

1 **Epidemiological characteristics of the B.1.526 SARS-CoV-2 variant**

2 Wan Yang, PhD,^{1*} Sharon K. Greene, PhD, MPH,² Eric R. Peterson, MPH,² Wenhui Li, PhD,³
3 Robert Mathes, MPH,² Laura Graf, MPH,² Ramona Lall, PhD,² Scott Hughes, PhD,⁴ Jade
4 Wang, MS,⁴ Anne Fine, MD^{2,#}

5 ¹Department of Epidemiology, Mailman School of Public Health, Columbia University; ²Bureau
6 of Communicable Disease, New York City Department of Health and Mental Hygiene; ³Bureau
7 of Vital Statistics, New York City Department of Health and Mental Hygiene; ⁴Public Health
8 Laboratory, New York City Department of Health and Mental Hygiene

9 *correspondence to: wy2202@cumc.columbia.edu

10 #current affiliation: Council of State and Territorial Epidemiologists

11 **Abstract**

12 **Background:** The B.1.526 variant (WHO designation: Iota) is a SARS-CoV-2 variant of interest, as
13 classified by both the US CDC and the WHO. Due to a lack of extensive genomic sequencing and
14 contact tracing data, its key epidemiological properties have not been well characterized.
15

16 **Methods:** We utilized nine epidemiological and population datasets collected in New York City
17 (NYC), where B.1.526 emerged, and comprehensive modeling to estimate the changes in
18 transmissibility, immune escape potential, and infection fatality risk (IFR) for B.1.526.
19

20 **Findings:** Estimated transmission rate in the neighborhood where B.1.526 was initially detected
21 was consistently higher than other neighborhoods in NYC and further increased during the
22 weeks preceding B.1.526 detection, likely due to its early spread there. Overall, models
23 estimated that B.1.526 had transmissibility about 15-25% higher than previously circulating
24 variants and that it could escape immunity in 0-10% of previously infected persons. In addition,
25 B.1.526 substantially increased IFR in older adults: by 46% (95% CI: 7.4 – 84%) among 45-64
26 year-olds, 82% (95% CI: 20 – 140%) among 65-74 year-olds, and 62% (95% CI: 45 – 80%) among
27 75+ during Nov 2020 – Apr 2021, compared to baseline IFR estimated for preexisting variants.
28

29 **Interpretation:** New variants like B.1.526 likely spread in the population weeks prior to
30 detection, and partial immune escape (e.g., resistance to therapeutic antibodies) could offset
31 prior medical advances and increase IFR. Early preparedness for and close monitoring of SARS-
32 CoV-2 variants, their epidemiological characteristics, and disease severity are thus crucial to
33 COVID-19 pandemic response as it remains a global public health threat.
34

35 **Key words:** COVID-19; variant of interest; B.1.526; transmissibility; immune escape; infection
36 fatality risk
37

39 Research in context

40 **Evidence before this study**

41 We searched PubMed for studies published through June 23, 2021 on the B.1.526 (Iota) SARS-
42 CoV-2 variant, using the terms “B.1.526”, “Iota variant” and “COVID-19”, or “Iota variant” and
43 “SARS-CoV-2”. Our search returned 14 papers, from which we read the abstracts and identified
44 8 relevant studies. The majority (7 of 8) of these studies focused on identifying the specific
45 mutations, namely L5F, T95I, D253G, additionally E484K or S477N for two different subclades
46 (i.e., B.1.526-S:E484K and B.1.526-S:S477N) and a closely related subclade (i.e., B.1.526-S:L452R
47 or B.1.526.1) as defined by spike mutations D80G, Δ144, F157S, L452R, T859N, and D950H;
48 these studies also examined changes in neutralizing ability of convalescent plasma and vaccinee
49 sera (mild reductions by ~2-4 fold were reported) and changes in neutralizing ability of
50 monoclonal antibodies (substantial reductions for some monoclonal antibodies were reported).
51 Two studies used genomic data from GISAID to estimate the growth rate of B.1.526 during
52 January - March 2021 and reported that it was similar to the Alpha variant (i.e., B.1.1.7), a SARS-
53 CoV-2 variant of concern with around 50% higher transmissibility than preexisting variants.
54 However, an epidemiological study using individual patient data from January 1 to April 5, 2021
55 reported no increases in secondary attack rate among contacts of B.1.526 cases, compared to
56 non-VOC/VOI variants. This latter study also showed preliminary evidence that B.1.526 did not
57 increase the risk of hospitalization or death, nor of vaccine breakthrough or reinfection.

58
59 **Added value of this study**

60 We utilized multiple datasets collected in New York City – where B.1.526 emerged – and model-
61 inference methods to reconstruct the transmission dynamics, including the initial emergence of
62 B.1.526 in the fall of 2020. We estimated that B.1.526 had a moderate increase (15-25%) in
63 transmissibility and could escape immunity in 0-10% of previously infected individuals. In
64 addition, B.1.526 substantially increased the infection-fatality risk (IFR) among adults 65 or
65 older by >60% during Nov 2020 – Apr 2021, compared to baseline risk estimated for preexisting
66 variants.

67
68 **Implications of all the available evidence**

69 New SARS-CoV-2 variants like B.1.526 could increase transmissibility, evade prior immunity, and
70 increase disease severity. Early preparedness for and close monitoring of SARS-CoV-2 variants
71 and their epidemiological characteristics are crucial to COVID-19 pandemic response as it
72 remains a global public health threat.

73

74 INTRODUCTION

75 The SARS-CoV-2 virus spread quickly worldwide in early 2020, causing the COVID-19 pandemic.
76 As the virus spread, it also diversified, and multiple novel SARS-CoV-2 variants emerged in
77 different populations, producing both local and global waves of infection. Several variants have
78 been characterized as variants of concern (VOC) or of interest (VOI), based on evidence
79 regarding their ability to increase transmissibility, evade immunity conferred by either prior
80 infection or vaccination, or cause more severe disease. Accurately estimating the
81 epidemiological characteristics and impact of these variants is thus important for informing
82 public health response, such as monitoring effectiveness of vaccines and therapeutic
83 antibodies. More broadly, such findings can also provide insights into the long-term trajectory
84 of SARS-CoV-2 beyond the pandemic phase.

85
86 The B.1.526 variant (WHO designation: Iota),¹ a SARS-CoV-2 VOI, was identified during Nov
87 2020 and quickly became a predominant variant in the New York City (NYC) area.²⁻⁴ It has also
88 been detected in all 52 states/territories in the US and at least 27 other countries (GISAID data,⁵
89 as of 6/9/2021). An initial laboratory study² suggested that this variant is to some extent
90 resistant to two therapeutic monoclonal antibodies in clinical use and neutralization by
91 convalescent plasma and vaccinee sera. However, another study⁴ examined all sequenced
92 B.1.526 cases in NYC identified as of April 5, 2021 (n = 3,679) and showed preliminary evidence
93 that this variant did not increase risk for infection after vaccination or reinfection. Both studies
94 may be limited due to the small number of specimens available for analysis as well as delay in
95 observation and reporting. Given these discrepancies, here we utilize detailed population
96 epidemiological data collected since the beginning of the COVID-19 pandemic in NYC (March 1,
97 2020 – April 30, 2021) and multiple model-inference methods to estimate the transmissibility,
98 immune escape potential, and disease severity of B.1.526.

99

100 METHODS

101 Study design and data

102 This study included three interconnected modeling analyses, synthesizing nine epidemiological
103 and population datasets (Fig 1). The first analysis applied a network model-inference system to
104 construct underlying SARS-CoV-2 transmission dynamics in NYC, accounting for under-detection
105 of infection; it also enabled estimation of key population variables and parameters (e.g., the
106 infection rate including those not detected as cases and transmission rate). The second analysis
107 applied a city-level multi-variant, age-structured model to simulate and estimate the changes in
108 transmissibility and immune escape potential for B.1.526 based on the network model-
109 inference estimates and additional data (e.g., variant prevalence data). The last analysis utilized
110 estimates from the first two model systems to estimate variant-specific infection fatality risk
111 (IFR, i.e. the fraction of all persons with SARS-CoV-2 infection who died from the disease), for

112 B.1.526 and B.1.1.7, separately. Of note, here we refer to the combination of B.1.526-S:E484K
113 and B.1.526-S:S477N as B.1.526; we did not include B.1.526.1 (i.e., B.1.526-S:L452) based on
114 earlier lineage classification and available data showing lower prevalence of B.1.526.1.

115
116 For the network model-inference system, we utilized multiple sources of epidemiological data,
117 including confirmed and probable COVID-19 cases, emergency department (ED) visits, and
118 deaths, as well as vaccination data. As done previously,⁶ we aggregated all COVID-19 confirmed
119 and probable cases^{7,8} and deaths⁸ reported to the NYC Department of Health and Mental
120 Hygiene (DOHMH) by age group (<1, 1-4, 5-14, 15-24, 25-44, 45-64, 65-74, and 75+ year-olds),
121 neighborhood of residence (42 United Hospital Fund neighborhoods in NYC⁹) and week of
122 occurrence (i.e., week of diagnosis for cases or week of death for decedents). COVID-19-related
123 ED visit data were obtained from the NYC syndromic surveillance system, comprised of all 53
124 hospital EDs in the city.¹⁰ This system identified individuals presenting at the EDs with COVID-
125 like-illness (CLI; defined as having a fever and cough or sore throat or respiratory illness, or
126 pneumonia, or a COVID-19 discharge diagnosis code, excluding those with a discharge diagnosis
127 code of influenza only); in addition, CLI patients were matched to electronic laboratory reports
128 of SARS-CoV-2 tests with diagnosis date within ± 7 days of ED visit. We estimated the number of
129 COVID-19-related ED visits as the number classified as CLI multiplied by the percentage of those
130 who tested positive for SARS-CoV-2 RNA, stratified by the same age and neighborhood groups
131 in weekly intervals. To account for the impact of vaccination, we also included COVID-19
132 vaccination data (partially and fully vaccinated, separately), aggregated to the same
133 age/neighborhood strata by week.

134
135 In addition, as in our previous study,⁶ we used mobility data from SafeGraph¹¹ to model
136 changes in SARS-CoV-2 transmission rate due to non-pharmaceutical interventions. These data
137 were aggregated to the neighborhood level by week without age stratification.

138
139 For the multi-variant model analysis, we additionally utilized four city-level, weekly datasets: 1)
140 COVID-19 confirmed and probable cases, 2) hospitalizations,¹² 3) deaths, and 4) the percentage
141 of different variants in NYC based on genomic sequencing of samples submitted to the NYC
142 DOHMH Public Health Laboratory and Pandemic Response Laboratory.^{4,13} The additional
143 hospitalization and variant percentage data were published by the NYC DOHMH^{12,13} and
144 accessed on June 22, 2021. We used the variant data from the week starting Jan 31, 2021 to the
145 week starting April 25, 2021 in this analysis, because earlier weeks had very low sample sizes
146 (<200 samples sequenced per week).¹³

147

148 This study was classified as public health surveillance and exempt from ethical review and
 149 informed consent by the Institutional Review Boards of both Columbia University and NYC
 150 DOHMH.

