

Adaptive vaccination may be needed to extirpate COVID-19: Results from a runtime-alterable strain-drift and waning-immunity model

Wayne M. Getz^{1,2,3}, Richard Salter^{3,4}, Ludovica Luisa Vissat¹,
James S. Koopman⁵, Carl P. Simon⁶

¹Dept. ESPM, UC Berkeley, CA 94720-3114, USA

²School of Mathematical Sciences, University of KwaZulu-Natal, South Africa

³Numerus, 850 Iron Point Rd., Folsom, CA 95630, USA

⁴Computer Science Dept., Oberlin College, Oberlin, Ohio, OH 44074, USA

⁵School of Public Health, University of Michigan, Ann Arbor, MI 48109, USA

⁶Gerald R. Ford School of Public Policy, University of Michigan, Ann Arbor, MI 48109, USA

Corresponding Author: Wayne M. Getz, wgetz@berkeley.edu

Author Contributions: JSK and CPS initiated the project, provided feedback on early drafts and participated in discussions. WMG designed and formulated the SEIVD-IBM model, drafted the manuscript, ran simulations, and produced the figures. RS designed and built the J-RAMP platform and wrote the code and informational text. LLV checked all the mathematical equations and code. Both RS and LLV contributed to the structure of the manuscript. All authors contributed to editing and adding text as needed.

Competing Interest Statements: none

Classification. Major: Biological Sciences; Minor: Biophysics and Computational Biology

Keywords: SARS-CoV-2, SEIR models, multivalent vaccines, escape mutations, strain/variant proliferation

1 **Abstract**

2 We developed an elaborated susceptible-infected-recovered (SIR) individual-
3 based model (IBM) with pathogen strain drift, waning and cross immunity,
4 implemented as a novel Java Runtime-Alterable-Model Platform (J-RAMP).
5 This platform allows parameter values, process formulations, and scriptable
6 runtime drivers to be easily added at the start of simulation. It includes facility
7 for integration into the R statistical and other data analysis platforms. We se-
8 lected a set of parameter values and process descriptions relevant to the current
9 COVID-19 pandemic. These include pathogen-specific shedding, environmen-
10 tal persistence, host transmission and mortality, within-host pathogen mutation
11 and replication, adaptive social distancing, and time dependent vaccine rate and
12 strain valency specifications. Our simulations illustrate that if waning immu-
13 nity outpaces vaccination rates, then vaccination rollouts may fail to contain the
14 most transmissible strains. Our study highlights the need for adaptive vaccina-
15 tion rollouts, which depend on reliable real-time monitoring and surveillance of
16 strain proliferation and reinfection data needed to ensure that vaccines target
17 emerging strains and constrain escape mutations. Together with such data, our
18 platform has the potential to inform the design of vaccination programs that
19 extirpate rather than exacerbate local outbreaks. Finally, our RAMP concept
20 promotes the development of highly flexible models that can be easily shared
21 among researchers and policymakers not only addressing healthcare crises, but
22 other types of environmental crises as well.

23
24 **Significance Statement:** Effective COVID-19 vaccine development has been
25 unprecedented, but vaccinations are being delivered at contrasting rates across
26 the globe. Here, using an innovative epidemiological Java runtime alterable
27 modeling platform (J-RAMP), we demonstrate that seemingly reasonable vac-
28 cination programs may exacerbate rather than mitigate regional outbreaks when
29 immune waning outpaces vaccinations and reinfection promotes escape muta-
30 tions. Our simulations suggest that adaptive vaccination programs, requiring
31 adequate strain and serology monitoring and surveillance, are needed to extir-
32 pate outbreaks. Our platform provides policymakers with an easy-to-use tool
33 for designing effective vaccination programs. Our RAMP concept points the
34 way to the development of highly flexible models that are easily shared among
35 researchers and policymakers in all areas of systems analysis.

36 Introduction

37 Vaccines have either eliminated, or all but eliminated, diphtheria, measles, po-
38 lio, rubella, small pox, and tetanus within developed countries, such as the
39 USA [1]. The global COVID-19 pandemic has amply illustrated the overwhelm-
40 ing economic importance that vaccines can play in bringing an epidemic un-
41 der control [2]. Unlike the above mentioned diseases and more inline with in-
42 fluenza, however, the SARS-CoV-2 pathogen may re-emerge with vengeance
43 due both to waning immunity and proliferation of mutant strains or variants
44 with increased transmissibility that escape both naturally-acquired and vaccine-
45 induced immunity [3–8]. To investigate how effective vaccination programs may
46 be in preventing future large-scale outbreaks of COVID-19 [9, 10], we built a
47 Java Runtime-Alterable-Model Platform (J-RAMP) of an SEIVD (Susceptible,
48 Exposed, Infectious, V=Immune & Vaccinated, Dead) individual-based model
49 (IBM) [11–14] that includes the following processes (Fig. 1; see Materials and
50 Methods, as well as Appendix A for details):

- 51 • pathogen strain-specific shedding [15], environmental persistence [16], with-
52 host replication [17] and mortality rates [18] (some prefer the term variant
53 rather than strain)
- 54 • immunological waning with strain cross immunity [6, 19]
- 55 • pathogen strain drift during transmission and within-host replication [20]
- 56 • an adaptive contact rate [21]
- 57 • a time-dependent, uni- or multivalent vaccine rollout [22, 23]

