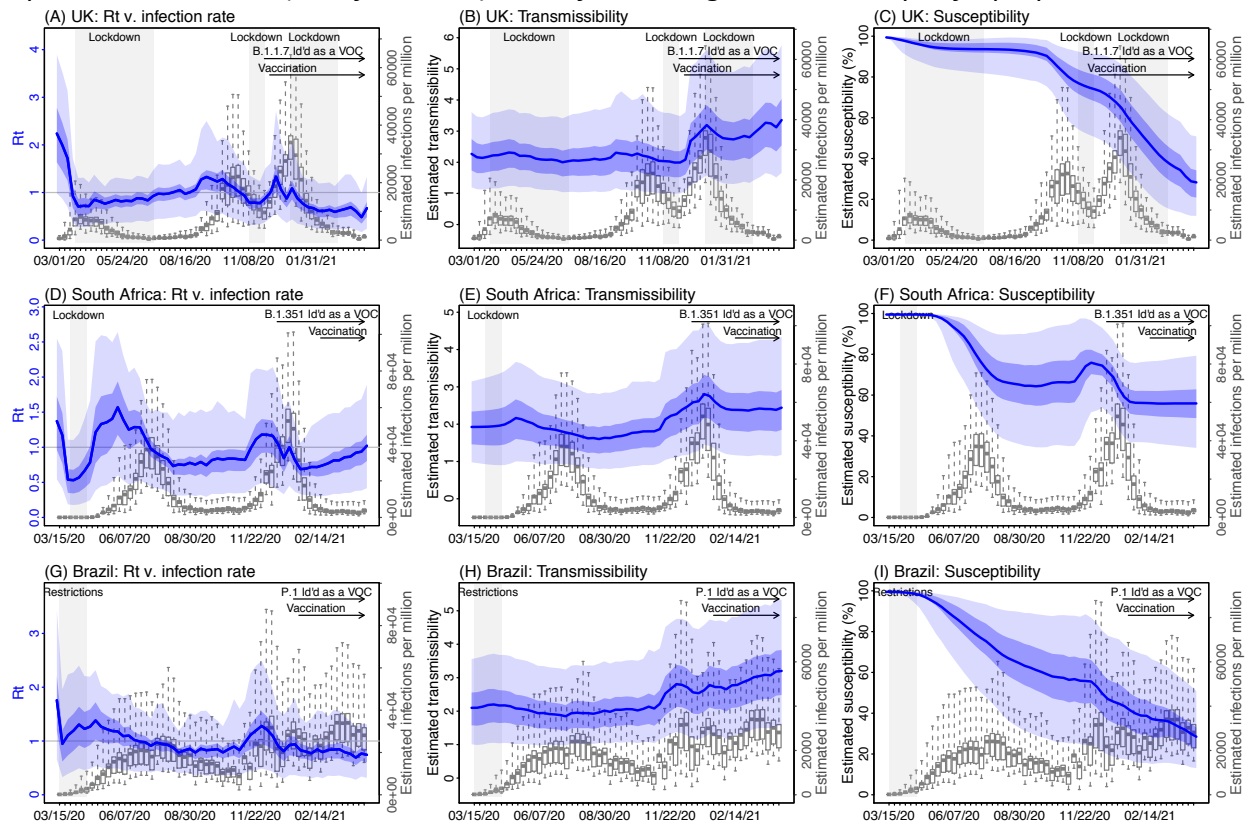
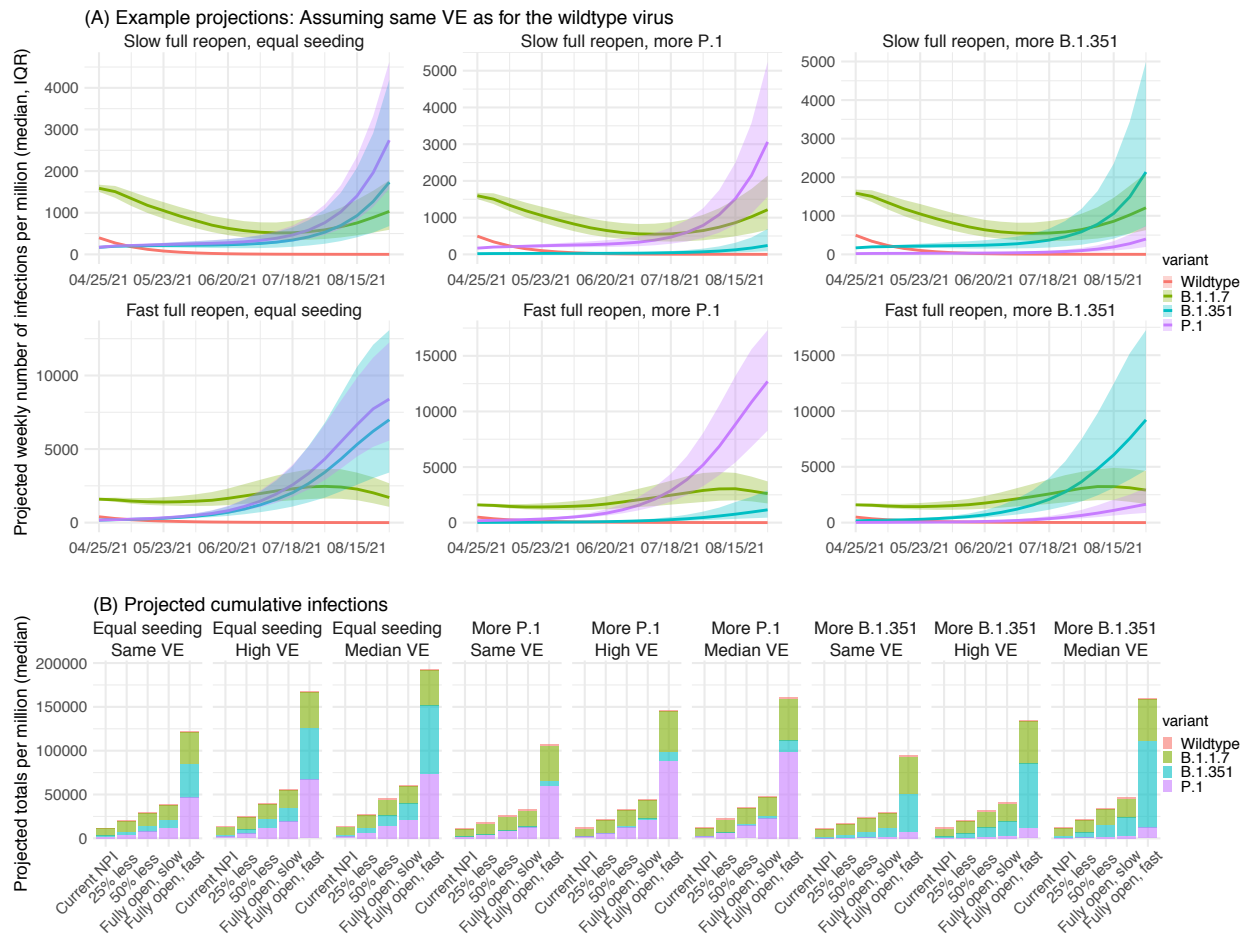


Fig 4. Model projections of infection, under different scenarios of VOC co-circulation, NPIs, and vaccine efficacy (VE). (A) Example projected epidemic trajectories for each variant assuming the VE is as high as for the wild-type virus. Lines and surrounding areas show model projected median and interquartile range (color-coded for each variant as indicated by the legend). (B) Tallies over the entire simulation period (May – Aug 2021) for different scenarios of seeding, VE (as indicated in the subtitles), and NPIs (as indicated by the x-axis labels). All numbers are scaled for 1 million people. For the projected percentages for each variant and uncertainty bounds, see Table S1.



Supplemental Information

for

Epidemiological characteristics of three SARS-CoV-2 variants of concern and implications for future COVID-19 pandemic outcomes

Wan Yang and Jeffrey Shaman

1. Estimating seasonal trends

Many respiratory infections tend to occur seasonally and are predominantly prevalent during certain months of the year (e.g., cold months in temperate climates).¹ This seasonal pattern has been documented for influenza viruses,² respiratory syncytial viruses,³ and endemic human coronaviruses.⁴ In addition, studies have showed that this seasonality may be associated with climate conditions – particularly, temperature and humidity – as they may modulate the survival and transmission of respiratory viruses.⁵⁻⁸ For the SARS-CoV-2 virus, our work has also shown that a winter-time seasonality exists, similar to endemic human coronaviruses in New York City (NYC), and that models accounting for this seasonality enable more accurate projection of COVID-19 pandemic dynamics than those do not.^{9,10} However, to date, no mechanistic models exist that quantify the response of the SARS-CoV-2 virus to temperature and humidity and in turn the seasonality of COVID-19. In addition, seasonal trends may differ by climate. For instance, epidemics of influenza can occur any time of the year in subtropical and tropical climates; it is thus more difficult to characterize the seasonality of respiratory infections in these climates. To address these challenges, we recently developed a flexible climate-forced model of epidemic dynamics for subtropical and tropical climates; results with this model also describe the response to temperature and humidity conditions common in temperate climates.^{11,12} Thus, to account for the potentially diverse seasonal trends of COVID-19 in the UK (temperate climates), South Africa (mostly temperate climates), and Brazil (mostly tropical climates), we applied this climate-forcing to temperature and humidity data for each country and computed the relative seasonal trend for each country.