151

152 **Network model-inference system**

153 The network model-inference system used here is similar to the approach described in Yang et
 154 al;⁶ however, here we further accounted for waning immunity and vaccination and additionally
 155 used COVID-19-related ED visit data for model optimization. Briefly, the model-inference
 156 system uses an epidemic model (Eqn 1) to simulate the transmission of SARS-CoV-2 by age
 157 group and neighborhood, under implemented public health interventions and mass-vaccination
 158 when vaccines became available starting Dec 14, 2020:

159

$$\left\{ \begin{array}{l} \frac{dS_i}{dt} = \frac{R_i}{L} - \left(S_i \sum_{j=1}^{j=42} \frac{b_s b_j \beta_{city} m_{ij} I_j}{N_j} \right) - v_{i,1} - v_{i,2} \\ \frac{dE_i}{dt} = \left(S_i \sum_{j=1}^{j=42} \frac{b_s b_j \beta_{city} m_{ij} I_j}{N_j} \right) - \frac{E_i}{Z} \\ \frac{dI_i}{dt} = \frac{E_i}{Z} - \frac{I_i}{D} \\ \frac{dR_i}{dt} = \frac{I_i}{D} + v_{i,1} + v_{i,2} \end{array} \right. \quad [1]$$

160

161 where S_i , E_i , I_i , R_i , and N_i are the number of susceptible, exposed (but not yet infectious),
 162 infectious, and removed (either recovered or deceased) individuals and the total population,
 163 respectively, from a given age group in neighborhood- i . β_{city} is the average citywide
 164 transmission rate; b_s is the estimated seasonal trend.⁶ The term b_i represents the
 165 neighborhood-level transmission rate relative to the city average. The term m_{ij} represents the
 166 changes in contact rate in each neighborhood (for $i=j$) or spatial transmission from
 167 neighborhood- j to i (for $i \neq j$) and was computed based on the mobility data.⁶ Z , D , and L are the
 168 latency period, infectious period, and immunity period, respectively. The term $v_{i,1}$ represents
 169 the number of individuals in neighborhood- i successfully immunized after the first dose of the
 170 vaccine and is computed using vaccination data and vaccine efficacy (VE) for 1st dose; $v_{i,2}$ is the
 171 additional number of individuals successfully immunized after the second vaccine dose
 172 (excluding those successfully immunized after the first dose). Because 97% of vaccine doses
 173 administered in NYC during our study period (through April 30, 2021) were the Pfizer-BioNTech
 174 or Moderna vaccines, we assumed a VE of 85% fourteen days after the first dose and 95% seven
 175 days after the second dose based on clinical trials and real-world data.¹⁴⁻¹⁶

176

177 Using the model-simulated number of infections occurring each day, we further computed the
 178 number of cases, ED visits, and deaths each week to match with the observations.⁶ Similar to
 179 the procedure for cases and deaths described in Yang et al.,⁶ to compute the number of ED
 180 visits, we multiplied the model-simulated number of new infections per day by the ED-
 181 consultation rate (i.e. the fraction of model-simulated persons with new SARS-CoV-2 infections
 182 presenting at the EDs), and further distribute these estimates in time per a distribution of time-
 183 from-infection-to-ED-consultation (Table S1); we then aggregated the daily lagged, simulated
 184 estimates to weekly totals for model inference.

185
 186 Each week, the system uses the ensemble adjustment Kalman filter (EAKF)¹⁷ to compute the
 187 posterior estimates of model state variables and parameters based on the model (prior)
 188 estimates and observed case, ED visit, and mortality data per Bayes' rule.⁶ In particular, using
 189 this model-inference, we estimated the citywide transmission rate (β_{city}), neighborhood
 190 relative transmission rate (b_i), and IFR by age group for each week, from the week starting
 191 March 1, 2020 (i.e. the beginning of the COVID-19 pandemic in NYC) to the week starting April
 192 25, 2021.

193
 194 **Multi-variant, age-structured model**

195 Due to model complexity, the model-inference system described above does not account for
 196 the circulation of different variants. To model variants, we used a city-level multi-variant, age-
 197 structured model,¹⁸ per Eqn 2:

198

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

199
 200 Model variables and parameters in Eqn 2 are similar to those in Eqn 1 with the same symbols.
 201 For instance, β_i is the transmission rate for variant- i . However, instead of modeling the spatial
 202 structure, Eqn 2 focuses on the interactions among different variants (indicated by the
 203 subscript, i) and age structure (indicated by the superscript, a or A). Specifically, c_{ij} measures
 204 the strength of cross-immunity to variant- i conferred by infection of variant- j (e.g., close to 0 if
 205 it is weak and $c_{ii}=1$ for infection by the same variant). The vaccination model component $v_{i,1}^A$

206 and $v_{i,2}^A$ are also variant-specific and can additionally account for the reduction in VE against the
207 new variants if needed; however, here we used the same VE estimates for all variants included
208 (i.e., B.1.526, B.1.1.7, B.1.427 and B.1.429) based on observations.^{4,16,19,20} Additionally, the term
209 ϵ_i represents travel-related importation of infections of variant- i (see Table S2).

210

211 We restricted this simulation to Nov 2020 – Apr 2021 (i.e., from the initial identification of
212 B.1.526 to before the further detection and increase of other variants such as Gamma and
213 Delta). In addition to preexisting variants of SARS-CoV-2 prior to Nov 2020 (hereafter referred
214 to as “non-VOC/VOI variants” for simplicity) and B.1.526, the model included B.1.1.7 and
215 B.1.427/B.1.429 (combined for simplicity), based on available genomic surveillance data
216 showing consistent detection of these variants during the simulation period and very low levels
217 for others if detected. Model parameters for B.1.1.7 and B.1.427/B.1.429 were listed in Table
218 S2. For simplicity, we did not account for other VOCs/VOIs variants because their percentages
219 were very low during either analysis period.¹³

220

221 Initial analysis based on the model-inference estimates suggested B.1.526 was around 20%
222 more infectious than non-VOC/VOI variants, without accounting for changes in immunity due to
223 potential immune escape (see details in Results). Therefore, in this analysis, we tested
224 combinations of change in transmissibility ranging from 10 – 30% increases and immune escape
225 ranging from 0 – 30%, both with a 5% increment and $\pm 5\%$ intervals (35 combinations in total).
226 For instance, for the combination centering at 10% transmissibility increase and 0% immune
227 escape, the model is initialized using values in the range of 5-15% (i.e., $10 \pm 5\%$) transmissibility
228 increase and 0-5% (i.e., $0 \pm 5\%$ and setting negatives to 0) immune escape. In addition, due to
229 uncertainty on the initial prevalence, we tested three different levels of initial seeding for the
230 week starting Nov 1, 2020, i.e., low (0.5 – 2.5%), high (1.5 – 3.5%), and wider range (0.5 – 3.5%).
231 For reference, Washington Heights – Inwood (WHts), which is the neighborhood where the first
232 patients identified with B.1.526 resided and sought care, constituted 3.2% of the NYC
233 population in 2018. We initialized the model using the model-inference estimates (e.g.,
234 population susceptibility and transmission rates by age group; Table S2) and ran the model for
235 each parameter combination 10 times, each with 1000 realizations to account for model
236 stochasticity. Results are summarized from the 10,000 model realizations.

237

238 To identify the most plausible combination of transmissibility and immune escape properties
239 for B.1.526, we compared the model-estimated weekly number of cases, hospitalization, and
240 deaths as well as the percentage of the variants to available data. Evaluation was made based
241 on 1) accuracy, i.e., if the observation falls within the model-estimated interquartile range, it is
242 deemed accurate; 2) relative root-mean-square-error (RMSE) between the observed and the
243 model-estimated; and 3) Pearson correlation between the two time-series. Because results

244 show that model accuracy and relative RMSE had a wider spread among the combinations
245 tested (i.e., more distinctive), we first subset those having accuracy within the highest 25th
246 percentile and relative RMSE within the lowest 25th percentile (2-4 out of 35 combinations
247 remained for each setting of initial prevalence); we then selected the one with the highest
248 correlation in the subset as the best-performing and most plausible combination.

249

250 **Estimating the changes in IFR due to B.1.526**

251 The network model-inference system enables estimation of the IFR by age group over time.
252 These estimates are made combining all variants and do not distinguish by variant. However,
253 we reasoned that the combined IFR is a weighted average of individual, variant-specific
254 estimates given the relative prevalence of each variant. Accordingly, we built two linear
255 regression models to estimate the variant-specific IFR. Model 1 restricted the analysis to Nov
256 2020 – Jan 2021 (when the relative prevalence of B.1.1.7 in NYC was likely <10%; n = 14 weeks)
257 and only included two categories of variants:

$$IFR_{combined} \sim IFR_{B.1.526} i_{B.1.526} + IFR_{baseline} i_{others} \quad [3]$$

258

259 Model 2 extended the analysis to Nov 2020 – Apr 2021 (n = 26 weeks) and included both
260 B.1.526 and B.1.1.7, in addition to other variants:

$$IFR_{combined} \sim IFR_{B.1.526} i_{B.1.526} + IFR_{B.1.1.7} i_{B.1.1.7} + IFR_{baseline} i_{others} \quad [4]$$

261

262 In both models, $IFR_{combined}$ is the overall IFR for each week, estimated using the model-inference
263 system; $i_{B.1.526}$, $i_{B.1.1.7}$, i_{others} are the percentage of infection by the corresponding variant for each
264 week, estimated using the multi-variant age-structured model with the most plausible
265 parameter combination as data are not available. $IFR_{baseline}$ is the baseline IFR for the preexisting
266 variants, set to the average of model-inference estimates over the period of Oct – Nov 2020
267 (i.e., prior to the increase of the new variants). The variant-specific IFRs, $IFR_{B.1.526}$ and $IFR_{B.1.1.7}$,
268 are then estimated using the regression models (n = 14 weekly data points for Model 1; and n =
269 26 weekly data points for Model 2). For either model, the change in IFR due to a given variant is
270 then computed as:

$$\Delta IFR = \frac{IFR_{variant} - IFR_{baseline}}{IFR_{baseline}} \times 100\% \quad [5]$$

271

272 Both model analyses were performed for each age group or all ages combined, separately; we
273 also combined all those aged under 25 as the IFRs were similarly low for the four sub-age
274 groups (i.e. <1, 1-4, 5-14, and 15-24 year-olds).

275

276 **RESULTS**

277 **Epidemic dynamics of the second pandemic wave in NYC**

278 NYC experienced a very large first pandemic wave during spring 2020. Similar to our previous
279 work,⁶ the model-inference system here estimates that 16.6% of the population (95% CrI: 13.6
280 – 21.5%; or 1.1 – 1.8 million people) had been infected by the end of May 2020 (i.e., end of the
281 first wave; Fig 2D). The city was able to gradually reopen part of its economy during summer
282 2020 after a 3-month long stay-at-home mandate for all non-essential workers. However,
283 infection resurged beginning in the fall of 2020 and the city experienced a second pandemic
284 wave around Nov 2020 – April 2021 (Fig 2). Following the second wave, an estimated total of
285 41.7% (95% CrI: 35.4 – 49.3%; or 3.0 – 4.1 million people) had been infected by the end of Apr
286 2021, including all those infected during the first wave. Note these estimates accounted for
287 under-detection of infections (Fig S1), for which the overall infection-detection rate increased
288 to 37.1% (95% CrI: 33.3 – 43.0%) during the 2nd wave from 15.1% (95% CrI: 11.7 – 18.5%) during
289 the 1st wave. This large number of infections occurred despite the non-pharmaceutical
290 interventions implemented throughout the pandemic and rollout of mass-vaccination starting
291 mid-Dec 2020. In addition, unlike the first wave that predominantly affected older age groups,
292 the second pandemic wave affected all age groups (Figs S2-3).

293

294 **Transmission rate increased earliest in the neighborhood where B.1.526 was initially** 295 **identified**

296 The emergence and rapid increase of B.1.526 coincided with the second pandemic wave in NYC.
297 While first reported in Feb 2021,² testing initially identified the B.1.526 variant in patient
298 samples dated back to early Nov 2020 from the city's WHts neighborhood.² As such, we first
299 examined potential changes in the transmission rate there. Indeed, prior to the identification of
300 B.1.526, estimated neighborhood relative transmission rate (b_i in Eqn 1) in WHts gradually
301 increased, remained at high levels during Nov 2020 – Feb 2021, and decreased to the baseline
302 level afterwards when B.1.526 became a predominant variant citywide (~40% of all cases
303 sequenced by end of Feb 2021). In comparison, the estimates were relatively stable for other
304 neighborhoods (Fig 3A), suggesting the changes in WHts were likely due to the early spread of
305 B.1.526. Averaging over this period, we estimate that the relative transmission rate in WHts
306 increased by 8.4% (95% CI: -5.8 – 22.5%). Concurrently, the citywide transmission rate
307 increased by 13.3% (95% CI: -21.1 – 47.8%; Fig 3B). These two preliminary estimates in
308 combination suggest that the transmission rate of B.1.526 likely is 22.8% (95% CI: -12.4 –
309 58.0%) higher than preexisting non-VOC/VOI variants, without accounting for potential change
310 due to immune evasion.