58 Our SEIVD-IBM J-RAMP, is an example of our novel runtime alterable-model
59 platform (RAMP) concept, which includes panels, windows and sliders that al-
60 low users to specify and manipulate model parameter values in real time, as well
61 as modify process function descriptions, even during the course of a simulation,
62 while protecting the integrity of the underlying code. Thus our SEIVD-IBM
63 J-RAMP has epidemiological applications well beyond COVID-19, including
64 diseases such as influenza [24], Ebola [25, 26], and any other directly transmit-
65 ted disease [27]. To illustrate its application, we used our J-RAMP to explore
66 the relative efficacy of uni- and multivalent vaccines applied at various time-
67 dependent rates, where choice of valency may switch in response to realtime
68 monitoring and surveillance data. Such adaptive vaccination programs may be
69 required to combat the evolutionary arms race between vaccine efficacy and the
70 evolution of new pathogen strains [20, 23, 28]. We demonstrate the possible con-
71 sequences of vaccinating too few individuals in the context of the emergence of
72 more transmissible strains, as well as the comparative efficacy of non-adaptive
73 versus adaptive vaccines responsive in realtime to current strain valency. We
74 simulate plausible COVID-19 outbreaks over a three year period in which vac-
75 cination rollout commences at the start of year two and continues until the end
76 of the year three.

77 Our study is meant to highlight issues related to vaccination rollout rates [10]
78 and strain valency that may turn out to provide insufficient coverage of the pop-
79 ulation or being outrun by waning immunity in vaccinated hosts or those that
80 have survived COVID-19. We undertake these simulation with the knowledge
81 that though a typical set of parameters may cover most outbreaks of COVID-19

82 in various regions around the world, regional differences in the contact behavior
83 of individuals, with additional contact rate time dependency in each locality
84 due to social distancing regulations and behavior, have led to markedly differ-
85 ent outbreak patterns across regions [29–31]. Thus the result presented here are
86 meant to inform us of potential pitfalls and dangers in rolling out inadequate
87 vaccination programs and extract general principles, where possible, as well as
88 illustrate an approach to its local application rather than prescribe a universal
89 program: adaptive programs need to operate in response to local surveillance
90 data and contact rate conditions.

91 The local nature of contact rate patterns is well established as an important
92 driver of outbreak patterns [30]. The definition of a contact where transmission
93 may occur is rather challenging though. Effective or “close” contact may be
94 defined, as in [32], in terms of direct physical contact or close conversation, or it
95 may involve concepts of a threshold distance and time period [33]. Taken a step
96 further, it may relate to infectious dose levels and the relationship between dose
97 and severity of the infection of individuals [34]. This effective contact rate issue,
98 however, is side-stepped by fitting models to local incidence data to obtain a
99 concatenated contact-and-transmission rate parameter value that best fits local
100 outbreak patterns, as discussed in depth elsewhere [30].

101 Our results also suggest that adaptive vaccination programs that closely
102 monitor the emergence of new strains and adapt vaccine valency to match pre-
103 dominating strains are much more likely to extirpate local epidemics than non-
104 adaptive policies, though our simulation results are greatly influenced by as-
105 sumptions regarding rates of waning immunity and strain cross-immunity char-
106 acteristics as well. Accurate forecasting of the future course of an epidemic in
107 terms of strain composition and vaccination program efficacy require data on
108 the rates at which immunity wanes, how such immunity depends on particu-
109 lar strains, and the levels of cross immunity among strains [4, 6] that are more
110 comprehensive than currently available. It also requires better data on within-
111 host replication and mutation rates. Thus, in truth, we present our model
112 and results fully aware that answers to questions relating the design of optimal
113 region-specific vaccination programs await more reliable data on immunity pro-
114 cesses, strain mutation rates [35], and the transmission potential and virulence
115 of dominant strains [36]. We hope, however, that our results and subsequent
116 investigations using our model provide the kinds of quantitative analyses that
117 can help formulate highly effective local or country level vaccination programs
118 that avoid many of the pitfalls revealed by our analysis, as well as encourage
119 the adoption of rapidly adaptive vaccination programs.

120 Results

121 Single strain simulations

122 Our first simulation, which we regard as a “baseline run” has $J = 0$ (i.e.,
123 only one strain is possible), assumes a constant contact rate ($\kappa(t) = \kappa_0$ for all t ;
124 which implies no social distancing or effective contact reducing actions taken) in
125 an a homogeneously, well-mixed population of ten thousand individuals ($N_0 =$
126 10,000), with an immune waning half-life of one year ($t^{\text{half}} = 365$), and no
127 vaccination rollout. The resulting simulation depicts a classic outbreak pattern

128 (Fig. 2) in which incidence rises to a peak of about 250 cases around 96 days
129 after “patient zero” has been introduced, and then burns out around day 166
130 days because an insufficient number of susceptible individuals are now available
131 to keep the epidemic going. In the upper left-hand panel of our SEIVD-IBM
132 J-RAMP dashboard in Fig. 2, we see that around 3/4 of individuals have been
133 infected at the point where the epidemic burns out ($V(166) = 7700$, which is
134 77% of N_0): this suggests that herd immunity may well be in the ball park of
135 75-80% (as supported by additional runs presented next), although this would
136 be an underestimate if more transmissible strains come to prevail.