Specifically, the climate-forcing takes the following form:

$$R_0(t) = [aq^2(t) + bq(t) + c] \left[\frac{T_c}{T(t)} \right]^{T_{exp}} \quad [S1]$$

where $R_0(t)$ is the basic reproduction number at time t ; $q(t)$ is specific humidity (i.e. a measure of absolute humidity) at time t ; and $T(t)$ is temperature at time t . In essence, the forcing function assumes that specific humidity has a bimodal effect on R_0 , with both low and high humidity conditions favoring transmission; in addition, this effect is moderated by temperature, where low temperatures promote transmission and temperatures above a certain threshold (i.e., T_c in Eqn. S1) limit transmission. Further, to link the coefficients a , b , and c to humidity q

36 and R_0 , Yuan et al.^{11,12} reparametrized the forcing function by solving the parabola with a nadir
37 at $(q_{mid}, R_{0max} - R_{0diff})$ and maxima at both (q_{min}, R_{0max}) and (q_{max}, R_{0max}) , such that:

38

$$\begin{cases} a = \frac{-b}{q_{max} + q_{min}} \\ b = \frac{R_{0diff}(q_{max} + q_{min})}{(q_{max} - q_{mid})(q_{min} - q_{mid})} \\ c = (R_{0max} - R_{0diff}) - aq_{mid}^2 - bq_{mid} \end{cases} \quad [S2]$$

39

40 Yuan et al.^{11,12} estimated the parameters R_{0max} (i.e., the maximum R_0), R_{0diff} (i.e., the difference
41 between the maximum and minimum R_0), q_{min} , q_{mid} , and q_{max} (i.e., the minimum, median, and
42 maximum specific humidity for the response), T_c (the threshold temperature) and T_{exp} (the
43 exponent in Eqn S1) for influenza in Hong Kong, a subtropical city, based on long-term epidemic
44 data collected therein during 1998 - 2018. Here we use their mean estimates for these
45 parameters and temperature and humidity data for each country (see main text and Fig S5) to
46 compute the seasonal trend for each country using Eqns. S1-2. However, as these parameters
47 were estimated for influenza, the outputs do not represent the actual R_0 for the SARS-CoV-2
48 virus. Thus, we instead compute the *relative* seasonal trend, by scaling the weekly country
49 output from Eqn S1 by the country mean output, such that this scaled output provides a
50 *relative*, seasonality-related transmissibility for each week of the year (see results in Fig S2).
51 These relative estimates also decouple the seasonality-related and variant-specific
52 transmissibility (assuming no interaction; see below).

53

54 **2. Model-inference system**

55 The model-inference system developed for this study consists of an SEIRSV model to simulate
56 the transmission dynamics of SARS-CoV-2 and the ensemble Kalman adjustment filter (EAKF)¹³
57 to estimate the model state variables and parameters, based on case and mortality data. Here
58 we describe the model and the filtering method in detail.

59

60 **2.1. Epidemic model.**

61 The SEIRSV (susceptible-exposed-infectious-recovered-susceptible-vaccination) model uses the
62 following set of equations to simulate the transition of sub-populations between different
63 disease stages, while accounting for disease seasonality, concurrent non-pharmaceutical
64 interventions (NPIs), and vaccination:

65

$$\begin{cases} \frac{dS}{dt} = \frac{R}{L} - \frac{b_t e_t m_t \beta_t IS}{N} - \varepsilon - v_1(t) - v_2(t) \\ \frac{dE}{dt} = \frac{b_t e_t m_t \beta_t IS}{N} - \frac{E}{Z} + \varepsilon \\ \frac{dI}{dt} = \frac{E}{Z} - \frac{I}{D} \\ \frac{dR}{dt} = \frac{I}{D} + v_1(t) + v_2(t) \end{cases} \quad [S3]$$

66
 67 where S , E , I , R are the number of susceptible, exposed (but not yet infectious), infectious, and
 68 recovered/deceased individuals, respectively; N is the population size. The parameter ε
 69 represents travel-related importation of infections (nominally set to 1 per 20 days per 1 million
 70 population, unless specified otherwise). To account for local seasonality, b_t , the estimated
 71 relative seasonal trend for each country (see Section 1 above and Fig S2) is used to adjust the
 72 relative transmission rate at time t . To account for concurrent NPIs, the term m_t , the relative
 73 population mobility at time t (in this study, we use data from Google Community Mobility
 74 Reports;¹⁴ see main text and Fig S2), is used to adjust the transmission rate. In addition, as the
 75 effectiveness of NPIs is unknown and variable, the model further includes a parameter, e_t , to
 76 scale NPI effectiveness at time t . For virus-specific characterization, β_t is the variant-specific
 77 transmission rate at time t , Z is the latency period, D is the infectious period, and L is the
 78 immunity period. Note that the parameters e_t , β_t , Z , D , and L are estimated by the model-
 79 inference system as described below.

80
 81 To incorporate changes in population susceptibility due to vaccination, the model accounts for
 82 two-dose vaccination via $v_1(t)$ and $v_2(t)$. Specifically, $v_1(t)$ is the number of individuals
 83 successfully immunized after the first dose of the vaccine and is computed using vaccination
 84 data and vaccine efficacy for 1st dose (see detailed settings in Table S2). Similarly, $v_2(t)$ is the
 85 additional number of individuals successfully immunized after the second vaccine dose
 86 (excluding those successfully immunized after the first dose).

87 88 2.2. Observation model.

89 We compute the model-simulated number of cases and deaths for each week using the model-
 90 simulated infection rate, as done in Yang et al.¹⁰ Specifically, we include 1) a time-lag from
 91 infectiousness to detection (i.e., an infection being diagnosed as a case) – drawn from a gamma
 92 distribution with a mean of T_m and standard deviation (SD) of T_{sd} days – to account for delays in
 93 diagnosis and detection; 2) an infection-detection rate (r), i.e. the fraction of infections
 94 (including subclinical or asymptomatic infections) reported as cases, to account for under-
 95 detection; 3) a time-lag from infectiousness to death, drawn from a gamma distribution with a
 96 mean of 14 days and a standard deviation of 10 days, empirically based on mortality data; and

97 4) an infection-fatality risk (IFR), i.e., the fraction of infections that die from COVID-19. To
98 compute the model-simulated number of new cases per week, we multiply the model-
99 simulated number of new infections per day by the infection-detection rate, and further
100 distribute these simulated cases in time per the distribution of time-from-infectiousness-to-
101 detection. We then aggregate the daily lagged, simulated cases to weekly totals for model
102 inference (see below). Similarly, to compute the model-simulated deaths per week and account
103 for delays in time to death, we multiply the simulated-infections by the IFR and then distribute
104 these simulated deaths in time per the distribution of time-from-infectiousness-to-death, and
105 aggregate these daily numbers to weekly totals. For each week, the infection detection rate (r),
106 the mean (T_m) and standard deviation (T_{sd}) of time-from-infectiousness-to-detection, and the
107 IFR are estimated based on weekly case and mortality data, along with other model
108 parameters.

109

110 2.3. Inference using the EAKF

111 At the end of each week, the inference system uses the EAKF to update the state variables and
112 parameters based on model-generated prior estimates and case and mortality data. Briefly, the
113 EAKF uses an ensemble of model realizations ($n=500$ here), each with initial parameters and
114 variables randomly drawn from a prior range (see Table S2). After model initialization, the
115 system integrates the model ensemble forward in time for a week (per Eqn S3) *stochastically* to
116 compute the prior distribution for each model state variable or parameter, as well as the
117 model-simulated number of cases and deaths for that week as described in Section 2.2. The
118 system then combines the prior estimates with the observed case and death data for the same
119 week to compute the posterior per Bayes' theorem.¹³ During this filtering process, the system
120 updates the posterior distribution of all model parameters and variables for each week.¹³ As
121 such, it is able to capture the time-varying changes in transmission dynamics including the
122 variant-specific transmission rate (β_t) and infectious period (D) – the two parameters that we
123 use to compute variant-specific transmissibility over time.

124

125 However, unlike previous studies using similar model-inference approaches, here we further
126 modify the EAKF filtering process to test different potential combinations of changes in
127 transmissibility and immune escape. To enable this exploration of systemic changes (e.g. due
128 to the emergence of a new variant), we randomly replace a small fraction of ensemble
129 members (3-10%) using values randomly drawn from specified ranges. This technique, termed
130 space reprobing (SR), was developed in order to explore state space without corrupting
131 performance of the filter.¹⁵ Specifically for this application, we apply SR to a given related set of
132 parameters/variables and restrict the EAKF update of non-related parameters/variables, for 14
133 different hypothesized behaviors. These hypothesized changes are as follows:

134

- 135 1) Hypothesis 1 (minor changes in transmissibility, no immune escape): Large updates are
136 only allowed for the two transmissibility-related parameters β_t and D ; to explore the
137 changes, the system applies SR to these two parameters using values drawn from prior
138 ranges 10-20% higher than the initial priors.
- 139 2) Hypothesis 2 (major changes in transmissibility only, no immune escape): Similar to 1);
140 but to explore the changes, the system applies SR to β_t and D using values drawn from
141 prior ranges 30-40% higher than the initial priors.
- 142 3) Hypothesis 3 (minor immune escape only, no changes to transmissibility): Large updates
143 are only allowed for S , the population susceptibility, up to a total loss of 50% of the prior
144 immunity.
- 145 4) Hypothesis 4 (major immune escape only, no changes to transmissibility): Large updates
146 are only allowed for S , the population susceptibility, up to a total loss of 95% of the prior
147 immunity.
- 148 5) Hypothesis 5 (minor changes in transmissibility + minor immune escape): combining 1)
149 and 3) above.
- 150 6) Hypothesis 6 (major changes in transmissibility + major immune escape): combining 2)
151 and 4) above.
- 152 7) Hypothesis 7 (minor changes in transmissibility + major immune escape): combining 1)
153 and 4) above.
- 154 8) Hypothesis 8 (major changes in transmissibility + minor immune escape): combining 2)
155 and 3) above.
- 156 9) Hypothesis 9 (changes in both transmissibility and immune escape, no restriction on
157 magnitude of change): Large updates are allowed for β_t and D as well as S . To explore
158 the changes, initial SR uses values drawn from prior ranges 10-20% higher than the
159 initial priors, and values up to 30-40% higher than the initial priors if the inference
160 system detects the prior continues to underestimate the observed cases and deaths
161 with the 10-20% initial increase in SR values. In addition, SR allows updates of S up to
162 95% of the prior immunity.

163
164 To account for slower changes in overall population immunity (i.e., in the entire country) as the
165 new variant gradually spreads to different sub-regions across a large geographic space, such as
166 in Brazil, we also explore the fitting using the following five additional settings:

- 167
168 10) Hypothesis 10 (immune escape only and changes to overall population immunity occur
169 slowly over time): Large updates are only allowed for S , up to a total loss of 95% of the
170 prior immunity; however, SR is applied to a smaller fraction of ensemble members than
171 in 1)-9) such that changes in S occur gradually.