311

312 **B.1.526 likely causes a moderate increase in transmissibility (15-25%) and slight immune** 313 **evasion (0-10%)**

314 We further examine model estimations under a wide range of transmissibility and immune
315 escape settings for B.1.526. Under all three possible scenarios of initial prevalence (i.e., 0.5 –

316 2.5%, 1.5 – 3.5%, and 0.5 – 3.5%), model simulations consistently show that B.1.526 likely
317 increases transmissibility by 15-30% and can escape immunity in 0-10% of previously infected
318 persons (Fig 4 A-C). Overall, a higher initial prevalence (1.5 – 3.5% at the beginning of Nov 2020)
319 combining with a 15-25% increase in transmissibility and 0-10% immune escape (Fig 4 A-C,
320 middle column; and Fig 4D) generated the most accurate estimates of cases, hospitalizations
321 and deaths as well as variant percentages during the second wave. Model simulations show
322 that, with this moderate increase in transmissibility and small immune escape, B.1.526 was able
323 to outcompete preexisting variants and gradually increase its percentage from Nov 2020 to
324 March 2021; however, afterwards its percentage decreased with the surge of B.1.1.7, a more
325 infectious variant (Fig 4D, bottom right panel).

326

327 **B.1.526 likely increases disease severity substantially**

328 During the second wave, estimated IFR increased gradually in later months, particularly among
329 older age groups, despite the decline in deaths following mass-vaccination (Fig 5). Modeling
330 accounting for infections and deaths due to B.1.526, B.1.1.7, and non-VOC/VOI variants
331 suggests that B.1.526 increased IFR in older adults: by 46% (95% CI: 7.4 – 84%) among 45-64
332 year-olds [absolute IFR: 0.42% (95% CI: 0.31 – 0.54%) vs. 0.29% (95% CI: 0.15 – 0.44%) baseline
333 risk]; 82% (95% CI: 20 – 140%) among 65-74 year-olds [absolute IFR: 1.9% (95% CI: 1.2 – 2.5%)
334 vs. 1.0% (95% CI: 0.57 – 2.5%) baseline risk], and 62% (95% CI: 45 – 80%) among 75+ [absolute
335 IFR: 6.7% (95% CI: 5.9 – 7.4%) vs. 4.1% (95% CI: 2.2 – 6.3%) baseline risk], during Nov 2020 – Apr
336 2021; overall, B.1.526 increased the IFR by 60% (95% CI: 38 – 82%), compared to estimated
337 baseline risk (Table 1). The analysis restricting to Nov 2020 – Jan 2021 suggests similar IFR
338 increases (Table S3). These estimated IFR increases were lower than for B.1.1.7 but
339 comparable. Of note, the IFRs for B.1.1.7 estimated here were higher than but in line with those
340 reported in the UK [e.g., overall increase: 100% (75-130%) vs. 61% (42–82%) in the UK²¹].

341

342 **DISCUSSION**

343 The B.1.526 variant is one of the SARS-CoV-2 variants designated as a VOI by both the WHO¹
344 and the US CDC.¹⁸ However, due to a lack of extensive genomic sequencing and contact tracing
345 data particularly during the early phase of its emergence, its key epidemiological properties
346 have not been well characterized. Utilizing multiple epidemiological datasets and
347 comprehensive modeling, here we have estimated the changes in transmissibility, immune
348 escape potential, and disease severity for B.1.526. Results suggest that, compared to
349 preexisting non-VOC/VOI variants, B.1.526 causes a moderate increase in transmissibility and
350 minimal immune evasion; however, it might substantially increase IFR in older adults. As such,
351 continued monitoring of the circulation of this variant is warranted.

352

353 Our study offers several lessons for future outbreak response. First, prior to the emergence of
354 B.1.526, the estimated transmission rate in WHts, where it likely emerged, was consistently
355 higher than other neighborhoods in NYC throughout the pandemic. Population characteristics
356 (e.g., household structure) that may contribute to this higher transmission rate need further
357 investigation; however, the higher transmission rate may have facilitated the spread of new
358 mutants between hosts and its emergence population-wide. It is thus important to closely
359 monitor populations with sustained higher transmission rates for new variants, particularly in
360 areas lacking robust and timely sequencing of samples from newly identified cases. In addition,
361 the estimated transmission rate in WHts further increased in conjunction with the emergence
362 of B.1.526; such changes thus may serve as an early indicator for in-depth epidemiological
363 investigation (e.g., to assess changes in circulating variants and transmissibility). A similar
364 approach has been applied in the UK, where subregions with higher estimated growth rates
365 were prospectively investigated, leading to identification of B.1.1.7 as a VOC.²²⁻²⁵

366
367 Second, we did not find a higher B.1.526-related IFR among younger age groups (those under
368 45 years); this finding is consistent with the findings of Thompson et al.⁴ based on analysis of all
369 sequenced cases, the majority of whom (67%) were under 45 years. However, for older ages,
370 we found substantially higher B.1.526-related IFRs (e.g. >60% higher for those above 65 years).
371 This latter finding appears to be consistent with the report by Annavajhala et al.² showing
372 resistance of B.1.526 to therapeutic antibodies. Over the course of the pandemic, SARS-CoV-2
373 IFR has decreased substantially (about a 3-fold difference between the two pandemic waves),
374 likely due to improved medical treatments (e.g., therapeutic antibodies), better patient
375 management, and earlier diagnosis. As older adults are more likely to suffer from severe
376 COVID-19 and thus receive therapeutic antibodies,^{26,27} the resistance of B.1.526 may render
377 these treatments ineffective despite their prior success against other variants, leading to
378 increases in IFR among older adults. These findings highlight the importance of monitoring the
379 efficacy of therapeutics against different variants and timely update of treatments. In addition,
380 a better understanding of factors contributing to the higher IFRs in certain variants is warranted
381 to inform countermeasures.^{28,29}

382
383 Lastly, our analyses suggest both B.1.526 and B.1.1.7 likely had been spreading in the
384 population for weeks or months prior to detection by the surveillance system.^{2,3,30} Expanding
385 genomic sequencing programs for SARS-CoV-2 and improving linkage to epidemiologic data can
386 improve detection of new VOIs/VOCs. Such efforts are underway (e.g., in the US) but more
387 efforts and resources are urgently needed globally. In addition, to support more timely
388 detection and control, targeted screening of key subpopulations (e.g., those prospectively
389 identified from modeling as having high transmission rates) and viral traits (e.g., mutations
390 linked to increased transmissibility and/or immune evasion as done in Annavajhala et al.²) is

391 needed, as well as timely sharing of key information globally. The documentation of new
392 VOIs/VOCs anywhere in the world then should prompt preparedness measures to detect and
393 rapidly respond to the introduction of those variants into local areas. More fundamentally, to
394 limit emergence of new VOIs/VOCs and end the COVID-19 pandemic, all populations worldwide
395 should have timely access to vaccination, and multiple layers of mitigation efforts are needed
396 until a sufficient portion of the population is protected by vaccination.

397

398 Our study also has several limitations. First, most of our analyses are based on population-level
399 data without variant-specific information, given limited variant testing during most of the study
400 period. We circumvented this data deficiency by analyzing estimates of a key subpopulation
401 (e.g., the WHts neighborhood where B.1.526 was initially detected) and leveraging prior
402 knowledge (e.g., estimated IFR prior to B.1.526 emergence). Second, our study did not
403 distinguish the two subclades within the B.1.526 lineage – one containing the E484K mutation
404 and the other containing the S477N mutation. Both the E484K and S477N mutations have been
405 shown to mediate immune escape;^{29,31-33} in addition, the percentages of these two subclades
406 were similar during our study period, suggesting they likely have similar epidemiological
407 characteristics. Lastly, there is a likely larger uncertainty in B.1.526-related and B.1.1.7-related
408 IFR estimates for younger ages (those under 45), due to the smaller number of deaths and
409 larger uncertainty in baseline IFR estimates. Future investigation addressing these issues is
410 warranted should a large sample of variant-specific data become available.

411

412 In summary, our study has reconstructed the early epidemic trajectory and subsequent rise of
413 B.1.526 in NYC and estimated its key epidemiological properties. Findings highlight the
414 importance of monitoring the viral diversity of SARS-CoV-2, epidemiological characteristics of
415 new variants, and disease severity, as COVID-19 remains a global public health threat.

416

417 **Acknowledgments:**

418 This study was supported by the National Institute of Allergy and Infectious Diseases (AI145883)
419 and the NYC DOHMH. We thank Columbia University Mailman School of Public Health for high
420 performance computing, Safe Graph (safegraph.com) for providing the mobility data, and
421 Sasikiran Kandula at Columbia University for compiling the mobility data used in this study. We
422 thank the NYC DOHMH Incident Command System Surveillance and Epidemiology Section for
423 processing, cleaning, and managing COVID-19 surveillance data, the NYC DOHMH Public Health
424 Laboratory and Pandemic Response Laboratory for generating and analyzing sequence data,
425 and Iris Cheng, Mohammed Almashhadani, Charles Ko, and Jaimie Shaff from the NYC DOHMH
426 for providing the vaccination data and helpful suggestions on the manuscript.

427

428 **Author contributions:**

429 WY designed the study, conducted the analysis, and wrote the first draft; SKG contributed to
430 study coordination and specification of COVID-19 case data, and provided input on parameter
431 estimation; ERP led aggregation and provision of COVID-19 case data; LG contributed to
432 management of COVID-19 case data; WL provided the COVID-19-associated mortality data; RM
433 and RL oversaw the collection of and provided the COVID-19 ED data; SH and JW provided input
434 on SARS-CoV-2 variants and interpretation of the NYC variant percentage data; AF oversaw data
435 collection and management processes at DOHMH. All authors contributed to the final draft.

436

437 **Conflict of Interest**

438 The authors declare no conflict of interest.

439

440 **References**

- 441 1. World Health Organization. Tracking SARS-CoV-2 variants. 2021.
442 <https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/>.
- 443 2. Annavajhala MK, Mohri H, Wang P, et al. A Novel and Expanding SARS-CoV-2 Variant,
444 B.1.526, Identified in New York. *medRxiv* 2021: 2021.02.23.21252259.
- 445 3. West AP, Wertheim JO, Wang JC, et al. Detection and characterization of the SARS-CoV-2
446 lineage B.1.526 in New York. *bioRxiv* 2021: 2021.02.14.431043.
- 447 4. Thompson CN, Hughes S, Ngai S, et al. Rapid Emergence and Epidemiologic Characteristics
448 of the SARS-CoV-2 B.1.526 Variant - New York City, New York, January 1-April 5, 2021.
449 *MMWR Morbidity and mortality weekly report* 2021; **70**(19): 712-6.
- 450 5. Global Initiative on Sharing All Influenza Data (GISAID). GISAID. 2021.
451 <https://www.gisaid.org> (accessed 6/9/2021 2021).
- 452 6. Yang W, Kandula S, Huynh M, et al. Estimating the infection-fatality risk of SARS-CoV-2 in
453 New York City during the spring 2020 pandemic wave: a model-based analysis. *The Lancet*
454 *Infectious diseases* 2021; **21**(2): 203-12.
- 455 7. Centers for Disease Control and Prevention. National Notifiable Diseases Surveillance
456 System (NNDSS) - Coronavirus Disease 2019 (COVID-19). 4/16/2021 2021.
457 <https://ndc.services.cdc.gov/conditions/coronavirus-disease-2019-covid-19/> (accessed
458 6/22/21 2021).
- 459 8. New York City Department of Health and Mental Hygiene. Defining confirmed and probable
460 cases and deaths. 2020. <https://www1.nyc.gov/site/doh/covid/covid-19-data.page>
461 (accessed 5/27/2021 2021).
- 462 9. New York City Department of Health and Mental Hygiene. NYC UHF 42 Neighborhoods.
463 <http://a816-dohbsp.nyc.gov/IndicatorPublic/EPHTPDF/uhf42.pdf>.
- 464 10. Lall R, Abdelnabi J, Ngai S, et al. Advancing the Use of Emergency Department Syndromic
465 Surveillance Data, New York City, 2012-2016. *Public Health Rep* 2017; **132**: 23s-30s.