137 Repeated runs of this first simulation with different runtime seeds (viz., 10
138 runs with runtime seeds ranging from 0 to , where each seed generates its own,
139 essentially unique, but repeatedly sequence of pseudo random numbers; Fig. 2
140 is for the case runtime seed = 0) produces the following means and standard
141 deviations as a percentage of the population size (of S, V and D), as well as the
142 length of the epidemic (days) (see Fig. C.1, Appendix C): $S_{\text{fml}} = 21.9 \pm 0.6\%$,
143 $V_{\text{fml}} = 76.4 \pm 0.6\%$, $D_{\text{fml}} = 1.6 \pm 0.1\%$, and $t_{\text{fml}} = 167 \pm 9$ days. These results
144 are almost the same when the initial population size is increased threefold to
145 $N_0 = 30,000$ individuals, the primary difference being a small expected decrease
146 in variability (around $1/3^{\text{rd}}$ smaller) and slightly longer epidemic period (18%
147 longer) (see Fig. C.1 for details). To obtain full plots of the incidence, prevalence,
148 and daily-death trajectories from many runs of our SEIV-IBM, we exploited
149 the R-platform integrability of our J-RAMP to generate mean and standard
150 deviation plots of incidence, prevalence, and daily mortality, as described in
151 Appendix B (and Fig. C.2).

152 Our second single-strain ($J = 0$) simulation (Fig. 3A) illustrates that, for the
153 selected set of parameter values, if the rate of waning is tripled from that of our
154 first (i.e., t^{half} is reduced from 365 days to 120 days) then herd immunity is still
155 reached. Our third single strain simulation, however, shows that herd immunity
156 is eluded for the selected set of parameter values when this rate is quadrupled
157 ($t^{\text{half}} = 90$ days; Fig. 3B): viz., the population enters a periodic outbreak mode
158 because previously infected individuals are sufficiently susceptible after several
159 months post-initial infection to be reinfected and thus keep the epidemic going.
160 Herd immunity in this caricatured population (i.e., relatively small, isolated,
161 homogeneous) would also be reached within a year if the waning immunity
162 half-life is increased back to $t^{\text{half}} = 180$, but not if an adaptive contact rate
163 is in operation (Fig. 3C & D; $k(t)$ specified by Eq. A.10). More specifically,
164 for the case $p_{\text{I}}^{\text{half}} = 0.05$, that is the “effective” contact rate $\kappa(t)$ drops from
165 $\kappa(t) = \kappa_0 = 4$ when $I/N_0 = 0$ to $\kappa(t) = \kappa_0/2 = 2$ when $I(t)/N(t) = 0.05$, our
166 fourth single strain simulation indicates that prevalence peaks at roughly 400
167 individuals (i.e., just around 0.4% of the population) around day 85 (Fig. 3C),
168 drops to around 50 between days 210 and 240, and then rises back to over
169 200 by the end of one year. In our fifth single strain simulation (Fig. 3D), an
170 additional 5-fold reduction in our adaptive contact rate (i.e., $\kappa(t) = \kappa_0/2 = 2$
171 when $I(t)/N(t) = 0.01$) produces a prevalence of just under $100/10,000=0.01\%$
172 that appears to persist with stochastic noise from day 100 onwards.

173 We carried out two additional single strain simulations, this time for three
174 years, and applied a vaccination rollout program that vaccinated individuals
175 drawn at random in the second and third year at rates $v = 0.001$ and $v =$

176 0.002 (0.1% and 0.2% of the population vaccinated each day; Fig. 4A & B),
177 respectively, from a pool consisting of individuals not currently infectious or
178 previously vaccinated. In these runs, we set $t^{\text{half}} = 365$ days, $p_1^{\text{half}} = 0.002$
179 (reflecting US case percentages after 1 year), with a population size of $N_0 =$
180 100,000. In the case $v = 0.001$, the vaccination rate was insufficient to overcome
181 the rate at which infected individuals were becoming reinfected due to waning
182 immunity. In the case $v = 0.002$, the vaccination rate was sufficiently high to
183 bring the epidemic under control. In the latter case, however, if we no longer
184 allow individuals who had previously been infected to be part of the vaccination
185 pool as well (i.e. we confine vaccinations to the **S** class only), the vaccination
186 rollout rate $v = 0.002$ is now insufficient to extinguish the outbreak and a
187 resurgence occurs around the end of the second year (Fig. 4C).

188 Multi-strain simulations

189 In all our multi-strain simulations we set $p_1^{\text{half}} = 0.005$. This latter value, as
190 previously mentioned, produces a cumulative number of cases of around 10%
191 in the first year in a population of size $N_0 = 100,000$ (Appendix C, Fig. C.3),
192 which is in the ballpark of 8.5% for the US outbreak during its first year.

193 In our first set of multistrain simulations, we set $J = 4$ (i.e., permitting
194 16 strains) and compared the situations where cross immunity is absent ($c = 0$;
195 Fig. 5A) to where it has the intermediate value $c = 0.5$ (Fig. 5B). The remaining
196 parameters were the same for all strains and used the same basic epidemiological
197 parameters, as in our previous single strain simulations, although mutation rates
198 were now added to enable new strains to proliferate (see caption to Fig. 5 for
199 details on parameter values).