- 172 11) Hypothesis 11 (minor changes in transmissibility + minor immune escape; both occur
173 slowly over time): Large updates are allowed for β_t and D as well as S . Adjustment to S
174 is allowed as in 10) but up to only 50% of prior immunity. In addition, for
175 transmissibility, the system applies SR to β_t and D using values drawn from prior ranges
176 10-20% higher than the initial priors.
- 177 12) Hypothesis 12 (major changes in transmissibility + minor immune escape; both occur
178 slowly over time): Similar to 11); however, for β_t and D , initial SR uses values drawn
179 from prior ranges 10-20% higher than the initial priors, and values up to 30-40% higher
180 than the initial priors if the inference system detects the prior continues to
181 underestimate the observed cases and deaths with the 10-20% initial increase in SR
182 values.
- 183 13) Hypothesis 13 (minor changes in transmissibility + major immune escape; both occur
184 slowly over time): Similar to the settings specified in 11) but adjustment to S is allowed
185 up to 95% of prior immunity.
- 186 14) Hypothesis 14 (major changes in transmissibility + major immune escape; both occur
187 slowly over time): Similar to settings specified in 12) but adjustment to S is allowed up
188 to 95% of the prior immunity.

189
190 We carry out the model-inference process for each of the 14 settings described above and for
191 each country dataset. We then select the most plausible hypothesis for each country based on
192 the following criteria: 1) model fitting to case and mortality data, as indicated by the relative
193 root-mean-squared-error (RRMSE) between the *posterior* estimates for the corresponding
194 variable (i.e. case rate or mortality rate) and data; 2) the accuracy of one-step ahead prediction,
195 as indicated by the RRMSE between the *prior* estimates for the corresponding variable (i.e. case
196 rate or mortality rate) and data; 3) the level of adjustment needed for two key variables, i.e.,
197 infection rate and case rate, as indicated by the RRMSE between the *prior* and *posterior*
198 estimates for each variable; 4) a penalty on the number of variables needing SR adjustment;
199 and 5) a penalty on the frequency of SR adjustment. We combine all these metrics by weighting
200 them heuristically using the following set of weights: 0.27 for the two metrics in 1); 0.13 for the
201 two metrics in 2), 0.03 for the two metrics in 3), and 0.07 for both 4) and 5). We also tested
202 other sets of weights and found that higher weights should be given to 1) and 2) based on
203 results from the synthetic testing where the ‘true’ values of the state variables and parameters
204 are known; however, in general, the final results are similar if there are minor changes to these
205 weights.

206
207 To account for model stochasticity, we repeat each model-inference 100 times for each
208 dataset, each with initial parameters and variables randomly drawn from the prior distributions
209 (Table S2). Each model-inference tests the 14 hypotheses described above, selects the one with

210 the best performance (i.e. minimizing the combined metric described above), and outputs the
 211 estimates of the best-performing run. That is, the model estimates reported in the main text
 212 are aggregated from 100 best-performing model runs (each with 500 ensemble members and
 213 totaling 50,000 individual model realizations).

214

215 **3. Model validation using model-generated synthetic data**

216 To test the accuracy of our model-inference system, we generate 10 synthetic datasets using a
 217 separate multi-variant SEIRS model, similar to models developed in Yang et al.¹⁶ and Gog &
 218 Grenfell.¹⁷ Within this model, variants can interact via cross-immunity, which protects a
 219 portion of individuals with prior infection (i.e. polarized immunity) by reducing transmission.
 220 Specifically, the model takes the following form:

221

$$\left\{ \begin{array}{l} \frac{dS_i}{dt} = \frac{R_i}{L_i} - \sum_j \frac{b_t e_t m_t c_{ij} \beta_j S_j I_j}{N} - \varepsilon \\ \frac{dE_i}{dt} = \frac{b_t e_t m_t \beta_i S_i I_i}{N} - \frac{E_i}{Z_i} + \varepsilon \\ \frac{dI_i}{dt} = \frac{E_i}{Z_i} - \frac{I_i}{D_i} \\ \frac{dR_i}{dt} = \frac{I_i}{D_i} - \frac{R_i}{L_i} \end{array} \right. \quad [S4]$$

222

223 where N is the population size; S_i , E_i , I_i , and R_i , are, respectively, the numbers of susceptible,
 224 exposed-but-not-yet infectious, infectious, and recovered individuals, with respect to variant- i
 225 (here, the wild-type SARS-CoV-2 virus or a new variant); β_i , D_i , and L_i are, respectively, the
 226 transmission rate, mean infectious period, and mean immunity period, for variant- i ; and c_{ij}
 227 measures the strength of cross-immunity to variant- i conferred by infection with variant- j (e.g.,
 228 close to 0 if it is weak and $c_{ii}=1$ for infection by the same variant). The parameter ε represents
 229 travel-related importation of infections; to generate the synthetic data (i.e., “truths”), we set ε
 230 to 1 per week for the first 5 weeks and 1 every 3 days for the rest of the first wave (here weeks
 231 1-20); for the second wave, as transmission has been established locally, we set ε to 0 for
 232 simplicity. The terms b_t , m_t , and e_t are the same as in Eqn S3 and account for seasonality and
 233 NPIs over time. For simplicity, we omit birth, death, and vaccination.

234

235 To generate the synthetic data (i.e., “truths”), we seed the Eq. S4 model with 2 infections of
 236 wild-type virus at the start of each simulation and 50 infections of a new variant at the start of
 237 Week 21, for $N = 1$ million people; we run the model stochastically with a daily time-step from
 238 the week starting 3/1/2020 to the week starting 2/21/2021 (i.e. 52 weeks in total) using the
 239 parameters listed in Table S3. To compute the weekly number of cases and deaths, we use the

240 same procedure as described in Section 2.2 above for each variant. We then combine the
 241 case/mortality estimates for both variants, add random noises drawn from a Poisson
 242 distribution to mimic observational error. The final noise-added weekly case and mortality time
 243 series are then used as synthetic data for testing the model-inference system (described in
 244 Section 2 above). To compare the posterior estimates of key parameters and variables (e.g.
 245 transmissibility and population susceptibility) from the model-inference system, we compute
 246 the true values of population susceptibility and transmissibility over time as the weighted
 247 average of the two variants based on the relative prevalence during each week. Fig 1 and Fig S1
 248 show the 10 model-generated truths including cases, deaths, and the computed “true” values
 249 of population susceptibility and transmissibility for each week of the simulation.

250

251 **4. Multi-variant, age-structured model for simulation to test the relative competitiveness of** 252 **VOCs and project future SARS-CoV-2 dynamics.**

253 We modify the multi-variant model in Eqn S4 to further include age structure and vaccination.
 254 The inclusion of age structure here allows incorporation of age-specific parameters (e.g.,
 255 transmission rate and infection-fatality risk) as well as age-specific vaccination coverage and
 256 rates. Specifically, this multi-variant, age-structured SEIRSV model takes the following form:

$$\left\{ \begin{array}{l} \frac{dS_i^A}{dt} = \frac{R_i^A}{L_i^A} - \sum_j b_t e_t m_t c_{ij} \sum_a \frac{\beta_j^{Aa} S_j^A I_j^a}{N^a} - \varepsilon_i - v_{i,1}^A(t) - v_{i,2}^A(t) \\ \frac{dE_i^A}{dt} = \sum_j b_t e_t m_t c_{ij} \sum_a \frac{\beta_j^{Aa} S_j^A I_j^a}{N^a} - \frac{E_i^A}{Z_i^A} + \varepsilon_i \\ \frac{dI_i^A}{dt} = \frac{E_i^A}{Z_i^A} - \frac{I_i^A}{D_i^A} \\ \frac{dR_i^A}{dt} = \frac{I_i^A}{D_i^A} - \frac{R_i^A}{L_i^A} + v_{i,1}^A(t) + v_{i,2}^A(t) \end{array} \right. \quad [S5]$$

257

258 Model parameters in Eqn S5 are similar to those in Eqn S4, except for those related to age,
 259 which are indicated by the superscripts. The vaccination model component is also similar to
 260 Eqn S3, but with age-stratification. However, the terms $v_{i,1}^A(t)$ and $v_{i,2}^A(t)$ are variant-specific,
 261 as indicated by the additional subscript i ; that is, they additionally account for the reduction in
 262 vaccine efficacy against the new variants, based on scenario assumptions specified in Table S5.

263

264 As an example, we simulate the transmission dynamics under different scenarios of variant
 265 prevalence, vaccine efficacy, and NPIs for a city like NYC, from the week of 4/25/2021 to the
 266 week of 8/22/2021 (i.e., approximately May - August 2021). We use data or estimates available
 267 for NYC to initialize the parameters and state variables needed for model simulations. In
 268 addition, we use our model-inference estimates for the VOCs for related parameters and

269 variables. Specifically, as in previous work,^{9,10,18} we include 8 age-groups (i.e. <1, 1-4, 5-14, 15-
270 24, 25-44, 45-64, 65-74, and 75+ year-olds) to account for age-specific differences. To focus on
271 the three VOCs, here we only include the B.1.1.7, B.1.351, and P.1 variants and attribute the
272 rest as “wildtype” virus. NYC data on variant prevalence among tested infections during the
273 weeks of March – April 2021 are used to inform the initial range of seeding for each variant
274 (Table S5). Initial conditions for the state variables (e.g., susceptibility and SARS-CoV-2
275 prevalence) for each age group are taken from estimates^{9,10,18} made for the week of 4/18/2021,
276 using detailed data (including case, mortality, COVID-19-related emergency visit, mobility, and
277 vaccination) during 3/1/2020 – 4/24/2021. For each age group, to compute the initial variant-
278 specific population susceptibility (S_i) at the start of a simulation, we move the estimated
279 proportion with immune escape for variant i among those who have had prior infection with
280 the wild-type virus but have not been vaccinated back to the susceptible compartment. The
281 number of people losing vaccine-induced immunity is computed based on scenario assumptions
282 determining the reduction in vaccine efficacy (see scenarios in Table S5).

283
284 The cross-immunity settings, i.e., values of c_{ij} 's in Eqn S5, come from our posterior model-
285 inference estimates of immune escape and are used for all age groups. To reduce uncertainty,
286 here we use the 80% CI estimates (see Table S5). For instance, as our model-inference estimate
287 of immune escape for B.1.351 has an 80% CI of 40.1 – 82.8%, we set $c(\text{B.1.351} \leftarrow \text{wildtype})$, the
288 cross-immune protection against B.1.351 conferred by prior infection of the wildtype virus
289 relative to variant-specific immunity, to values drawn from a uniform distribution ranging from
290 0.172 to 0.599 (i.e. cross-immunity is set to the complement of estimated immune escape). We
291 set all $c(\text{wildtype} \leftarrow \text{new variant})$ to 1 – that is, we assume full cross-immune protection against
292 the wild-type virus conferred by infection due to any VOC.