- 466 11. SafeGraph. Weekly Patterns: Foot Traffic Data To Understand The COVID-19 Pandemic.
467 2020. <https://www.safegraph.com/weekly-foot-traffic-patterns>.
- 468 12. New York City Department of Health and Mental Hygiene. NYC Coronavirus Disease 2019
469 (COVID-19) Data. <https://github.com/nychealth/coronavirus-data>.
- 470 13. New York City Department of Health and Mental Hygiene. Variants. 2021.
471 <https://github.com/nychealth/coronavirus-data/tree/master/variants> (accessed 6/22/21
472 2021).
- 473 14. Polack FP, Thomas SJ, Kitchin N, et al. Safety and Efficacy of the BNT162b2 mRNA Covid-19
474 Vaccine. *New Engl J Med* 2020.
- 475 15. Baden LR, El Sahly HM, Essink B, et al. Efficacy and Safety of the mRNA-1273 SARS-CoV-2
476 Vaccine. *N Engl J Med* 2021; **384**(5): 403-16.
- 477 16. Haas EJ, Angulo FJ, McLaughlin JM, et al. Impact and effectiveness of mRNA BNT162b2
478 vaccine against SARS-CoV-2 infections and COVID-19 cases, hospitalisations, and deaths
479 following a nationwide vaccination campaign in Israel: an observational study using national
480 surveillance data. *The Lancet* 2021; **397**(10287): 1819-29.
- 481 17. Anderson JL. An Ensemble Adjustment Kalman Filter for Data Assimilation. *Mon Weather*
482 *Rev* 2001; **129**(12): 2884-903.
- 483 18. Yang W, Shaman J. Epidemiological characteristics of three SARS-CoV-2 variants of concern
484 and implications for future COVID-19 pandemic outcomes. *medRxiv* 2021:
485 2021.05.19.21257476.
- 486 19. Abu-Raddad LJ, Chemaitelly H, Butt AA. Effectiveness of the BNT162b2 Covid-19 Vaccine
487 against the B.1.1.7 and B.1.351 Variants. *New Engl J Med* 2021.
- 488 20. Garcia-Beltran WF, Lam EC, St Denis K, et al. Multiple SARS-CoV-2 variants escape
489 neutralization by vaccine-induced humoral immunity. *Cell* 2021; **184**(9): 2372-83 e9.
- 490 21. Davies NG, Jarvis CI, Group CC-W, et al. Increased mortality in community-tested cases of
491 SARS-CoV-2 lineage B.1.1.7. *Nature* 2021; **593**(7858): 270-4.
- 492 22. New and Emerging Respiratory Virus Threats Advisory Group. Minutes of the NERVTAG
493 COVID-19 Fortieth Meeting: 11 December 2020. 2020.
494 <https://app.box.com/s/3lkcbxepqixkg4mv640dpvvg978ixjtf/file/789097594594> (accessed
495 7/27/21 2021).
- 496 23. New and Emerging Respiratory Virus Threats Advisory Group. Minutes of the extraordinary
497 meeting of NERVTAG COVID-19 and SPI-M on SARS-CoV-2 variants: 21 December 2020.
498 2020. <https://app.box.com/s/3lkcbxepqixkg4mv640dpvvg978ixjtf/file/789098622117>
499 (accessed 7/27/21 2021).
- 500 24. New and Emerging Respiratory Virus Threats Advisory Group. NERVTAG/SPI-M
501 Extraordinary meeting on SARS-CoV-2 variant of concern 202012/01 (variant B.1.1.7). 2020.
502 <https://app.box.com/s/3lkcbxepqixkg4mv640dpvvg978ixjtf/file/756964987830> (accessed
503 7/27/21 2021).

- 504 25. Volz E, Mishra S, Chand M, et al. Transmission of SARS-CoV-2 Lineage B.1.1.7 in England:
505 Insights from linking epidemiological and genetic data. *medRxiv* 2021:
506 2020.12.30.20249034.
- 507 26. National Institutes of Health. COVID-19 Treatment Guidelines: Anti-SARS-CoV-2 Monoclonal
508 Antibodies. 5/24/2021. [https://www.covid19treatmentguidelines.nih.gov/therapies/anti-](https://www.covid19treatmentguidelines.nih.gov/therapies/anti-sars-cov-2-antibody-products/anti-sars-cov-2-monoclonal-antibodies/)
509 [sars-cov-2-antibody-products/anti-sars-cov-2-monoclonal-antibodies/](https://www.covid19treatmentguidelines.nih.gov/therapies/anti-sars-cov-2-antibody-products/anti-sars-cov-2-monoclonal-antibodies/) (accessed 7/6/2021).
- 510 27. Centers for Disease Control and Prevention. COVID-19: Treatment considerations for
511 healthcare providers. 2021. [https://www.cdc.gov/coronavirus/2019-ncov/variants/variant-](https://www.cdc.gov/coronavirus/2019-ncov/variants/variant-info.html)
512 [info.html](https://www.cdc.gov/coronavirus/2019-ncov/variants/variant-info.html).
- 513 28. Starr TN, Greaney AJ, Addetia A, et al. Prospective mapping of viral mutations that escape
514 antibodies used to treat COVID-19. *Science* 2021; **371**(6531): 850-+.
- 515 29. Greaney AJ, Starr TN, Gilchuk P, et al. Complete Mapping of Mutations to the SARS-CoV-2
516 Spike Receptor-Binding Domain that Escape Antibody Recognition. *Cell Host & Microbe*
517 2021; **29**(1): 44-57.e9.
- 518 30. Washington NL, Gangavarapu K, Zeller M, et al. Emergence and rapid transmission of SARS-
519 CoV-2 B.1.1.7 in the United States. *Cell* 2021; **184**(10): 2587-+.
- 520 31. Greaney AJ, Loes AN, Crawford KHD, et al. Comprehensive mapping of mutations in the
521 SARS-CoV-2 receptor-binding domain that affect recognition by polyclonal human plasma
522 antibodies. *Cell Host & Microbe* 2021; **29**(3): 463-76.e6.
- 523 32. Liu Z, VanBlargan LA, Bloyet L-M, et al. Identification of SARS-CoV-2 spike mutations that
524 attenuate monoclonal and serum antibody neutralization. *Cell host & microbe* 2021; **29**(3):
525 477-88.e4.
- 526 33. Harvey WT, Carabelli AM, Jackson B, et al. SARS-CoV-2 variants, spike mutations and
527 immune escape. *Nature Reviews Microbiology* 2021; **19**(7): 409-24.
- 528

529 **Table and Figure Captions**

530 **Table 1.** Estimated IFR for different variants and changes compared to the baseline risk
531 estimated for preexisting variants during Oct – Dec 2020, using Eqn 4.

532
533 **Fig 1.** Study design. This study included three modeling analyses: 1) spatial network model-
534 inference to construct the transmission dynamics and estimate key population variables and
535 parameters by United Hospital Fund (UHF) neighborhood of residence and age group; 2) city-
536 level multi-variant, age-structured modeling to simulate and estimate the changes in
537 transmissibility and immune escape potential for B.1.526; and 3) linear regression models to
538 estimate variant-specific infection fatality risk (IFR), for B.1.526 and B.1.1.7, separately. Nine
539 datasets (listed in the black open boxes) were used as model inputs or to evaluate the accuracy
540 of model estimates (indicated for each dataset below). Models used are shown in the blue filled
541 boxes and model outputs are listed in the blue open boxes (key estimates reported in detail in
542 the Results are bolded). Connections among the analyses are indicated by the arrows and
543 associated annotations.

544
545 **Fig 2.** Model fit and key estimates. Upper panel shows model-fit to weekly number of cases (A),
546 ED visits (B), and deaths (C), for all ages combined. Lower panel shows key model-inference
547 estimates of weekly number of infections including those not detected as cases (D), cumulative
548 number of infections in NYC overall (E), and cumulative infection rate by neighborhood (F).
549 Boxes show model estimates (thick horizontal lines and box edges show the median, 25th, and
550 75th percentiles; vertical lines extending from each box show 95% CrI) and red dots show
551 corresponding. For the weekly estimates, week starts (mm/dd/yy) are shown in the x-axis
552 labels. Star (*) in the map indicates the location of the Washington Heights – Inwood
553 neighborhood.

554
555 **Fig 3.** Changes in transmission rate. (A) Changes in neighborhood-level relative transmission
556 rate. (B) Changes in citywide transmission rate. Vertical dashed lines indicate the earliest date
557 B.1.526 was identified as reported in Annavajhala et al. Labels of x-axis show the week starts
558 (mm/dd/yy).

559
560 **Fig 4.** Comparison of different combinations of changes in transmissibility and immune escape
561 property for B.1.526. Left panel shows the overall accuracy (A), relative RMSE (B), and
562 correlation (C) of model estimates under different transmissibility and immune escape settings.
563 White crosses (x) indicate the best-performing parameter combination. Right panel shows
564 model estimates using the overall best-performing parameter combination (i.e., 1.5-3.5% initial
565 prevalence, 15-25% higher transmissibility, and 0-10% immune escape). Lines and surrounding
566 areas show model-simulated median estimates and interquartile range; dots show

567 corresponding observations; colors indicate different variants as specified in the legend. Note
568 that these model simulations used same infection-detection rate, hospitalization-rate and IFR
569 (i.e., average during Nov 2020 – Apr 2021); that is, they did not account for changes in case
570 ascertainment or disease severity by week during this period, due to, e.g., increases in disease
571 severity by the new variants. As such, there were larger deviations from the observations
572 during later months of the simulation with more infections by the new variants.

573

574 **Fig 5.** Estimated infection fatality risk. Red lines show the estimated median IFR with
575 surrounding areas indicating the 50% (darker color) and 95% (lighter color) CrI. For comparison,
576 the grey bars show the number of deaths reported for each week from the week of Oct 4, 2020
577 to Apr 25, 2021. X-axis labels show the week starts (mm/dd/yy).

578

579

580 **Supplemental Table and Figure Captions**

581 **Table S1.** Prior ranges for the network model-inference system. The prior ranges are similar to
582 Table S1 of Yang et al.¹ but include additional parameters in Eqn 1. The spatial, temporal, and
583 age resolution of each parameter or variable, estimated in the model-inference system, is
584 specified in the column "Resolution". Note posterior parameter estimates can extend outside
585 the specified prior ranges.