200 In the no cross immunity (Fig. 5A) case, we see that the initial surge is
201 driven by strain $0 \equiv (0, 0, 0, 0)$ (see discussion in Appendix A for use of binary
202 representation), which is first replaced by strain $8 \equiv (1, 0, 0, 0)$ around the end
203 of year 1, which in turn is somewhat displaced by strains $7 \equiv (0, 1, 1, 1)$ and
204 $6 \equiv (0, 1, 1, 0)$, with strain $1 \equiv (0, 0, 0, 1)$ coming to dominate at the start of year
205 3, only to be finally challenged by strain $3 \equiv (0, 0, 1, 1)$ towards the end of year
206 3. In the nearest-neighbor cross immunity case (Fig. 5B), given the retarding
207 effect of cross immunity on the proliferation of strains that differ in only one
208 allele from the initial dominating strain $0 \equiv (0, 0, 0, 0)$, this initial strain was
209 displaced by strains $3 \equiv (0, 0, 1, 1)$ and $11 \equiv (1, 0, 1, 1)$ around the end of year
210 1. These latter 2 strains were then displaced by strain $7 \equiv (0, 1, 1, 0)$ during the
211 first half of year 3, with strain $13 \equiv (1, 1, 0, 1)$ dominating at the end of year 3
212 somewhat supported at a lower level by strain $14 \equiv (1, 1, 1, 0)$. We also noticed
213 a moderate resurgence of strain $0 \equiv (0, 0, 0, 0)$ in the middle of year 3.

214 In a second set of multistrain simulations, we set $J = 7$ for a total of 128
215 possible strains. Unlike the previous set of multistrain simulations, we now
216 provided a strain specific set of transmission parameters $\bar{\beta}_j$ by selecting these at
217 random on the interval $[0, 0.3]$, apart from a dominant sequence of transmission
218 values that steps in increments of 0.25 from 0.3 in strain 0 to 0.475 in strain
219 127 along the sequence mentioned in the caption to Fig. 6 and documented
220 in Fig. C.4. The remaining parameters are the same as for the simulation
221 depicted in Fig. 5B (i.e., the cross-immunity parameter $c = 0.5$). Unlike the

222 case, $J = 4$, only two strains now played any serious role in the outbreak in the
223 non-vaccination scenario (Fig. 6A): strain 0 $\equiv (0, 0, 0, 0, 0, 0, 0)$ for the first year
224 and strain 127 $\equiv (1, 1, 1, 1, 1, 1, 1)$, the most transmissible, for the remaining 2
225 years.

226 Vaccination rollout simulations

227 Following our no-vaccination simulation of our 128-strain case (Case A; Fig. 6A),
228 we explored three different vaccination rollout programs applied in years 2 and
229 3 with the following vaccination rates and vaccine strain valencies: Case B, rate
230 $v = 0.002$ (i.e., 0.2% vaccinated per day), strain valency 0 (Fig. 6B); Case C, rate
231 $v = 0.002$, strain valency 127 (Fig. 6C); and Case D, $v = 0.01$, strain valency
232 $(0, 7, 31, 127)$ (Fig. 6D). Each of these vaccination programs begin with the initial
233 strain 0 $\equiv (0, 0, 0, 0, 0, 0, 0)$ being replaced by strain 127 $\equiv (1, 1, 1, 1, 1, 1, 1)$
234 towards the end of year 1 or start of year 2.

235 Alarming, the first vaccination rollout scenario (Case B) made the epi-
236 demic worse rather than better when compared with the no vaccination scenario
237 (number of deaths were 1366 versus 1122; compare Figs. 6A and B) because the
238 vaccination rollout with its strain 0 valence vaccine, though delaying the emer-
239 gence of strain 127, led to an increased prevalence of this most transmissible
240 strain ($\beta_{127} = 0.475$ versus $\beta_0 = 0.300$). Using a strain 127 valence vacci-
241 nation (Case C) had some effect in reducing the 3-year death total back to
242 1109, but the efficacy of this rollout was thwarted by the emergence of strain
243 48 $\equiv (0, 1, 1, 0, 0, 0, 0)$ ($\beta_{48} = 0.295$, Fig. 6C).

244 Given the inability of these two monovalent vaccination rollout programs in
245 years 2 and 3 to control the outbreak, we simulated the much more aggressive
246 rollout program of a quadravalent $(0, 7, 31, 127)$ vaccine applied at a rate of 1% per
247 day ($v = 0.01$; a five-fold increase over the previous two monovalent programs)
248 in a system with increased levels of cross immunity (cascading versus simple)
249 and including revaccination of previously vaccinated individuals to combat the
250 effects of waning immunity. Even this aggressive vaccination rollout program
251 failed, although it did knock the incidence way down at the start of the rollout.
252 The efficacy of this rollout, however, was thwarted by the rise of strain 23 \equiv
253 $(0, 0, 1, 0, 1, 1, 1)$ ($\beta_{23} = 0.214$) during the mid to latter part of the second year,
254 which was then supplanted by strain 3 $\equiv (0, 0, 0, 0, 0, 1, 1)$ ($\beta_3 = 0.350$), which
255 dominated the third year (Fig. 6D).