293
294 Similarly, the variant-specific transmission rates, i.e. β_i 's, come from our posterior model-
295 inference estimates of the relative transmissibility for each variant. For instance, as our model-
296 inference estimate of transmissibility for B.1.351 is 18.5 – 45.7% (80% CI) higher than that of
297 the wildtype virus, we set $\beta_{\text{B.1.351}}$ to 1.185 – 1.457 times of the estimate for β_{wildtype} . The same
298 scaling is applied to all age groups.

299
300 Due to a lack of information, we do not account for potential differences in infection fatality
301 risk by variant; therefore, the simulated mortality under different scenarios only reflect the
302 relative infection rate by age group, for which we apply age-specific infection-fatality risk (see
303 Table S5). In addition, due to the uncertainty of the infection fatality risk among breakthrough
304 infections (i.e., those who have been vaccinated), we only show mortality-related simulations
305 for the “Same VE” scenario which assumes no reduction in VE.

306

307 We run the model for each scenario 1000 times stochastically, with the parameters and initial
308 conditions randomly drawn from uniform distributions with ranges specified in Table S5.
309 Results are summarized from the 1000 model runs for each scenario.

310

311 **References:**

- 312 1 Moriyama, M., Hugentobler, W. J. & Iwasaki, A. Seasonality of Respiratory Viral
313 Infections. *Annu Rev Virol* **7**, 83-101, doi:10.1146/annurev-virology-012420-022445
314 (2020).
- 315 2 Tamerius, J. *et al.* Global influenza seasonality: reconciling patterns across temperate
316 and tropical regions. *Environ Health Perspect* **119**, 439-445, doi:10.1289/ehp.1002383
317 (2011).
- 318 3 Baker, R. E. *et al.* Epidemic dynamics of respiratory syncytial virus in current and future
319 climates. *Nat Commun* **10**, 5512, doi:10.1038/s41467-019-13562-y (2019).
- 320 4 Kissler, S. M., Tedijanto, C., Goldstein, E., Grad, Y. H. & Lipsitch, M. Projecting the
321 transmission dynamics of SARS-CoV-2 through the postpandemic period. *Science* **368**,
322 860+, doi:10.1126/science.abb5793 (2020).
- 323 5 Tamerius, J. D. *et al.* Environmental predictors of seasonal influenza epidemics across
324 temperate and tropical climates. *PLoS Pathog* **9**, e1003194,
325 doi:10.1371/journal.ppat.1003194 (2013).
- 326 6 Shaman, J. & Kohn, M. Absolute humidity modulates influenza survival, transmission,
327 and seasonality. *Proc Natl Acad Sci U S A* **106**, 3243-3248, doi:10.1073/pnas.0806852106
328 (2009).
- 329 7 Shaman, J., Pitzer, V. E., Viboud, C., Grenfell, B. T. & Lipsitch, M. Absolute humidity and
330 the seasonal onset of influenza in the continental United States. *Plos Biol* **8**, e1000316,
331 doi:10.1371/journal.pbio.1000316 (2010).
- 332 8 Yang, W. & Marr, L. C. Mechanisms by which ambient humidity may affect viruses in
333 aerosols. *Appl Environ Microbiol* **78**, 6781-6788, doi:10.1128/AEM.01658-12 (2012).
- 334 9 Yang, W., Shaff, J. & Shaman, J. Effectiveness of non-pharmaceutical interventions to
335 contain COVID-19: a case study of the 2020 spring pandemic wave in New York City. *J R*
336 *Soc Interface* **18**, 20200822, doi:10.1098/rsif.2020.0822 (2021).
- 337 10 Yang, W. *et al.* Estimating the infection-fatality risk of SARS-CoV-2 in New York City
338 during the spring 2020 pandemic wave: a model-based analysis. *The Lancet. Infectious*
339 *diseases* **21**, 203-212, doi:10.1016/S1473-3099(20)30769-6 (2021).
- 340 11 Yuan, H., Kramer, S. C., Lau, E. H. Y., Cowling, B. J. & Yang, W. Modeling Influenza
341 Seasonality in the Tropics and Subtropics. *PLoS Comput Biol* **In press** (2021).
- 342 12 Yuan, H., Kramer, S. C., Lau, E. H. Y., Cowling, B. J. & Yang, W. Modeling Influenza
343 Seasonality in the Tropics and Subtropics. *medRxiv*, 2021.2002.2004.21251148,
344 doi:10.1101/2021.02.04.21251148 (2021).

- 345 13 Anderson, J. L. An Ensemble Adjustment Kalman Filter for Data Assimilation. *Mon.*
346 *Weather Rev.* **129**, 2884-2903, doi:10.1175/1520-0493(2001)129<2884:aeakff>2.0.co;2
347 (2001).
- 348 14 Google Inc. *Community Mobility Reports*, <<https://www.google.com/covid19/mobility/>>
349 (2020).
- 350 15 Yang, W. & Shaman, J. A simple modification for improving inference of non-linear
351 dynamical systems. *arXiv*, 1403.6804 (2014).
- 352 16 Yang, W., Lau, E. H. Y. & Cowling, B. J. Dynamic interactions of influenza viruses in Hong
353 Kong during 1998-2018. *PLOS Computational Biology* **16**, e1007989,
354 doi:10.1371/journal.pcbi.1007989 (2020).
- 355 17 Gog, J. R. & Grenfell, B. T. Dynamics and selection of many-strain pathogens. *Proc Natl*
356 *Acad Sci U S A* **99**, 17209-17214, doi:10.1073/pnas.252512799 (2002).
- 357 18 Yang, W., Kandula, S. & Shaman, J. *Projections of COVID19 Epidemic Outcomes and*
358 *Healthcare Demands for New York City (NYC)*, <[https://github.com/wan-](https://github.com/wan-yang/COLUMBIA-COVID19-PROJECTIONS-FOR-NYC)
359 [yang/COLUMBIA-COVID19-PROJECTIONS-FOR-NYC](https://github.com/wan-yang/COLUMBIA-COVID19-PROJECTIONS-FOR-NYC)> (2020).
360

Fig S1. Model validation using model-generated synthetic data with an infection-detection rate set to 10%. (A) 5 sets of synthetic data (dots) and model-fits to each dataset; lines show mean estimates and surrounding areas show 50% (dark) and 95% (light) Crls. (B) weekly model estimated transmissibility and population susceptibility; lines show mean estimates and surrounding areas show 50% (dark) and 95% (light) Crls, compared to the true values (dots). (C) overall estimates of the change in transmissibility and immune evasion (boxes and whiskers show model estimated median, interquartile range, and 95% CI from 100 model-inference simulations) compared to the true values (dots).

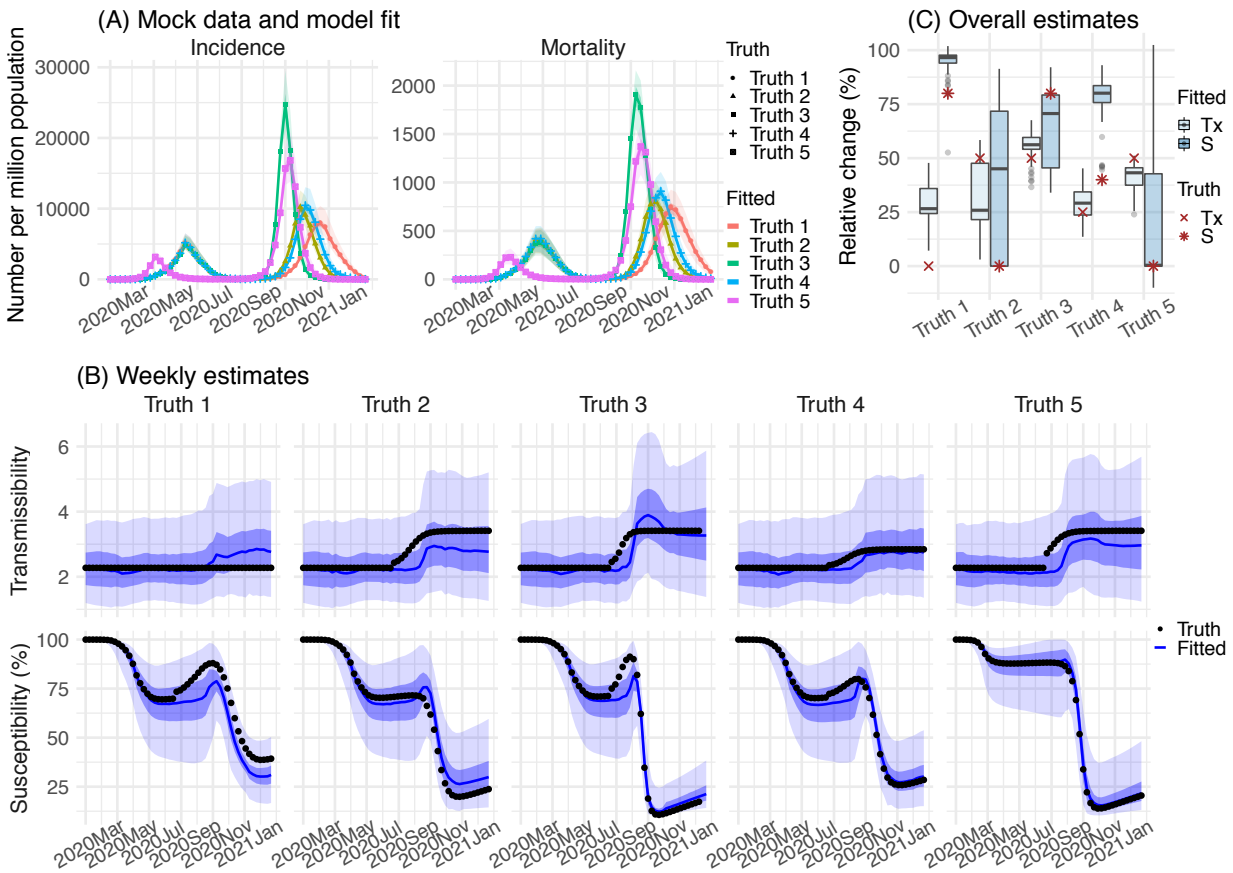


Fig S2. Pandemic dynamics, mobility, and estimated seasonal trends in the three countries.