586

587 **Table S2.** Initial conditions used to simulate co-circulation of different variants in the multi-
588 variant, age-structured model. To partially account for changing infection-detection rate, ED-
589 consultation rate (EDR) and IFR, for these three parameters, we used the model-inference
590 estimates averaged over the entire simulation period (i.e. Nov 2020 – April 2021). For the initial
591 transmission rate (for the preexisting non-VOC/VOI variants), we used the model-inference
592 estimates averaged over the week of 10/25/2020 – the week of 11/7/2020 (i.e. the 3 weeks
593 around the start of simulation). For the rest of model state variables and parameters, we used
594 model-inference estimates made at the week of 10/25/2020. For B.1.1.7, we used the
595 following ranges based on estimates from Yang and Shaman¹⁰: 40.3 – 52.3% higher
596 transmissibility (related to estimates for the preexisting non-VOC/VOI variants listed below) and
597 0 – 10% immune escape; for comparison, contact tracing data from the UK showed that B.1.1.7
598 was 30-50% more infectious.¹¹ For B.1.427/ B.1.429, we used the following ranges based on
599 estimates from Deng et al.¹²: 16 – 24% higher transmissibility and 0-10% immune escape (vs.
600 21.4 – 27.8% increase in transmission rate in Deng et al.¹² without accounting for changes in
601 immunity due to potential immune escape).

602

603 **Table S3.** Estimated IFR for different variants and changes compared to the baseline risk
604 estimated for preexisting variants during Oct – Dec 2020, using Eqn 3.

605

606 **Fig S1.** Estimated infection-detection rate by age group. Red lines show the estimated median
607 infection-detection rate with surrounding areas indicating the 50% (darker color) and 95%
608 (lighter color) CrI. For comparison, the grey bars show the number of cases reported for each
609 week from the week of Oct 4, 2020 to Apr 25, 2021. Labels of x-axis show the week starts
610 (mm/dd/yy).

611

612 **Fig S2.** Model-fit by age group. Boxes show model estimates (thick horizontal lines and box
613 edges show the median, 25th, and 75th percentiles; vertical lines extending from each box show
614 95% CrI) and red dots show corresponding.

615

616 **Fig S3.** Estimated cumulative infection rates by age group. Thick horizontal lines and box edges
617 show the median, 25th, and 75th percentiles; vertical lines extending from each box show 95%
618 CrI.

619

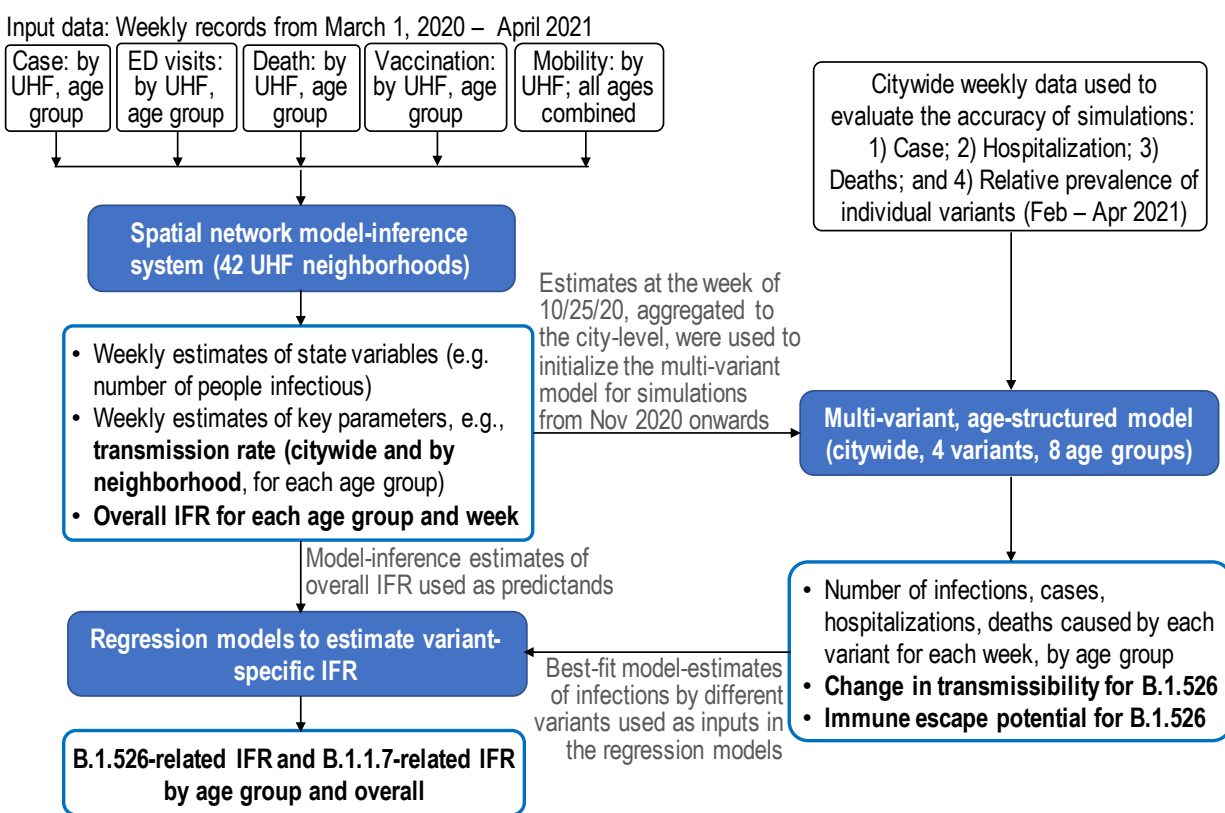
Table and Figures

620 **Table 1.** Estimated IFR for different variants and changes compared to the baseline risk estimated for preexisting variants during Oct
 621 – Dec 2020, using Eqn 4.

Age	IFR, baseline (%)	IFR, B.1.526 (%)	IFR, B.1.1.7 (%)	Changes, B.1.526 (%)	Changes, B.1.1.7 (%)	Model goodness-of-fit (R ²)
<25	0.004 (0.0021, 0.0059)	0.004 (0.0039, 0.0041)	0.004 (0.0039, 0.0042)	-0.61 (-3.5, 2.3)	0.97 (-2.8, 4.8)	1
25-44	0.04 (0.021, 0.059)	0.037 (0.034, 0.04)	0.043 (0.039, 0.047)	-6 (-13, 1.4)	8.4 (-1.9, 19)	1
45-64	0.29 (0.15, 0.44)	0.42 (0.31, 0.54)	0.51 (0.35, 0.67)	46 (7.4, 84)	76 (22, 130)	0.97
65-74	1 (0.57, 2.5)	1.9 (1.2, 2.5)	3.3 (2.4, 4.2)	82 (20, 140)	210 (130, 300)	0.96
75+	4.1 (2.2, 6.3)	6.7 (5.9, 7.4)	8 (7, 9)	62 (45, 80)	95 (71, 120)	0.99
all	0.35 (0.2, 0.58)	0.56 (0.48, 0.63)	0.72 (0.61, 0.82)	60 (38, 82)	100 (75, 130)	0.99

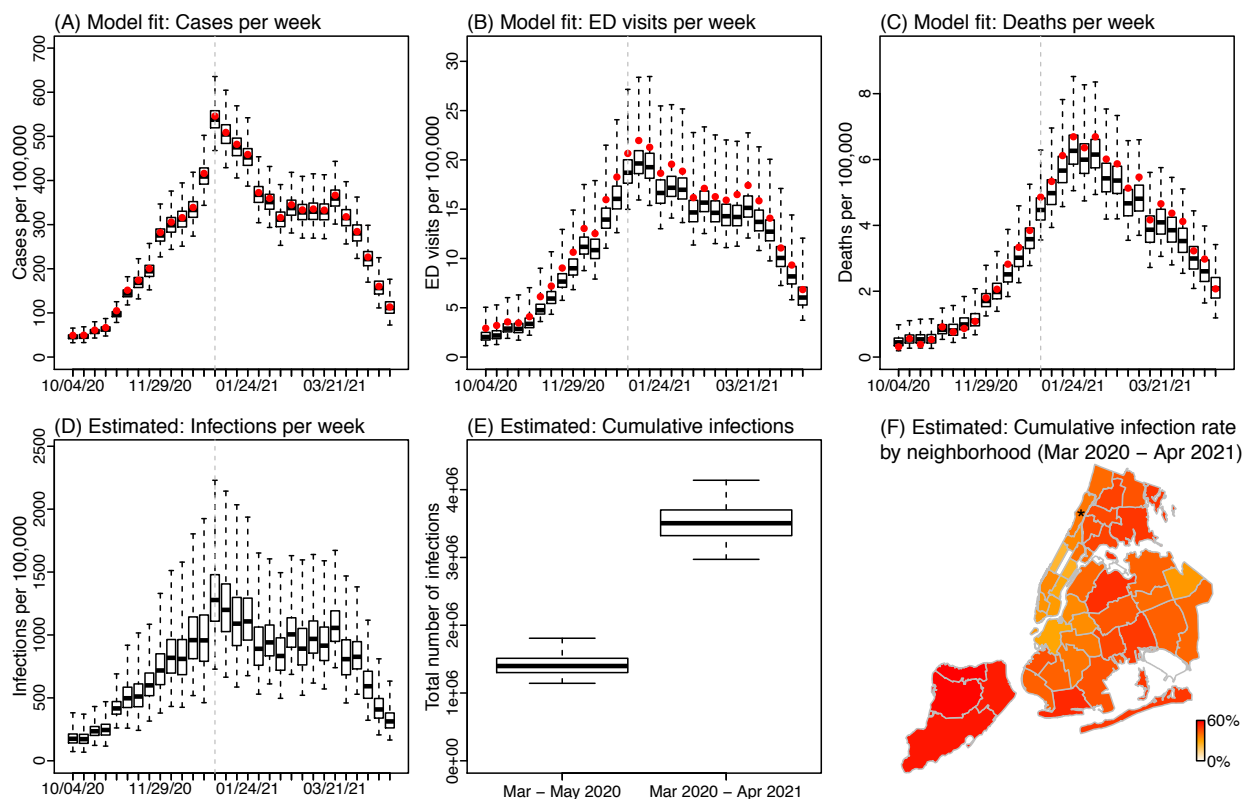
622

623 **Fig 1.** Study design. This study included three modeling analyses: 1) spatial network model-
 624 inference to construct the transmission dynamics and estimate key population variables and
 625 parameters by United Hospital Fund (UHF) neighborhood of residence and age group; 2) city-
 626 level multi-variant, age-structured modeling to simulate and estimate the changes in
 627 transmissibility and immune escape potential for B.1.526; and 3) linear regression models to
 628 estimate variant-specific infection fatality risk (IFR), for B.1.526 and B.1.1.7, separately. Nine
 629 datasets (listed in the black open boxes) were used as model inputs or to evaluate the accuracy
 630 of model estimates (indicated for each dataset below). Models used are shown in the blue filled
 631 boxes and model outputs are listed in the blue open boxes (key estimates reported in detail in
 632 the Results are bolded). Connections among the analyses are indicated by the arrows and
 633 associated annotations.