256 Unsurprisingly, the success of any vaccination program depends on the rela-
257 tive robustness of the immunity characteristics of hosts to the various strains.
258 Thus, in regard to the vaccination programs represented by cases B and C
259 (Fig. 6, we simulated the rollout of a bivalent vaccine with strain valency $(0, 127)$
260 implemented at the same vaccination rate of $v = 0.002$, but in a system with
261 much more robust immunity characteristics: a waning half-life increased from
262 $t^{\text{half}} = 365$ to 1825 days (a five-fold increase) and a cascading cross immunity
263 system with c increased from 0.5 to 0.9. In this case, the outbreak was rapidly
264 extirpated after 257 days (i.e., from start at day 365 to extirpation at day 622).

265 Given the failure of our first three 128 strain vaccination rollouts (cases B
266 to D in Fig. 6) to extirpate the simulated epidemics, we investigated whether
267 rollout programs that adapted vaccine valencies to track in real time the emer-

268 gence of dominant strains might be more effective. Thus, we used a JS scripting
269 window to write a JavaScript driver (as described in Appendix B; see Fig. B.4)
270 that adapted (i.e., switched as necessary) the valency of the vaccine every 90
271 days after the start of the vaccination rollout to match the top two strains
272 prevalent at time of switching, or only the top strain if no other strain exceeded
273 a prevalence of 10 individuals. The results of these two simulations, one for the
274 nearest-neighbor cross-immunity case (Eq. 1) and one for the cascading cross-
275 immunity case (Eq. A.7), are provided in Fig. 7. They should be compared
276 with the non-vaccination case depicted in Fig. 6A and the non adaptive vac-
277 cination rollout depicted in Fig. 6D. In contrast with the latter—and despite
278 having the same vaccination rates, the same strategies for selecting non-infected
279 individuals at random, and employing vaccines with fewer strain valency—both
280 of these adaptive vaccination rollouts extirpated the epidemic, the first within
281 half a year (Case A) and the second within 1.4 years (Case B) from the start of
282 vaccination rollout.

283 Discussion

284 The amount of structure and data needed in complex biological systems' mod-
285 els, depends on the questions that these models have been formulated to ad-
286 dress [26,37]. In this paper, we steered away from making specific predictions—
287 because universal solutions are not always locally applicable. Rather, we focused
288 on questions relating to the potential efficacy of different vaccination rollouts
289 (both vaccination rates and valencies of vaccines) in the context of strain emer-
290 gence and the potential for vaccination programs to go awry. To make specific
291 predictions requires localized data, particularly as it may relate to social dis-
292 tancing behavior or other factors that affect contact rates over time [30,38,39].
293 Additionally, we cannot expect models to make strain-specific informative pre-
294 dictions, unless they have been designed to do so and are supported by location
295 specific data regarding the relative transmissibility of strains and other strain
296 specific data. In the absence of such data, models become a tool for anticipat-
297 ing pitfalls and avoiding unintended consequences of well-intentioned actions.
298 Equally important, however, is the implementation of our model as a RAMP
299 (runtime alterable model platform), because this greatly facilitates the use of
300 our model by ourselves and others in testing out different hypothesis about the
301 process generating observed population level strain transmission dynamics.

302 The illustrative simulations we present demonstrate that if vaccination roll-
303 out rates and vaccine valencies are applied in blanket manner—that is, without
304 monitoring strain emergence and obtaining some assessment of the epidemio-
305 logical characteristics of those strains (e.g., relative transmissibility, virulence,
306 waning and cross immunity characteristics)—then the possibility exists for a
307 vaccination program to do more harm than good. This of course, is not to say
308 that vaccination programs should ever be avoided; only that once a vaccination
309 program is embarked upon, it should be implemented at a sufficiently vigorous
310 rate to ensure that it is not outrun by relatively fast waning immunity rates.
311 Additionally, when escape mutations emerge, generally associated with reinfec-
312 tion cases [8], then the valency of vaccines should be switched rapidly enough to

313 thwart escape mutations, rather than aiding such mutations by thwarting less
314 transmissible competing strains instead.

315 At this time, the primary value of our SEIVD-IBM J-RAMP is in testing var-
316 ious vaccination strategies as they relate to strain emergence [40] and identifying
317 pitfalls related to the unintended promotion of more transmissible or virulent
318 strains. Strain emergence has both a local and global component. Particu-
319 lar strains may first be identified as occurring in specific geographical regions,
320 as in the strains that appear to have originated in the UK (B.1.7.7 variant),
321 South Africa (variants B.1.1.54, B.1.1.56 and the C.1 lineage) and Brazil (P.1
322 lineage) [41]. Thus more geographically focused application of our J-RAMP
323 may require more specific strain related information. Such information would
324 then be used to get estimates of the relative values of transmissibility β_j , viru-
325 lence α_j , and even of shedding ($\bar{\zeta}_{jm}$) and environmental persistence ($\bar{\eta}_j$). But
326 equally important in evaluating the impacts of vaccination strategies on local
327 outbreaks is obtaining strain-specific cross-immunity data (for characterization
328 of the elements $c_{j\ell}$ of the cross immunity matrix \mathbf{C}), waning rates (which we
329 have not made strain specific, but our model could be generalized to include
330 strain specific values t_j^{half}), and the relative within host competition values (i.e.,
331 λ_j , which we have set all equal to 1 at this time in the absence of data to the
332 contrary). Models are sorely needed to explore multistrain dynamics, particu-
333 larly the epidemiological properties with regard to shedding, environmental
334 persistence, transmission, mutation, and within-host replication rates, which
335 acting together determine the relative success of different strains and their ac-
336 tual impact on the severity of epidemics and the nature of vaccination programs
337 needed to suppress them.