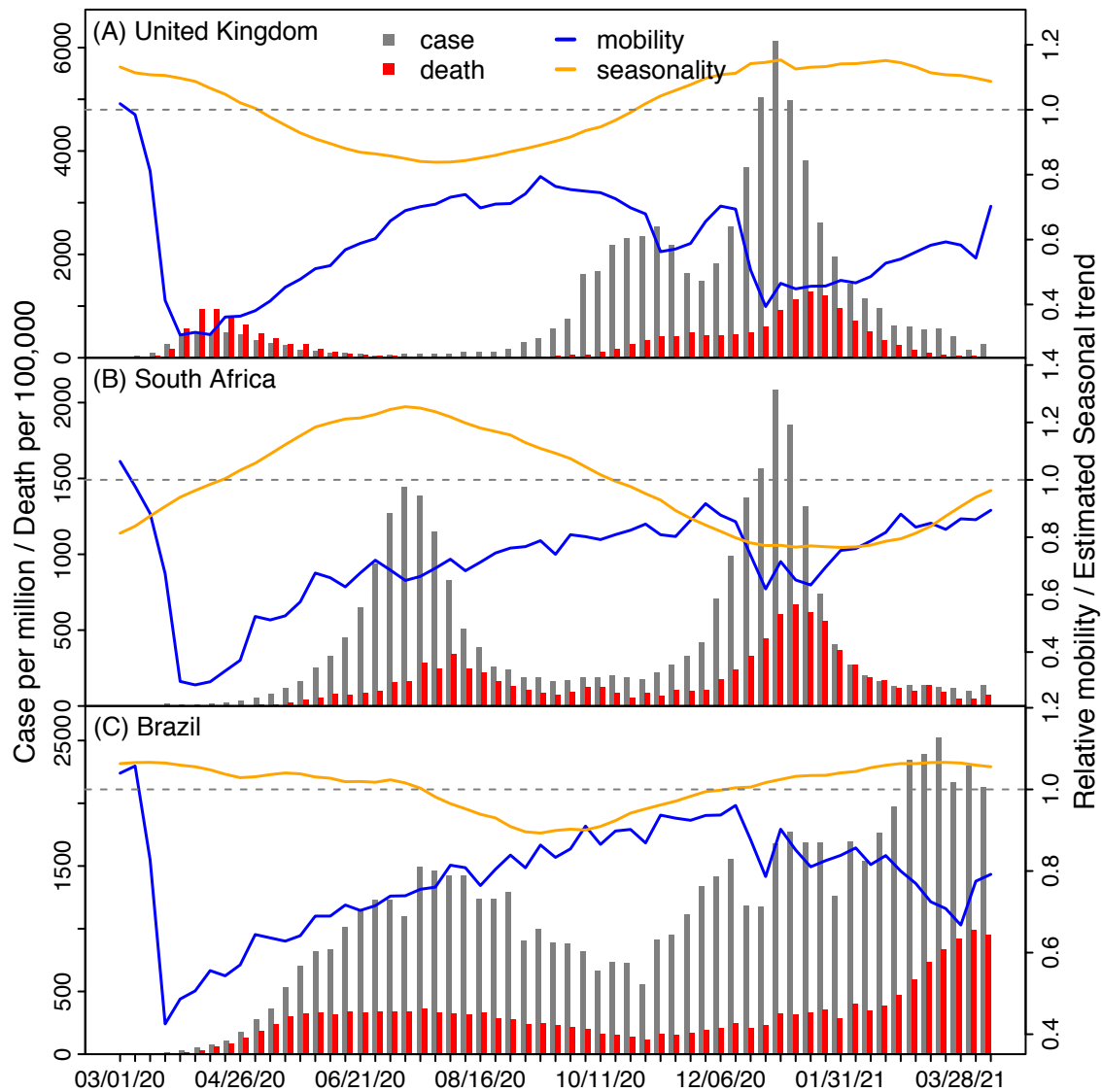


Fig S3. Other model estimates. Left panel shows the estimated infection-detection rate and right panel shows the estimated infection-fatality risk for each week during the study period for the three countries. For comparison, estimated weekly infection rates are superimposed in each plot (right y-axis). Blue lines and surrounding areas show model estimated mean, 50% and 95% CrIs. Boxes and whiskers show model-estimated weekly infection rates (mean, 50% and 95% CrIs). Grey shaded boxes indicate the timing of lockdowns or key period of restrictions; horizontal arrows indicate the timing of variant identification and vaccination rollout.

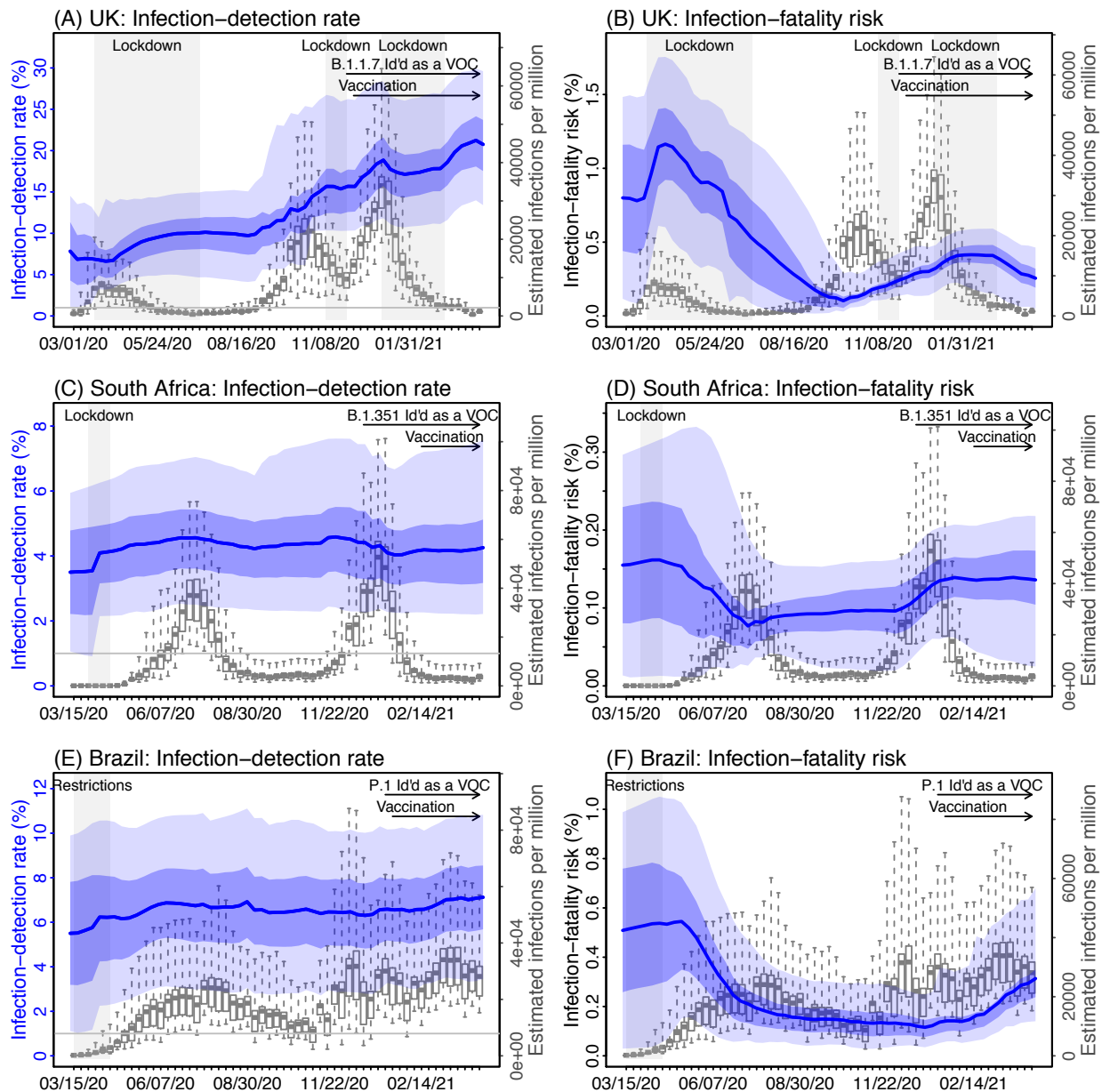


Fig S4. Model projections of *COVID-19* related mortality under different scenarios of VOC co-circulation and NPI. Due to the uncertainty on the infection fatality risk among breakthrough infections (i.e., those who have been vaccinated), all simulations shown here assume no reduction in vaccine efficacy (VE). All numbers are scaled for 1 million people.