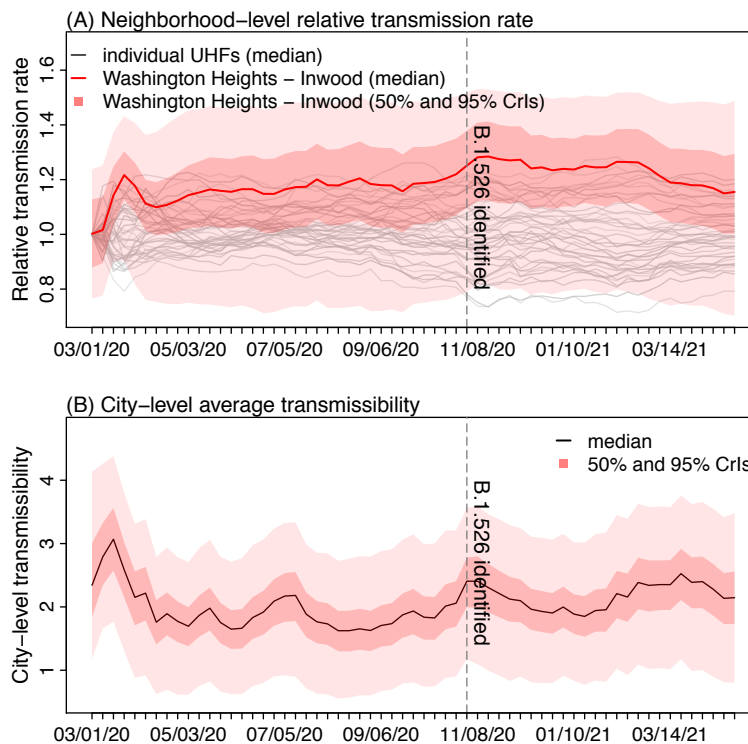
634

635 **Fig 2.** Model fit and key estimates. Upper panel shows model-fit to weekly number of cases (A),
636 ED visits (B), and deaths (C), for all ages combined. Lower panel shows key model-inference
637 estimates of weekly number of infections including those not detected as cases (D), cumulative
638 number of infections in NYC overall (E), and cumulative infection rate by neighborhood (F).
639 Boxes show model estimates (thick horizontal lines and box edges show the median, 25th, and
640 75th percentiles; vertical lines extending from each box show 95% CrI) and red dots show
641 corresponding. For the weekly estimates, week starts (mm/dd/yy) are shown in the x-axis
642 labels. Star (*) in the map indicates the location of the Washington Heights – Inwood
643 neighborhood.



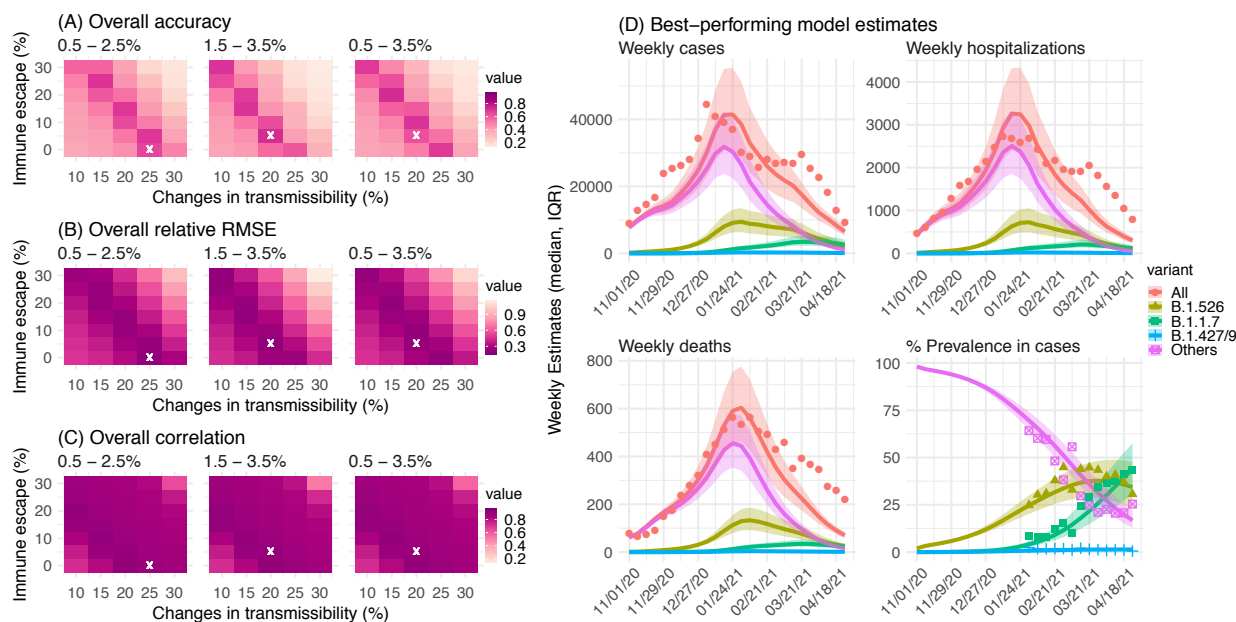
644
645

646 **Fig 3.** Changes in transmission rate. (A) Changes in neighborhood-level relative transmission
647 rate. (B) Changes in citywide transmission rate. Vertical dashed lines indicate the earliest date
648 B.1.526 was identified as reported in Annavajhala et al. Labels of x-axis show the week starts
649 (mm/dd/yy).



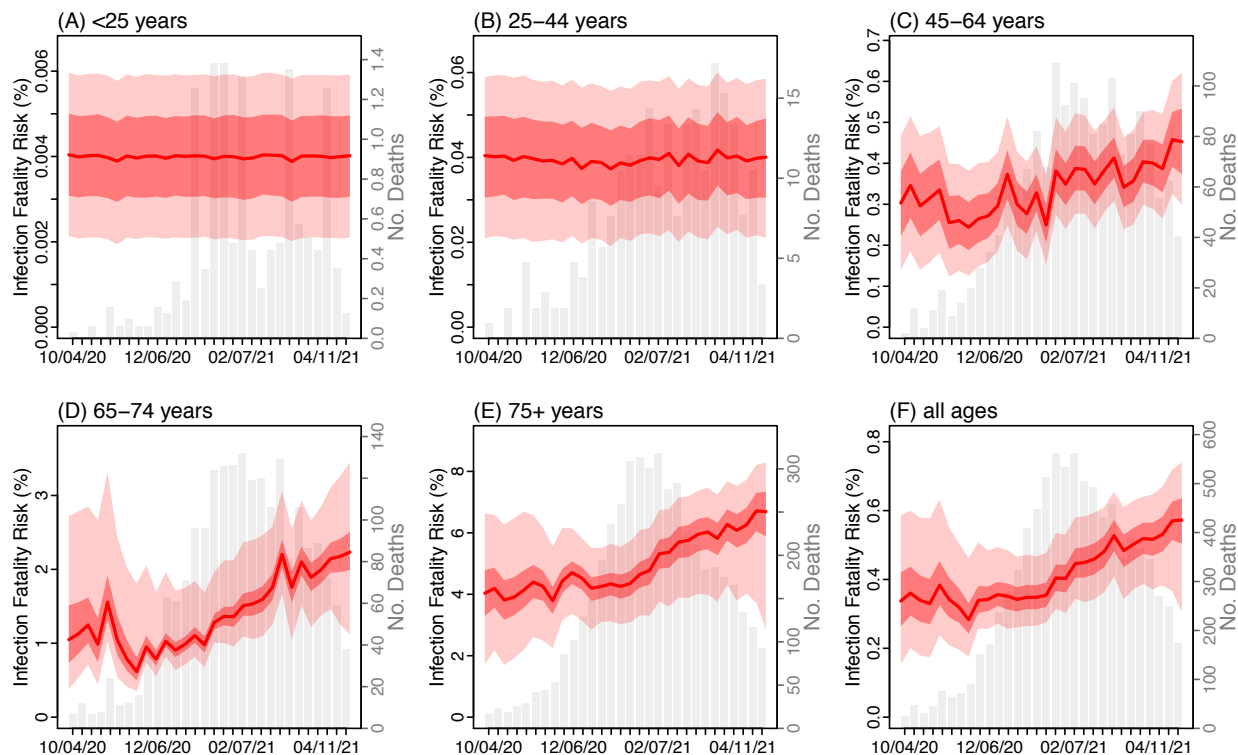
650
651

652 **Fig 4.** Comparison of different combinations of changes in transmissibility and immune escape
653 property for B.1.526. Left panel shows the overall accuracy (A), relative RMSE (B), and
654 correlation (C) of model estimates under different transmissibility and immune escape settings.
655 White crosses (x) indicate the best-performing parameter combination. Right panel shows
656 model estimates using the overall best-performing parameter combination (i.e., 1.5-3.5% initial
657 prevalence, 15-25% higher transmissibility, and 0-10% immune escape). Lines and surrounding
658 areas show model-simulated median estimates and interquartile range; dots show
659 corresponding observations; colors indicate different variants as specified in the legend. Note
660 that these model simulations used same infection-detection rate, hospitalization-rate and IFR
661 (i.e., average during Nov 2020 – Apr 2021); that is, they did not account for changes in case
662 ascertainment or disease severity by week during this period, due to, e.g., increases in disease
663 severity by the new variants. As such, there were larger deviations from the observations
664 during later months of the simulation with more infections by the new variants.



665

666 **Fig 5.** Estimated infection fatality risk. Red lines show the estimated median IFR with
667 surrounding areas indicating the 50% (darker color) and 95% (lighter color) CrI. For comparison,
668 the grey bars show the number of deaths reported for each week from the week of Oct 4, 2020
669 to Apr 25, 2021. X-axis labels show the week starts (mm/dd/yy).



670
671

672
673
674
675
676
677

Supplemental Tables and Figures

Table S1. Prior ranges for the network model-inference system. The prior ranges are similar to Table S1 of Yang et al.¹ but include additional parameters in Eqn 1. The spatial, temporal, and age resolution of each parameter or variable, estimated in the model-inference system, is specified in the column "Resolution". Note posterior parameter estimates can extend outside the specified prior ranges.

Parameter/ variable	Symbol	Resolution	Prior range	Source/rationale
Initial exposed	$E(t=0)$	neighborhood- and age-group specific, estimated for the beginning of the Week of March 1, 2020	300 – 8000 total citywide, scaled by population size for each age group and neighborhood	Large uncertainties, used very wide range
Initial infectious	$I(t=0)$	neighborhood- and age-group specific, estimated for the beginning of the Week of March 1, 2020	150 – 4000 total citywide, scaled by population size for each age group and neighborhood	Assumed to be half the initial exposed
Initial susceptible	$S(t=0)$	neighborhood- and age-group specific, estimated for the beginning of the Week of March 1, 2020	$N - E - I$	Assumed all were susceptible except for those initially exposed/infectious
Population size in each age group and neighborhood	N	neighborhood- and age-group specific	N/A	NYC intercensal population estimates for 2018 ²
Citywide transmission rate	β_{city}	Citywide, age-group specific, estimated for each week	[0.5, 1] per day overall; scaled by contact rate for each age group based on contact data from the POLYMOD study ³ (averaged across 8 countries)	Based on R_0 estimates of around 1.5-4 for SARS-CoV-2 ⁴⁻⁶

Scaling of neighborhood transmission rate	b_i	neighborhood- and age-group specific, estimated for each week	[0.8, 1.2] for age groups under 65 years; [0.5, 1.5] for age groups 65 or older	Around 1; larger variation for elderly groups based on data
Latency period Z		Citywide, age-group specific, estimated for each week	[2, 5] days	Incubation period: 5.2 days (95% CI: 4.1, 7) ⁴ ; latency period is likely shorter than the incubation period
Infectious period	D	Citywide, age-group specific, estimated for each week	[2, 5] days	Time from symptom onset to hospitalization: 3.8 days (95% CI: 0, 12.0) in China, ⁷ plus 1-2 days viral shedding before symptom onset. We did not distinguish symptomatic/asymptomatic infections.
Immunity period	L	Citywide, age-group specific, estimated for each week	[2.5, 3.5] years	Based on estimated immunity period for endemic human coronaviruses (see Appendix of Yang et al. ¹)
Multiplicative factor for mobility; see Yang et al. ¹ for detail	m_1	Citywide, age-group specific, estimated for each week	[1, 2] for <1 year; [0.5, 1.5] for three age groups 1-24 years; [0.1, 1.5] for age group 25-44; [1, 2.5] for age groups 45 or older	Initial model testing showed transmission rates for younger age groups were more sensitive to changes in mobility whereas the two oldest age groups were not sensitive to mobility. For age groups with contact rates lower than the average (based

				on the POLYMOD study ³), we raised the diagonal elements in the mobility matrix to the power of the relative contact rate (<1) to account for insensitivity of transmission rate in these age groups to mobility.
Multiplicative factor for neighborhood connectivity; see Yang et al. ¹ for detail	m_2	Citywide, age-group specific, estimated for each week	[0.5, 2]	Likely around 1 but with large uncertainties
Mean of time from viral shedding to diagnosis; see Yang et al. ¹ for detail	T_m	Citywide, age-group specific, estimated for each week	[3, 8] days	From a few days to a week from symptom onset to diagnosis, ⁷ plus 1-2 days of viral shedding (being infectious) before symptom onset
Standard deviation (SD) of time from viral shedding to diagnosis; see Yang et al. ¹ for detail	T_{sd}	Citywide, age-group specific, estimated for each week	[1, 3] days	To allow variation in time to diagnosis
Infection-detection rate; see Yang et al. ¹ for detail	r	Citywide, age-group specific, estimated for each week	Starting from [0.001, 0.05] at time 0 and allowed to increase over time using space re-probing ⁸	Large uncertainties