338 The process of making our model both location and strain specific could be
339 undertaken using methods designed to enhance the relevancy of models, such
340 as Appropriate Complexity Modeling (ACM: [26,37]). Further, in some cases it
341 may be useful to add spatial or age-structure information to our SEIVD model
342 IBMs or include a contact network [42], which itself may contain spatial or re-
343 fined class category (e.g. age or work category) information. In addition, our
344 current implementation represents strain differences in terms of J loci with two
345 alleles (denoted by 0 and 1 respectively) at each locus. A more realistic repre-
346 sentation of the genetic basis of strain differences may involve a more involved
347 genetic representation in which several alleles are possible at each locus.

348 Although cross-immunity and immune waning are entangled in our immu-
349 nity modifier functions (i.e., ϕ_{ij} ; see Eq. A.11), cross-neutralization data can
350 be used to estimate the cross and waning immunity parameters using appro-
351 priate methods [43,44]. Such data are becoming more widely available through
352 the application of rapid PCR methods [3,45,46]. Strain or variant cross neu-
353 tralizing studies bring up a much neglected issue, which is the effect of dose
354 (number of pathogens involved in the initial infection, also know as viral load)
355 on the severity of the infection. This differs from the questions of the number
356 of doses—usually one or two—constituting and vaccine regimen versus the viral
357 load (or antigen load, or virus-like particle load when only antigens or parts of
358 viruses are administered) in each dose [4]. In the context of vaccination, both
359 these issues and the technology used to produce the vaccine [23] may well have
360 an impact on waning immunity half-lives and cross-immunity values. Thus,

361 the parameter values used in the model should ultimately be vaccine specific.
362 Further, when it comes to determining waning and cross immunity paramete-
363 ter values, our model does not distinguish between individuals that have been
364 infected with a particular strain or vaccinated with a valency related to that
365 strain.

366 In the coming year, as we obtain more information on the nature of immunity
367 to SARS-CoV-2, it will become more apparent to us whether or not COVID-19
368 will settle into global endemicity [19, 47]. If this is the case, then constant vigi-
369 lance and a well designed vaccination program with respect to vaccinating the
370 young and implementing booster vaccinations with appropriate strain valency
371 will become the order of the day. The J-RAMP presented here, with appropri-
372 ate elaborations that will become evident through its future application, such as
373 being able to compute the best time to administer booster shots of the same or
374 different strain valencies to individuals, should play a decisive role in the ratio-
375 nal design of effective and efficient COVID-19 vaccination programs worldwide.
376 This is made apparent from the simulations we present here that, depending on
377 the immunological characteristics of hosts and mutational rates of SARS-Cov-
378 2, it may be harder than currently anticipated to extinguish COVID-19 in the
379 next few years through non-adaptive vaccination programs. Thus, we should
380 endeavor to adapt the strain valency of our vaccines as rapidly as possible to
381 the predominating SARS-Cov-2 strains, identified from ongoing monitoring and
382 surveillance programs.

383 Finally, the concept of runtime-alterable, modeling platforms (RAMPs),
384 with driver script and other coding platform integration, in our case devel-
385 oped in Java with R platform integration and a JavaScript simulation driver
386 window, provides the first example of a new concept in model implementation
387 that facilitates model sharing and easy modification by users other than the
388 original developers. We believe such platforms can come to play an important
389 role not only in disease modeling, but in all fields of research that rely on models
390 for comprehensive analyses of the behavior of systems of interest.

391 **Materials and Methods**

392 **Our SEIVD-IBM in a nutshell**

393 We constructed an individual-based model (IBM) of a susceptible-exposed-
394 infectious-recovered (i.e., an SEIVD model, where removed R are split into
395 V =immune/vaccinated, and D =dead) epidemiological process [11, 12] in a ho-
396 mogeneous population with a random encounter contact rate parameter $\kappa_0 > 0$.
397 Our formulation includes a host immunological waning process [6, 19]. It also
398 includes the emergence of pathogen strains due to a mutational process that
399 impacts both transmission of mutant strains from the infectee and genetic
400 drift [3, 5, 48] of strains within the infector, with rates impacted by cross im-
401 munity effects. We allowed for variation in pathogen strain transmissibility
402 (i.e., in the $\beta > 0$ parameter of the frequency dependent transmission function
403 $\beta SI/N$ [49, 50]) and pathogen virulence as represented by the disease-induced
404 host mortality rate sensu Anderson and May [51] (and often represented by a
405 parameter $\alpha \geq 0$ [49]).