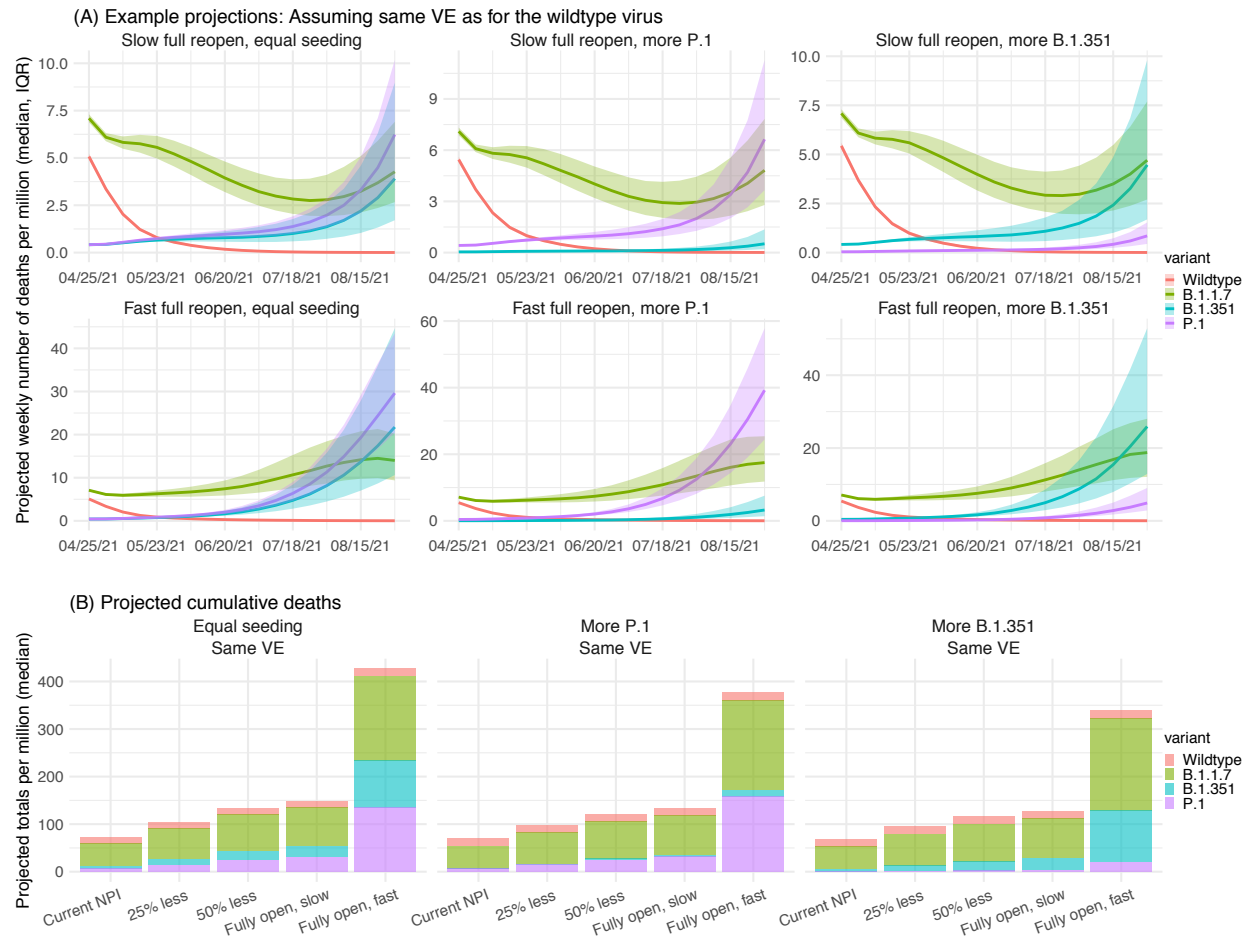


Fig S5. Weekly average temperature and specific humidity for the three countries.

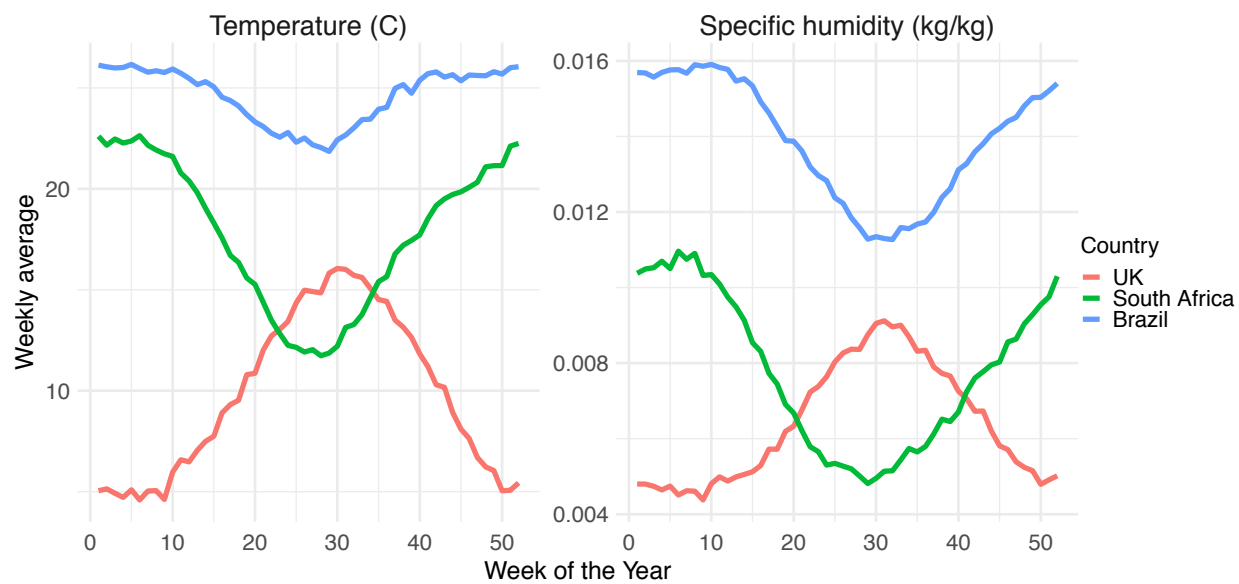


Table S1. Model-simulated prevalence of different variants under different scenarios. Numbers show the median (and interquartile range; all in percentage) of tallies over the entire simulation period (i.e. the week of 4/25/2021 to 8/22/2021) for each scenario, as specified in columns 1 (seeding of the B.1.351 and P.1 variant), 2 (NPI), and 3 (reduction on vaccine efficacy).

Seeding	NPI	VE	Wildtype	B.1.1.7	B.1.351	P.1
Equal seeding	Current NPI	Same VE	8.7 (8.1, 9.4)	65.4 (62.1, 67.9)	11.9 (9.8, 15)	13.2 (11.5, 15.2)
Equal seeding	Current NPI	High VE	7.7 (7.1, 8.3)	63.9 (60.2, 67)	13.3 (10.5, 17)	14.3 (12.4, 16.6)
Equal seeding	Current NPI	Median VE	7.6 (6.9, 8.3)	63.1 (59.5, 66.3)	13.7 (10.8, 17.6)	14.6 (12.6, 17)
Equal seeding	25% less	Same VE	5.7 (4.9, 6.5)	57.1 (51.4, 62.3)	15.9 (11.3, 21.8)	19.5 (15.9, 23.8)
Equal seeding	25% less	High VE	4.5 (3.8, 5.4)	52.6 (47, 59)	19 (12.8, 26.6)	20.9 (16.9, 26.4)
Equal seeding	25% less	Median VE	4.3 (3.5, 5.1)	50.2 (43.1, 56.8)	20.5 (13.9, 30)	22.2 (17.6, 27.2)
Equal seeding	50% less	Same VE	3.8 (3, 4.7)	48.2 (39.9, 54.8)	19.5 (12.5, 29.3)	25.5 (19.7, 32.3)
Equal seeding	50% less	High VE	2.8 (2.1, 3.5)	40.7 (32.2, 49.1)	24.3 (14.7, 37.4)	27.6 (21.4, 36)
Equal seeding	50% less	Median VE	2.5 (1.9, 3.3)	36.6 (28.2, 45)	27.8 (16.2, 41.2)	28.7 (21.5, 37.5)
Equal seeding	Fully open, slow	Same VE	2.9 (2.2, 3.7)	40.5 (32.4, 49.8)	20.9 (13.1, 35.3)	28.9 (21.8, 38)
Equal seeding	Fully open, slow	High VE	2 (1.5, 2.7)	33.9 (25.5, 43.7)	24.8 (15.1, 41.5)	32.5 (24, 42.2)
Equal seeding	Fully open, slow	Median VE	1.8 (1.3, 2.4)	30.6 (22.6, 40)	30.6 (17, 46.3)	32.5 (23.1, 42.3)
Equal seeding	Fully open, fast	Same VE	1 (0.9, 1.3)	26.8 (20.4, 35.4)	30.1 (17.2, 45.5)	36.7 (26.3, 48)
Equal seeding	Fully open, fast	High VE	0.8 (0.6, 0.9)	22.4 (16.7, 30)	33.7 (20.4, 50.8)	38.3 (26.7, 50.1)
Equal seeding	Fully open, fast	Median VE	0.7 (0.6, 0.8)	19.5 (15, 26)	39.9 (24.1, 56.1)	36.8 (24.9, 48.9)
More P.1	Current NPI	Same VE	11.9 (11.1, 12.8)	71.5 (69.4, 73.5)	1.6 (1.2, 2.1)	14.6 (12.9, 16.7)
More P.1	Current NPI	High VE	10.8 (9.9, 11.6)	70.9 (68.5, 73.3)	1.7 (1.3, 2.4)	16.1 (13.9, 18.6)
More P.1	Current NPI	Median VE	10.6 (9.8, 11.5)	70.9 (68.1, 73.1)	1.9 (1.4, 2.5)	16.3 (14.1, 18.9)
More P.1	25% less	Same VE	8.2 (7.1, 9.4)	66 (61.6, 69.9)	2.4 (1.6, 3.5)	22.4 (18.4, 27.5)
More P.1	25% less	High VE	6.9 (5.8, 7.9)	63.6 (57.7, 68.2)	2.7 (1.7, 4.4)	26 (21.2, 31.9)
More P.1	25% less	Median VE	6.6 (5.6, 7.5)	62 (56.4, 66.5)	3.1 (1.9, 5.1)	27.3 (22.6, 33)
More P.1	50% less	Same VE	5.7 (4.6, 7)	58.6 (50.8, 64.8)	3 (1.7, 5)	31.1 (24.6, 39.6)
More P.1	50% less	High VE	4.4 (3.4, 5.4)	52.5 (44.6, 59.7)	3.9 (2, 7.5)	36.5 (29.2, 45.7)