Infection fatality risk (IFR); see Yang et al. ¹ for detail	Citywide, age-group specific, estimated for each week	[5, 15]×10 ⁻⁵ for ages under 25; [5, 15]×10 ⁻⁴ for ages 25-44; [5, 15]×10 ⁻³ for ages 45-64; [0.01, 0.1] for ages 65-74; [0.02, 0.2] for ages 75+;	Based on previous estimates ⁹ but extend to have wider ranges
Time from diagnosis to death; see Yang et al. ¹ for detail	Citywide	Gamma distribution with mean of 9.36 days and SD of 9.76 days	Based on <i>n</i> =15,686 COVID-19 confirmed deaths in NYC as of May 17, 2020.
ED consultation rate (EDR)	Citywide, age-group specific, estimated for each week	[0.001, 0.02] for ages under 25; [0.003, 0.03] for ages 25-44; [0.006, 0.06] for ages 45-64; [0.01, 0.15] for ages 65-74; [0.02, 0.25] for ages 75+;	Based on the ratio of total ED visits and estimated infections during March – Dec 2020
Time-from-infectiousness-to-ED or hospitalization	Citywide, for all ages	Gamma distribution with mean of [5, 7] days and SD of [2, 4] days	

678
679

680 **Table S2.** Initial conditions used to simulate co-circulation of different variants in the multi-
 681 variant, age-structured model. To partially account for changing infection-detection rate, ED-
 682 consultation rate (EDR) and IFR, for these three parameters, we used the model-inference
 683 estimates averaged over the entire simulation period (i.e. Nov 2020 – April 2021). For the initial
 684 transmission rate (for the preexisting non-VOC/VOI variants), we used the model-inference
 685 estimates averaged over the week of 10/25/2020 – the week of 11/7/2020 (i.e. the 3 weeks
 686 around the start of simulation). For the rest of model state variables and parameters, we used
 687 model-inference estimates made at the week of 10/25/2020. For B.1.1.7, we used the
 688 following ranges based on estimates from Yang and Shaman¹⁰: 40.3 – 52.3% higher
 689 transmissibility (related to estimates for the preexisting non-VOC/VOI variants listed below) and
 690 0 – 10% immune escape; for comparison, contact tracing data from the UK showed that B.1.1.7
 691 was 30-50% more infectious.¹¹ For B.1.427/ B.1.429, we used the following ranges based on
 692 estimates from Deng et al.¹²: 16 – 24% higher transmissibility and 0-10% immune escape (vs.
 693 21.4 – 27.8% increase in transmission rate in Deng et al.¹² without accounting for changes in
 694 immunity due to potential immune escape).

variant	parameter	lower bound	upper bound
B.1.526	Low initial prevalence (%)	0.5	2.5
B.1.526	High initial prevalence (%)	1.5	3.5
B.1.526	Wide initial prevalence (%)	0.5	3.5
B.1.1.7	Increase in transmission rate	0.403	0.5227
B.1.1.7	Immune escape	0	0.1
B.1.427/9	Increase in transmission rate	0.16	0.24
B.1.427/9	Immune escape	0	0.1
non-VOC/VOI	Travel-related importation ϵ_i	Nominally set to 1 per week for the entire city (N = 8.4 million people)	
B.1.526	Travel-related importation ϵ_i	Set to 0 as it emerged locally	
B.1.1.7	Travel-related importation ϵ_i	For the entire city (N = 8.4 million), set to 1 per 2 days for 11/1 – 11/15/20 to reflect lower initial seeding, 1.5 per day for 11/16 -12/31/20 to reflect higher seeding during the holidays, and 2 per day for 1/1 – 4/30/21 to reflect higher seeding due to increases in these variants in the US. Same settings were used for B.1.1.7 and B.1.427/9, because once local transmission is established, travel-related importation plays a nominal role.	
B.1.427/9	Travel-related importation ϵ_i		

non-VOC/VOI	β_{11} (per day, same below)	0.14	0.21
non-VOC/VOI	β_{22}	0.15	0.2
non-VOC/VOI	β_{33}	0.16	0.22
non-VOC/VOI	β_{44}	0.16	0.23
non-VOC/VOI	β_{55}	0.24	0.36
non-VOC/VOI	β_{66}	0.17	0.25
non-VOC/VOI	β_{77}	0.17	0.22
non-VOC/VOI	β_{88}	0.19	0.25
non-VOC/VOI	β_{12}	0.071	0.1
non-VOC/VOI	β_{13}	0.018	0.027
non-VOC/VOI	β_{14}	0.0074	0.011
non-VOC/VOI	β_{15}	0.023	0.034
non-VOC/VOI	β_{16}	0.011	0.015
non-VOC/VOI	β_{17}	0.0075	0.011
non-VOC/VOI	β_{18}	0.0052	0.0075
non-VOC/VOI	β_{21}	0.074	0.1
non-VOC/VOI	β_{23}	0.019	0.026
non-VOC/VOI	β_{24}	0.0078	0.011
non-VOC/VOI	β_{25}	0.024	0.033
non-VOC/VOI	β_{26}	0.011	0.015
non-VOC/VOI	β_{27}	0.0079	0.011
non-VOC/VOI	β_{28}	0.0054	0.0074
non-VOC/VOI	β_{31}	0.02	0.028
non-VOC/VOI	β_{32}	0.02	0.028
non-VOC/VOI	β_{34}	0.0098	0.014
non-VOC/VOI	β_{35}	0.013	0.018
non-VOC/VOI	β_{36}	0.0072	0.01
non-VOC/VOI	β_{37}	0.0051	0.0071
non-VOC/VOI	β_{38}	0.0066	0.0091
non-VOC/VOI	β_{41}	0.0085	0.013
non-VOC/VOI	β_{42}	0.0085	0.013
non-VOC/VOI	β_{43}	0.013	0.02
non-VOC/VOI	β_{45}	0.014	0.021
non-VOC/VOI	β_{46}	0.01	0.015
non-VOC/VOI	β_{47}	0.0037	0.0054
non-VOC/VOI	β_{48}	0.0069	0.01
non-VOC/VOI	β_{51}	0.11	0.17
non-VOC/VOI	β_{52}	0.11	0.17

non-VOC/VOI	β_{53}	0.092	0.14
non-VOC/VOI	β_{54}	0.073	0.11
non-VOC/VOI	β_{56}	0.07	0.11
non-VOC/VOI	β_{57}	0.042	0.064
non-VOC/VOI	β_{58}	0.041	0.062
non-VOC/VOI	β_{61}	0.053	0.076
non-VOC/VOI	β_{62}	0.053	0.076
non-VOC/VOI	β_{63}	0.043	0.063
non-VOC/VOI	β_{64}	0.051	0.074
non-VOC/VOI	β_{65}	0.053	0.076
non-VOC/VOI	β_{67}	0.058	0.084
non-VOC/VOI	β_{68}	0.052	0.076
non-VOC/VOI	β_{71}	0.032	0.041
non-VOC/VOI	β_{72}	0.032	0.041
non-VOC/VOI	β_{73}	0.022	0.028
non-VOC/VOI	β_{74}	0.0096	0.012
non-VOC/VOI	β_{75}	0.023	0.029
non-VOC/VOI	β_{76}	0.032	0.042
non-VOC/VOI	β_{78}	0.066	0.085
non-VOC/VOI	β_{81}	0.028	0.036
non-VOC/VOI	β_{82}	0.028	0.036
non-VOC/VOI	β_{83}	0.03	0.038
non-VOC/VOI	β_{84}	0.022	0.028
non-VOC/VOI	β_{85}	0.027	0.034
non-VOC/VOI	β_{86}	0.047	0.06
non-VOC/VOI	β_{87}	0.073	0.093
all	Z_1 (days, same below)	2.9	4
all	Z_2	3.3	4.3
all	Z_3	3.4	4.4
all	Z_4	3.4	4.4
all	Z_5	3.5	4.4
all	Z_6	3.5	4.5
all	Z_7	3.3	4.2
all	Z_8	3.3	4.2
all	D_1	2.1	2.9
all	D_2	2.7	3.6
all	D_3	3.2	4.1
all	D_4	3.4	4.4

all	D ₅	3.1	4.1
all	D ₆	3.2	4.2
all	D ₇	2.8	3.7
all	D ₈	2.5	3.3
all	IFR ₁	3.10E-05	5.00E-05
all	IFR ₂	3.00E-05	4.90E-05
all	IFR ₃	3.10E-05	4.90E-05
all	IFR ₄	3.00E-05	4.90E-05
all	IFR ₅	3.00E-04	0.00048
all	IFR ₆	0.0029	0.004
all	IFR ₇	0.013	0.016
all	IFR ₈	0.046	0.055
all	EDR ₁	0.0087	0.013
all	EDR ₂	0.0059	0.0084
all	EDR ₃	0.0023	0.0033
all	EDR ₄	0.0046	0.006
all	EDR ₅	0.01	0.012
all	EDR ₆	0.019	0.023
all	EDR ₇	0.031	0.039
all	EDR ₈	0.057	0.07
all	<i>Infection detection rate, r₁</i>	0.14	0.19
all	<i>Infection detection rate, r₂</i>	0.21	0.26
all	<i>Infection detection rate, r₃</i>	0.23	0.29
all	<i>Infection detection rate, r₄</i>	0.28	0.34
all	<i>Infection detection rate, r₅</i>	0.39	0.47
all	<i>Infection detection rate, r₆</i>	0.36	0.42
all	<i>Infection detection rate, r₇</i>	0.33	0.41
all	<i>Infection detection rate, r₈</i>	0.34	0.41

695

696

697 **Table S3.** Estimated IFR for different variants and changes compared to the baseline risk
 698 estimated for preexisting variants during Oct – Dec 2020, using Eqn 3.