406 The detailed formulation of our model and its algorithmic implementation
407 is provided in Appendix A. In a nutshell we:

- 408 1. defined a set of 2^J pathogen strains (user selected value for J ranging from
409 0 to 8; pathogen index $j = 0, \dots, 2^J - 1$) with a genetic-relatedness topology
410 of a J -dimensional unit cube—i.e., each pathogen has J -loci that can take
411 on one of two allelic values at each locus with immediate neighboring
412 strains differing from each other by exactly one allelic value (0 or 1) at
413 only one of the J loci
- 414 2. defined a population of N_0 hosts as belonging at time t to either an epi-
415 demologically naïve set of susceptible individuals \mathbf{S} of size $N_S(t)$, a set
416 \mathbf{A} of $N_A(t)$ identified agents A_i ($i = 1, \dots, N_A(t)$) whose epidemiological
417 histories are known, or a set \mathbf{D} of $N_D(t)$ individuals that have died from
418 the disease
- 419 3. allow pathogen strain-specific transmission “force” ($\bar{\beta}_j > 0$) and virulence
420 ($\alpha_j \geq 0$) parameters to vary in value among one another within a defined
421 range $\bar{\beta}_j \in [\beta_{\min}, \beta_{\max}]$ and $\alpha_j \in [\alpha_{\min}, \alpha_{\max}]$
- 422 4. kept track of the total prevalence N_I as the sum of the prevalences of the
423 individual strains N_{I_j} , $j = 0, \dots, 2^J - 1$ —i.e. $N_I = \sum_{j=0}^{2^J-1} N_{I_j}$
- 424 5. introduced a random contact rate function $\kappa(t)$ with a constant param-
425 eter κ_0 that is Poisson distributed on $[t, t + 1)$, $t = 0, 1, \dots$, multiplied
426 by an adaptive response function that reduces the contact rate with in-
427 creasing disease prevalence, such that the $\kappa(t)$ is reduced to $\kappa_0/2$ when
428 the $N_I(t)/(N_0 - N_D) = p^{\text{half}}$ —see Eq. A.10 in Appendix A
- 429 6. update the epidemiological state of the agents A_i with respect to each
430 of the strains $j = 0, \dots, 2^J - 1$ where the state with respect to particular
431 strain j at time t is represented by
 - 432 (a) 0: has never been infected with this strain
 - 433 (b) $E_j(t, \tau_{ij})$: infected at time $\tau_{ij} \leq t$ with this strain, but not yet infec-
434 tious for an expected period of σ_E units of time
 - 435 (c) $I_j(t, \tau_{ij})$: infectious at time t with this strain, for an expected period
436 of σ_I units of time, having been most recently infected (reinfections
437 with the same strain may occur) with this strain at time $\tau_{ij} < t$
 - 438 (d) $V_j(t, \tau_{ij})$: has now recovered from its most recent infection at time
439 τ_{ij} with this strain and has some level of waning immunity to it
440 where we assume that agent A_i can be infected at time t with at most one
441 dominant strain (denoted by the index j), although it will have different
442 levels of waning immunity to all of the strains to which it has been infected
443 in the past
- 444 7. included waning immunity functions $\omega_{ij}(t)$ (symbol is omega: Eq. A.6)
445 used to compute the level of immunity that agent A_i has to its most
446 recent infection by strain j
- 447 8. included cross immunity constraints (a J^2 -matrix C) that apply both to
448 the *infecter* transmitting the pathogen and the *infectee* being invaded (inv)
449 by the pathogen, both of which reduce the likelihood of infection and strain
450 drift by strain j compared with closely related strains ℓ
- 451 9. computed an *infection probability* $\pi_{ih,j\ell}^{\text{inf}}$ that agent A_i infected with strain
452 j infects agent A_h with strain ℓ in terms of a concatenation of infecter
453 viral shedding ($\zeta_{i\ell}$), viral persistence in the environment (ε_ℓ), and viral
454 transmission ($\beta_{h\ell}$) processes (Fig. 1)

- 455 10. computed an *invasion probability* $\pi_{h\ell\ell'}^{\text{inv}}$ that an agent A_h infected with
456 strain ℓ becomes infectious with strain ℓ' as its major strain, in terms
457 of the multiplicative effects of viral mutation (μ) and viral replication
458 (λ_ℓ) processes ongoing within an infectee A_h during this infectees exposed
459 ($E_{\ell'\tau_{h\ell'}}$) and infectious ($I_{\ell'\tau_{h\ell'}}$) stages (Fig. 1)
- 460 11. computed the overall probability $\pi_{ih,j\ell'}$ that that an infector A_i infected
461 with major strain j results in an infectee A_h expressing ℓ' as its major
462 strain
- 463 12. implemented a discrete time individual-based stochastic SEIVD (here V
464 represents individuals that have either recovered from infection or have
465 been vaccinated, D represents cumulative dead) multi-strain model that
466 includes specifiable time-dependent univalent and multivalent vaccination
467 applications

468 **Our J-RAMP implementation**

469 Models of systems process can be coded as singular implementations model for-
470 mulations using: i) highly efficiently compilable computer languages (e.g., C++,
471 FORTRAN, Java) ; or ii) less efficiently, but more easily coded, scriptable (e.g.,
472 JavaScript, Python, Perl) computer languages. More conveniently and expedi-
473 tiously, they can be coded up, as discussed in [52], using a systems modeling
474 platform, such as Matlab's SIMULINK, Mathematica, Stella, AnyLogic, Num-
475 erus, or Berkeley Madonna. Advantages of the latter include more rapid and
476 accurate model development, though simulations may be slowed down by plat-
477 form overhead. Between these extremes, we propose a more general approach
478 to specific classes of systems' models, where the basic system structure is fixed,
479 but implementation of some elements can be easily and safely altered so that
480 optional implementations are presented at runtime. We call such a design *run-*
481 *time alterable-model platforms*. (RAMPs); and here we present a Java RAMP
482 implementation of the SEIVD-IBM described in the previous subsection.