More P.1	50% less	Median VE	4.1 (3.2, 5.1)	49.7 (42.3, 56.9)	4.5 (2.3, 8.5)	39.3 (32.1, 47.4)
More P.1	Fully open, slow	Same VE	4.4 (3.4, 5.8)	53.3 (44.5, 60.5)	3.5 (1.8, 6.8)	36.8 (28.8, 46.1)
More P.1	Fully open, slow	High VE	3.2 (2.5, 4.2)	45 (36.8, 53.6)	4.4 (2.3, 8.8)	45 (35.7, 53.9)
More P.1	Fully open, slow	Median VE	3 (2.3, 3.9)	43 (34.3, 50.8)	4.9 (2.5, 10.3)	45.6 (36.1, 55.6)
More P.1	Fully open, fast	Same VE	1.6 (1.3, 2)	36.1 (28.6, 45.2)	5.1 (2.3, 10.8)	53.2 (43.8, 63.2)
More P.1	Fully open, fast	High VE	1.2 (1, 1.4)	30.9 (24.3, 39.2)	6.9 (3, 14.2)	56.7 (47.8, 66.2)
More P.1	Fully open, fast	Median VE	1 (0.9, 1.3)	27.8 (21.9, 35.4)	8.4 (3.9, 17.1)	58.8 (47.7, 67.4)
More B.1.351	Current NPI	Same VE	12.1 (11.3, 12.9)	72.5 (69.8, 74.8)	13.2 (10.9, 16.6)	1.8 (1.5, 2.2)
More B.1.351	Current NPI	High VE	10.9 (10, 11.8)	72.2 (68.7, 74.9)	14.6 (11.5, 18.6)	1.9 (1.5, 2.4)
More B.1.351	Current NPI	Median VE	10.6 (9.8, 11.5)	71.5 (67.6, 74.6)	15.5 (12.3, 19.8)	2 (1.6, 2.5)
More B.1.351	25% less	Same VE	8.4 (7.2, 9.6)	68.9 (62, 73.6)	19.3 (13.7, 27)	2.8 (2, 3.8)
More B.1.351	25% less	High VE	7 (5.9, 8.1)	66.9 (58.1, 72.6)	22.3 (15.6, 32.3)	3.2 (2.3, 4.5)
More B.1.351	25% less	Median VE	6.5 (5.4, 7.8)	62.9 (53.8, 70.5)	26.3 (17.5, 36.5)	3.3 (2.4, 4.7)
More B.1.351	50% less	Same VE	5.9 (4.7, 7.3)	62.7 (51.1, 70.5)	26.4 (17.1, 39.2)	4.1 (2.7, 6)
More B.1.351	50% less	High VE	4.5 (3.5, 5.8)	56.6 (43.7, 67.4)	32.4 (20.1, 46.9)	4.9 (3.3, 7.3)
More B.1.351	50% less	Median VE	4.1 (3, 5.3)	53 (39.6, 63.4)	36.8 (24.3, 52.5)	5 (3.2, 7.6)
More B.1.351	Fully open, slow	Same VE	4.7 (3.5, 6.2)	58.2 (43.8, 68.5)	30.7 (18, 47)	4.9 (3.2, 7.6)
More B.1.351	Fully open, slow	High VE	3.4 (2.4, 4.6)	50 (35.9, 62.4)	38.5 (23.5, 56.1)	5.7 (3.5, 8.9)
More B.1.351	Fully open, slow	Median VE	3.1 (2.2, 4.2)	44.7 (30.9, 57.4)	43.6 (28.2, 61)	5.9 (3.3, 9.5)
More B.1.351	Fully open, fast	Same VE	1.7 (1.3, 2.3)	42.2 (28.6, 55.5)	44.4 (28.7, 63.1)	7.5 (4.4, 13.1)
More B.1.351	Fully open, fast	High VE	1.2 (1, 1.6)	34.4 (23, 48.2)	53.4 (35.1, 69.1)	8.2 (4.7, 13.7)
More B.1.351	Fully open, fast	Median VE	1.1 (0.8, 1.4)	29.2 (19.9, 41.9)	59.7 (42.1, 72.6)	7.7 (4.3, 13.2)

Table S2. Prior ranges for the parameters used in the model-inference system for the three countries.

Parameter/ variable	Symbol	Prior range	Source/rationale
Initial exposed	$E(t=0)$	5 – 50 times (for the UK), 1 – 10 times (for South Africa), or 2.5 – 25 times (for Brazil) of the reported number of cases on the first week of simulation	Low infection-detection rate in first weeks; likely higher initial introduction for the UK due to more international travel per capita.
Initial infectious	$I(t=0)$	Same as for $E(t=0)$	
Initial susceptible	$S(t=0)$	99 – 100% of the population	Almost everyone is susceptible initially
Population size	N	N/A	Based on data
Variant-specific transmission rate	β	[0.5, 0.8] for the UK; [0.4, 0.7] for South Africa; [0.4, 0.8] for Brazil	Based on R_0 estimates of around 1.5-4 for SARS-CoV-2. ¹⁻³ Slightly lower ranges are used for South Africa and Brazil, as initial testing showed that the priors tend to overestimate the observations.
Scaling of effectiveness of NPI	e	[0.5, 1.5]	Around 1, with a large bound to be flexible.

Latency period	Z	[2, 5] days	Incubation period: 5.2 days (95% CI: 4.1, 7) ¹ ; latency period is likely shorter than the incubation period
Infectious period	D	[2, 5] days	Time from symptom onset to hospitalization: 3.8 days (95% CI: 0, 12.0) in China, ⁴ plus 1-2 days viral shedding before symptom onset. We did not distinguish symptomatic/asymptomatic infections.
Immunity period	L	[730, 1095] days	Assuming immunity lasts for 2-3 years
Mean of time from viral shedding to diagnosis	T_m	[5, 7] days for the UK, [5, 8] days for South Africa and Brazil	From a few days to a week from symptom onset to diagnosis/reporting, ⁴ plus 1-2 days of viral shedding (being infectious) before symptom onset. There may be a slightly longer delay for South Africa and Brazil.
Standard deviation (SD) of time from viral shedding to diagnosis	T_{sd}	[1, 3] days	To allow variation in time to diagnosis/reporting

Infection-detection rate	r	<p><u>For the UK:</u> starting from U[0.01, 0.15] at time 0 and allowed to increase over time using space re-probing⁵ with values drawn from U[0.1, 0.24] starting at week 5 (the week of 3/29/2020), U[0.05, 0.2] during the summer (July – Aug 2020), U[0.1, 0.3] starting Sep 2020, and U[0.15, 0.3] starting Nov 2020.</p> <p><u>For South Africa:</u> starting from U[0.01, 0.06] at time 0 and allowed to change over time using space re-probing⁵ with values drawn from U[0.02, 0.08] starting at week 5 (the week of 4/12/2020).</p> <p><u>For Brazil:</u> starting from U[0.01, 0.1] at time 0 and allowed to change over time using space re-probing⁵ with values drawn from U[0.02, 0.1] starting at week 5 (the week of 4/12/2020).</p>	<p>Large uncertainties; therefore, in general we use large prior bounds and large bounds for space re-probing (SR). Note that SR is only applied to 3-10% of the ensemble members and r can migrate outside either the initial range or the SR ranges during EAKF update. Efforts were made in the UK to increase detection of infection; however, detection during the summer of 2020 was likely lower because more infections at the time occurred among younger age groups with no or mild symptoms. In South Africa, due to the younger age structure in the population, infection detection rates were likely lower. In Brazil, infection-detection rates were likely low throughout the pandemic.</p>
Infection fatality risk (IFR)		<p><u>For the UK:</u> starting from U[0.001, 0.015] at time 0 and allowed to change over time using space re-probing⁵ with values drawn from U[0.0001, 0.003] during Jun – Sep 2020 when infections occurred mostly among younger ages, values drawn from U[0.0001, 0.005] during Oct – Dec 2020, and [0.0001, 0.006] during Jan – Apr 2021 to account for higher IFR for B.1.1.7.</p> <p><u>For South Africa:</u> starting from [0.0001, 0.003] at time 0 and allowed to change over time using space re-probing⁵</p>	<p>Based on previous estimates⁶ but extend to have wider ranges. Note that SR is only applied to 3-10% of the ensemble members and IFR can migrate outside either the initial range or the SR ranges during EAKF update.</p>

with values drawn from $U[0.0001, 0.0015]$ during the week of 4/19/2020 to the week of 7/5/2020 when case fatality risk was lower as computed from the data, values drawn from $U[0.0001, 0.002]$ starting the week of 12/13/2020 with the rise of B.1.351.

For Brazil: starting from $U[0.001, 0.01]$ at time 0 and allowed to change over time using space re-probing⁵ with values drawn from $U[0.0001, 0.007]$ starting at week 5 (the week of 4/12/2020), values drawn from $U[0.0001, 0.0035]$ during the week of 6/21/2020 to the week of 11/8/2020 when case fatality risk was lower as computed from the data, values drawn from $U[0.0001, 0.006]$ starting the week of 1/3/2021 with the rise of P.1, and values drawn from $U[0.0001, 0.01]$ starting the week of 2/7/2021 when the healthcare systems began to be overwhelmed.

Vaccine efficacy (VE)

For the UK: VE = 85% fourteen days after the 1st dose, and 95% seven days after the 2nd dose.

For South Africa: VE = 60% fourteen days after the 1st dose; no 2nd dose (for J&J vaccine).

For Brazil: VE = 45% fourteen days after the 1st dose, and 55% seven days after the 2nd dose.

During our study period (up to mid-April 2021), the UK mostly used the Pfizer and later on Oxford/AstraZeneca vaccine, both shown to be highly effective against both the wildtype virus and B.1.1.7. South Africa mostly used the J&J vaccine with just one dose. Brazil mostly used the Sinovac and Oxford/AstraZeneca vaccines with relatively lower VE against P.1.

References:

- 1 Li, Q. *et al.* Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia. *New Engl J Med*, doi:10.1056/NEJMoa2001316 (2020).
- 2 Wu, J. T., Leung, K. & Leung, G. M. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *Lancet*, doi:10.1016/S0140-6736(20)30260-9 (2020).
- 3 Li, R. *et al.* Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science* **368**, 489-493, doi:10.1126/science.abb3221 (2020).
- 4 Zhang, J. *et al.* Evolving epidemiology and transmission dynamics of coronavirus disease 2019 outside Hubei province, China: a descriptive and modelling study. *The Lancet. Infectious diseases*, doi:10.1016/S1473-3099(20)30230-9 (2020).
- 5 Yang, W. & Shaman, J. A simple modification for improving inference of non-linear dynamical systems. *arXiv*, 1403.6804 (2014).
- 6 Verity, R. *et al.* Estimates of the severity of coronavirus disease 2019: a model-based analysis. *The Lancet. Infectious diseases*, doi:10.1016/S1473-3099(20)30243-7 (2020).