Age	IFR, baseline (%)	IFR, B.1.526 (%)	Changes, B.1.526 (%)	Model fit, R ²
<25	0.004 (0.0021, 0.0059)	0.004 (0.0038, 0.0041)	-0.03 (-3.8, 3.8)	1
25-44	0.04 (0.021, 0.059)	0.035 (0.031, 0.038)	-12 (-21, -3.1)	0.97
45-64	0.29 (0.15, 0.44)	0.42 (0.24, 0.59)	43 (-17, 100)	0.67
65-74	1 (0.57, 2.5)	1.5 (0.53, 2.5)	46 (-50, 140)	0.46
75+	4.1 (2.2, 6.3)	6.1 (5.1, 7)	47 (25, 69)	0.94
all	0.35 (0.2, 0.58)	0.5 (0.4, 0.61)	43 (13, 73)	0.89

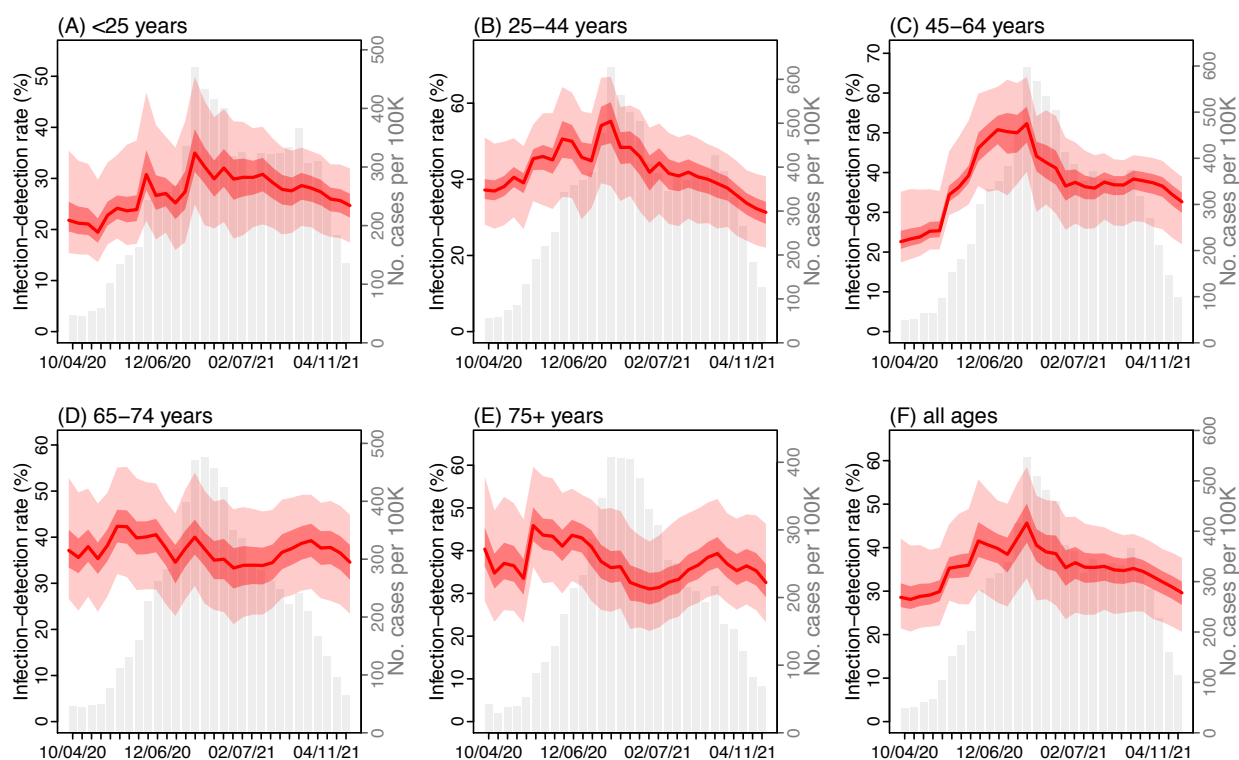
699

700 **References included in the tables:**

- 701 1. Yang W, Kandula S, Huynh M, et al. Estimating the infection-fatality risk of SARS-CoV-2 in
 702 New York City during the spring 2020 pandemic wave: a model-based analysis. *The Lancet*
 703 *Infectious diseases* 2021; **21**(2): 203-12.
- 704 2. New York City Department of Health and Mental Hygiene. NYC DOHMH population
 705 estimates, modified from US Census Bureau interpolated intercensal population estimates,
 706 2000-2018. . Updated August 2019. ed.
- 707 3. Mossong J, Hens N, Jit M, et al. Social contacts and mixing patterns relevant to the spread of
 708 infectious diseases. *PLoS Med* 2008; **5**(3): e74.
- 709 4. Li Q, Guan X, Wu P, et al. Early Transmission Dynamics in Wuhan, China, of Novel
 710 Coronavirus–Infected Pneumonia. *New Engl J Med* 2020.
- 711 5. Wu JT, Leung K, Leung GM. Nowcasting and forecasting the potential domestic and
 712 international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling
 713 study. *Lancet* 2020.
- 714 6. Li R, Pei S, Chen B, et al. Substantial undocumented infection facilitates the rapid
 715 dissemination of novel coronavirus (SARS-CoV-2). *Science* 2020; **368**(6490): 489-93.
- 716 7. Zhang J, Litvinova M, Wang W, et al. Evolving epidemiology and transmission dynamics of
 717 coronavirus disease 2019 outside Hubei province, China: a descriptive and modelling study.
 718 *The Lancet Infectious diseases* 2020.
- 719 8. Yang W, Shaman J. A simple modification for improving inference of non-linear dynamical
 720 systems. *arXiv* 2014: 1403.6804.
- 721 9. Verity R, Okell LC, Dorigatti I, et al. Estimates of the severity of coronavirus disease 2019: a
 722 model-based analysis. *The Lancet Infectious diseases* 2020.
- 723 10. Yang W, Shaman J. Epidemiological characteristics of three SARS-CoV-2 variants of concern
 724 and implications for future COVID-19 pandemic outcomes. *medRxiv* 2021:
 725 2021.05.19.21257476.

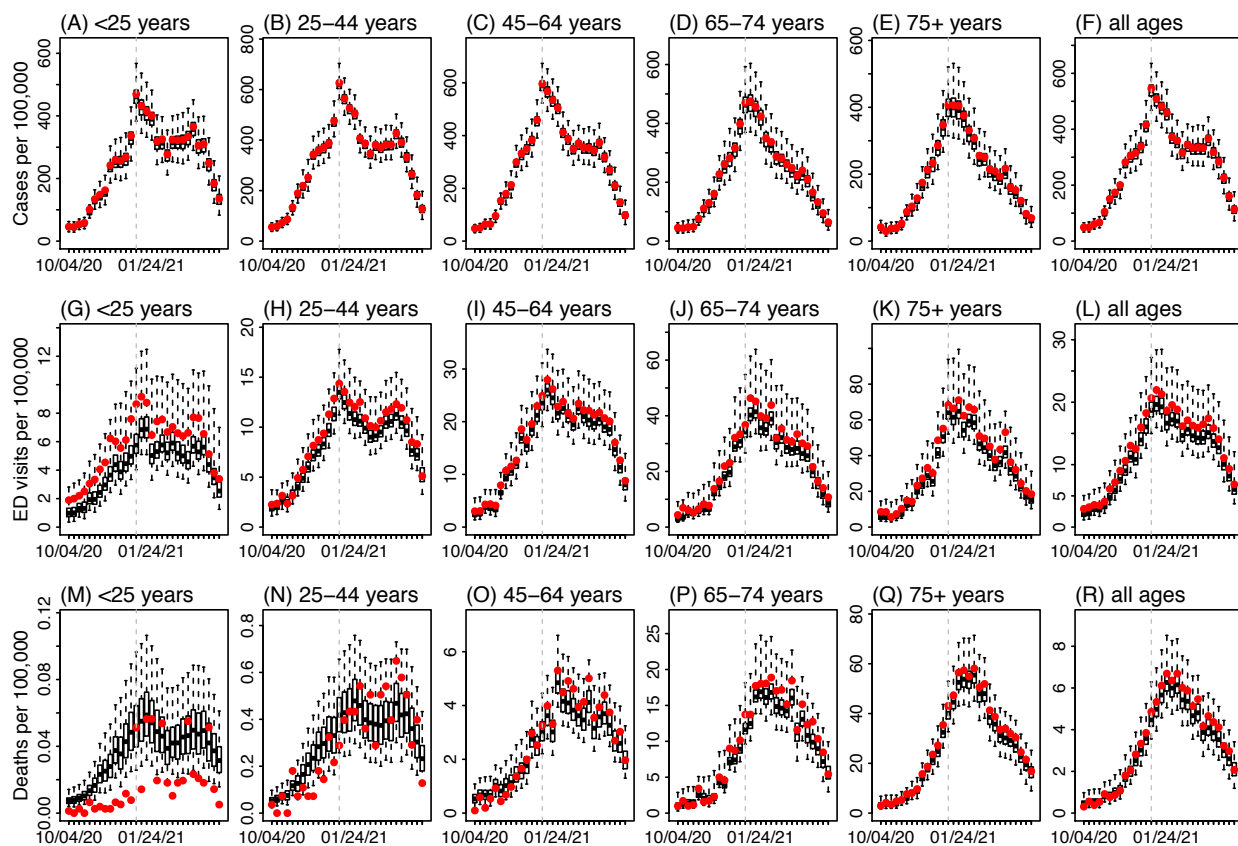
- 726 11. Public Health England. Investigation of novel SARS-CoV-2 variant, Variant of Concern
727 202012/01, Technical briefing 3, 2020.
728 12. Deng X, Garcia-Knight MA, Khalid MM, et al. Transmission, infectivity, and antibody
729 neutralization of an emerging SARS-CoV-2 variant in California carrying a L452R spike
730 protein mutation. *medRxiv* 2021: 2021.03.07.21252647.

731
732 **Fig S1.** Estimated infection-detection rate by age group. Red lines show the estimated median
733 infection-detection rate with surrounding areas indicating the 50% (darker color) and 95%
734 (lighter color) CrI. For comparison, the grey bars show the number of cases reported for each
735 week from the week of Oct 4, 2020 to Apr 25, 2021. Labels of x-axis show the week starts
736 (mm/dd/yy).



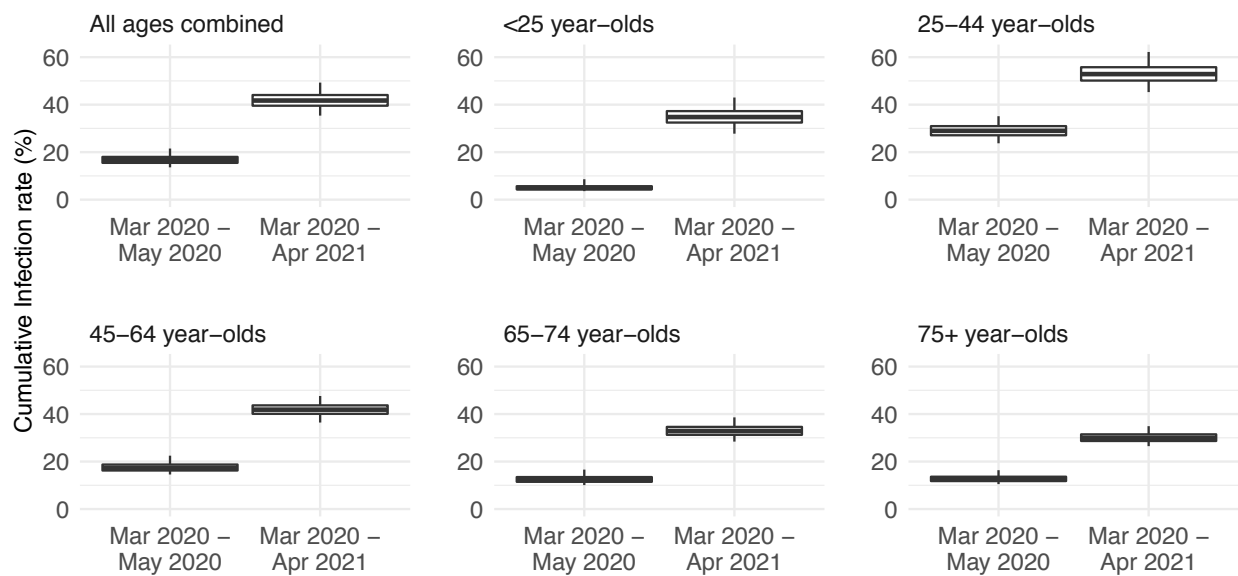
737
738

739 **Fig S2.** Model-fit by age group. Boxes show model estimates (thick horizontal lines and box
740 edges show the median, 25th, and 75th percentiles; vertical lines extending from each box show
741 95% CrI) and red dots show corresponding.



742

743 **Fig S3.** Estimated cumulative infection rates by age group. Thick horizontal lines and box edges
744 show the median, 25th, and 75th percentiles; vertical lines extending from each box show 95%
745 CrI.



746
747
748