Table S3. Parameters used to generate the synthetic data for model validation

Truth	Settings
Wildtype-virus, same for all truths	$\beta = 0.65$ per day; $Z = 3.5$ days; $D = 3.5$ days; $L = 2.5$ years; $T_m = 6$ days; $T_{sd} = 2$ days; IFR = 0.7%
Variant in Truth 1	Large outbreak during the first wave (See Figs 1 and S1). No increase in transmissibility (i.e. same β and D as for the wildtype virus); 80% increase in immune escape. Other parameters same as the wildtype virus.
Variant in Truth 2	Large outbreak during the first wave (See Figs 1 and S1). 50% increase in transmissibility (i.e. $\beta = 0.65 \times 1.5 = 0.975$ per day); no immune escape. Other parameters same as the wildtype virus.
Variant in Truth 3	Large outbreak during the first wave (See Figs 1 and S1). 50% increase in transmissibility (i.e. $\beta = 0.65 \times 1.5 = 0.975$ per day); 80% increase in immune escape. Other parameters same as the wildtype virus.
Variant in Truth 4	Large outbreak during the first wave (See Figs 1 and S1). 25% increase in transmissibility (i.e. $\beta = 0.65 \times 1.25 = 0.8125$ per day); 40% increase in immune escape. Other parameters same as the wildtype virus.
Variant in Truth 5	Small outbreak during the first wave (See Figs 1 and S1). 50% increase in transmissibility (i.e. $\beta = 0.65 \times 1.5 = 0.975$ per day); no immune escape. Other parameters same as the wildtype virus.
Infection- detection rate	20% for results shown in Fig 1 and 10% for results shown in Fig S1.

Table S4. Cumulative vaccination coverage used in the multi-variant, age-structured model simulations. Baseline vaccination coverage (as of 4/24/2021) is based on data for NYC. Projected vaccination rates for the simulation period (from the week starting 4/25/2021 to the week ending 8/29/2021) are based on data 10 days preceding the simulation and assuming a cumulative vaccination uptake of 80%. Note that ages under 15 (i.e., <1, 1-4, 5-14 years) are combined in the same column as they were not eligible to receive the vaccines at the time of this study. In addition, the numbers are aggregated from neighborhood-level estimates and thus could slightly exceed the 80% total for some age groups when some neighborhoods had actual vaccination coverage above 80% at baseline.

Time point	Vaccine dose	Vaccination coverage (%) by age group (in year)					
		0-14	15-24	25-44	45-64	65-74	75+
Start of simulation:	1st dose	0	25.4	41.7	52.9	69	58.2
4/24/2021	2nd dose	0	10.5	26.7	40.7	62.6	53.6
End of simulation:	1st dose	0	70	76	79.1	80.6	76.8
8/29/2021	2nd dose	0	65.4	73.4	78.7	80.3	74.5

Table S5. Parameter settings for different scenarios to simulate and project the impact of different variants of concern. Initial conditions and parameters are randomly drawn from uniform distributions with lower and upper bounds as specified below, based on estimates for NYC made for the week of 4/18/2021 using data during 3/1/2020 – 4/24/2021. Numbers associated with the parameter names denote the corresponding age groups.

scenario: Seeding/VE/NPI	variant	parameter	lower bound	upper bound
Equal seeding	wildtype	Initial seeding (%)	45	61
Equal seeding	B.1.1.7	Initial seeding (%)	35	45
Equal seeding	B.1.351	Initial seeding (%)	2	5
Equal seeding	P.1	Initial seeding (%)	2	5
More B.1.351	wildtype	Initial seeding (%)	49.4	62.8
More B.1.351	B.1.1.7	Initial seeding (%)	35	45
More B.1.351	B.1.351	Initial seeding (%)	2	5
More B.1.351	P.1	Initial seeding (%)	0.2	0.6
More P.1	wildtype	Initial seeding (%)	49.4	62.8
More P.1	B.1.1.7	Initial seeding (%)	35	45
More P.1	B.1.351	Initial seeding (%)	0.2	0.6
More P.1	P.1	Initial seeding (%)	2	5
All	All	Travel-related importation, ϵ_i	1 per 21 days, for the entire city ($N = 8.4$ million)	
Same VE	wildtype	VE = 85% fourteen days after the 1st dose and 95% seven days after the 2nd days		
Same VE	B.1.1.7	VE = 85% fourteen days after the 1st dose and 95% seven days after the 2nd days		
Same VE	B.1.351	VE = 85% fourteen days after the 1st dose and 95% seven days after the 2nd days		
Same VE	P.1	VE = 85% fourteen days after the 1st dose and 95% seven days after the 2nd days		
High VE	wildtype	VE = 85% fourteen days after the 1st dose and 95% seven days after the 2nd days		
High VE	B.1.1.7	VE = 85% x .95 = 80.75% fourteen days after the 1st dose and 95% seven days after the 2nd days		
High VE	B.1.351	VE = 85% x .80 = 68% fourteen days after the 1st dose and VE = 95% x .9 = 85.5% seven days after the 2nd days		
High VE	P.1	VE = 85% x .85 = 72.25% fourteen days after the 1st dose and VE = 95% x .95 = 90.25% seven days after the 2nd days		

Median VE	wildtype	VE = 85% fourteen days after the 1st dose and 95% seven days after the 2nd days		
Median VE	B.1.1.7	VE = 85% x .95 = 80.75% fourteen days after the 1st dose and VE = 95% x .95 = 90.25% seven days after the 2nd days		
Median VE	B.1.351	VE = 85% x .70 = 59.5% fourteen days after the 1st dose and VE = 95% x .8 = 76% seven days after the 2nd days		
Median VE	P.1	VE = 85% x .8 = 68% fourteen days after the 1st dose and VE = 95% x .9 = 85.5% seven days after the 2nd days		
Current NPI	all variants	no further increase in transmission rate		
25% less NPI	all variants	transmission rate increases by 5% per week up to 25% in total		
50% less NPI	all variants	transmission rate increases by 5% per week up to 50% in total		
Full reopen, slow	all variants	transmission rate increases by 5% per week up to fully reopen (~85% of current level)		
Full reopen, fast	all variants	transmission rate increases by 10% per week up to fully reopen (~85% of current level)		
	wildtype	β_{11} (per day, same below)	0.11	0.16
	wildtype	β_{22}	0.078	0.11
	wildtype	β_{33}	0.11	0.15
	wildtype	β_{44}	0.13	0.19
	wildtype	β_{55}	0.16	0.25
	wildtype	β_{66}	0.13	0.19
	wildtype	β_{77}	0.13	0.19
	wildtype	β_{88}	0.11	0.16
	wildtype	β_{12}	0.055	0.08
	wildtype	β_{13}	0.014	0.021
	wildtype	β_{14}	0.0057	0.0084
	wildtype	β_{15}	0.018	0.026
	wildtype	β_{16}	0.0082	0.012
	wildtype	β_{17}	0.0058	0.0085
	wildtype	β_{18}	0.004	0.0059
	wildtype	β_{21}	0.039	0.055
	wildtype	β_{23}	0.01	0.014
	wildtype	β_{24}	0.0041	0.0058
	wildtype	β_{25}	0.013	0.018
	wildtype	β_{26}	0.0059	0.0083

wildtype	β_{27}	0.0042	0.0059
wildtype	β_{28}	0.0029	0.004
wildtype	β_{31}	0.013	0.018
wildtype	β_{32}	0.013	0.018
wildtype	β_{34}	0.0066	0.0089
wildtype	β_{35}	0.0087	0.012
wildtype	β_{36}	0.0048	0.0066
wildtype	β_{37}	0.0034	0.0046
wildtype	β_{38}	0.0044	0.006
wildtype	β_{41}	0.0073	0.011
wildtype	β_{42}	0.0073	0.011
wildtype	β_{43}	0.012	0.017
wildtype	β_{45}	0.012	0.017
wildtype	β_{46}	0.0089	0.013
wildtype	β_{47}	0.0032	0.0046
wildtype	β_{48}	0.0059	0.0086
wildtype	β_{51}	0.075	0.12
wildtype	β_{52}	0.075	0.12
wildtype	β_{53}	0.061	0.096
wildtype	β_{54}	0.048	0.076
wildtype	β_{56}	0.046	0.073
wildtype	β_{57}	0.028	0.043
wildtype	β_{58}	0.027	0.042
wildtype	β_{61}	0.04	0.058
wildtype	β_{62}	0.04	0.058
wildtype	β_{63}	0.033	0.047
wildtype	β_{64}	0.039	0.056
wildtype	β_{65}	0.04	0.058
wildtype	β_{67}	0.044	0.064
wildtype	β_{68}	0.04	0.058
wildtype	β_{71}	0.024	0.035
wildtype	β_{72}	0.024	0.035
wildtype	β_{73}	0.017	0.025
wildtype	β_{74}	0.0074	0.011
wildtype	β_{75}	0.018	0.025
wildtype	β_{76}	0.025	0.036
wildtype	β_{78}	0.051	0.074

wildtype	β_{81}	0.017	0.024
wildtype	β_{82}	0.017	0.024
wildtype	β_{83}	0.018	0.025
wildtype	β_{84}	0.013	0.018
wildtype	β_{85}	0.016	0.023
wildtype	β_{86}	0.028	0.04
wildtype	β_{87}	0.043	0.061
wildtype	Z ₁ (days, same below)	3.3	4.3
wildtype	Z ₂	3.3	4.3
wildtype	Z ₃	3.4	4.4
wildtype	Z ₄	3.4	4.4
wildtype	Z ₅	3.5	4.4
wildtype	Z ₆	3.5	4.5
wildtype	Z ₇	3.4	4.5
wildtype	Z ₈	3.3	4.4
wildtype	D ₁	3.1	4
wildtype	D ₂	3.1	4.1
wildtype	D ₃	3	4
wildtype	D ₄	3.3	4.2
wildtype	D ₅	3.2	4.2
wildtype	D ₆	3.3	4.2
wildtype	D ₇	3.5	4.4
wildtype	D ₈	3.3	4.3
wildtype	IFR ₁	7.70E-05	0.00012
wildtype	IFR ₂	7.60E-05	0.00013
wildtype	IFR ₃	7.70E-05	0.00012
wildtype	IFR ₄	7.60E-05	0.00013
wildtype	IFR ₅	0.00031	0.00049
wildtype	IFR ₆	0.0033	0.0047
wildtype	IFR ₇	0.018	0.023
wildtype	IFR ₈	0.056	0.069
B.1.1.7	Increase in transmission rate	0.403	0.5227
B.1.1.7	Immune escape	0	0.1
B.1.351	Increase in transmission rate	0.1849	0.457
B.1.351	Immune escape	0.4414	0.8281
P.1	Increase in transmission rate	0.3682	0.4945
P.1	Immune escape	0.3588	0.6683

B.1.1.7	Increase in transmission rate	0.403	0.5227
B.1.1.7	Immune escape	0	0.1
B.1.351	Increase in transmission rate	0.1849	0.457
B.1.351	Immune escape	0.4414	0.8281
P.1	Increase in transmission rate	0.3682	0.4945
P.1	Immune escape	0.3588	0.6683