- 483 The characteristics we envision for a model platform to be a RAMP are:
- 484 1. RAMPs include a set of model parameters (constants) whose values can
485 be selected or specified (sometimes within a predefined range of values)
486 at simulation runtime using a switch, nob, slider, or text-entry window
487 accessed via a platform graphical interface or dashboard (e.g., see Fig. 2
488 and Table 1 which apply to our SEIVD-IBM J-RAMP).
 - 489 2. RAMPs include a specific set of *runtime alternative modules*, (RAMs),
490 where the original can be redefined in a graphical interface window, and
491 the unaltered original and the alternative routines are stored as a (prefer-
492 ably open-ended) numbered set. The original or any one of the alternatives
493 can be selected for use at runtime (for a list of functions in our RAMP
494 see Table 2).
 - 495 3. RAMP implementations also provide an API for both remote and on-
496 board scripting. This API enables control of all user aspects of the simu-
497 lation, including the parameter set, run management, RAM options, and
498 data retrieval. Script logic can alter parameter settings and RAM options
499 as the simulation progresses. A Nashorn-based Javascript interpreter en-
500 hanced with API methods is provided.
 - 501 4. The API can be accessed remotely using operating system facilities by
502 external applications running concurrently with the simulator. Of particu-
503 lar interest is the ability to control the simulation from the R statistical
504 platform. An R routine can be formulated to both manage the simulation
505 run and to retrieve and process the resulting data. The RAMP simulation
506 becomes a “virtual package” to the controlling R logic. See Appendix B.

507 We implemented our RAMP using Java and made ample use of all of the fea-
508 tures described above. Use of the RAM facility permitted experimentation with
509 the several versions of cross immunity presented in this paper. A Javascript pro-
510 gram was used to control an adaptive vaccination strategy. A small R package
511 serving as a driver was used in an R program that ran the simulation multiple
512 times, extracted results into R data structures, and produced graphs showing
513 statistical mean and standard deviation. More details on the graphical structure
514 and implementation of our SEIVD-IBM are both implied in the presentation of
515 results below, and in additional details provided in Appendix B.

516 COVID-19 model assumptions and parameter values

517 The first variable that needs to be determined is the unit of time we use for
518 our simulations because all process rate parameters are scaled by its selection.
519 Since the time resolution of empirical COVID-19 incidence and mortality data
520 is daily, we selected as unit of time t to be days. Additionally, based on various
521 sources including a metapopulation study of COVID-19 parameter estimates
522 [53], we set the latent and infectious periods to be 4 and 7 days respectively.
523 Basic SEIR epidemiological models do not separate out the processes of contact
524 and transmission-per-contact, so we had some leeway on what values to choose
525 for contact rates and transmission rates per contact because it is the value of
526 the product of these that is important in determining the reproductive value,
527 commonly referred to as “ R_0 ” for COVID-19. Hussein et al.’s [53] meta analysis
528 of COVID-19 zeroed in on $R_0 = 3.14$ as a mean value across a range of studies

529 (95% confidence interval [2.69, 3.59]) though these authors suggested that the
530 most recent studies they looked at suggested that R_0 was more likely between
531 2 and 3.

532 Assuming no seasonal effects (i.e., $\delta = 0$ in Eq. 3), in our single strain runs
533 we kept $R_0 \approx 3$ by setting $\kappa = 4$ (effective contacts per day, which nominally
534 involves being with 6 feet of one another at least 15 mins) thereby requiring
535 $\beta = 0.3$, because in the notation of the continuous time model introduced in
536 Eq. A.29, we have an inverse latent period value of $\gamma \equiv 1/\sigma_E = 1/4$ and an
537 inverse infectious period value of $\rho \equiv 1/\sigma_I = 1/7$ days (Table 1) for Eq. A.29 to
538 yield $R_0 = \frac{4 \times 0.3}{1/4 + 1/7} = 3.05$

539 In our multistrain runs, we did allow for the emergence of strains that were
540 more transmissible than this single strain value (in these runs new strains were
541 in the range [0,0.475] (see table depicted in Fig. C.4). Also, although the incu-
542 bation period is typically thought to have a median of 5 days, it can be much
543 faster [54].

544 In reality, the contact rate κ is not constant, but is time varying. It is influ-
545 enced by the social distancing behavior of individuals who may be responding
546 to fears of contracting the disease, or regulations imposed by local authorities
547 to reduce effective contact rates. To account for this, we formulated an adap-
548 tive contact rate function $\kappa(t)$ (Eq. A.10) in terms of a function that declines
549 adaptively from a non-epidemic (baseline maximum) value κ_0 to $\kappa_0/2$ when the
550 prevalence (the percentage of the population infectious at time t) increases to
551 $\frac{N_I(t)}{N(t) - N_D(t)} = p_I^{\text{half}}$. In the illustrative simulations presented here, we explore the
552 effect of varying p_I^{half} on the size of the shape of epidemic, but set $p_I^{\text{half}} = 0.002$
553 (i.e., a prevalence level of 0.2%) as a baseline value, because in our simulations
554 this baseline value, along with the parameter values selected above, results in a
555 cumulative number of cases to just over 11% in a population of $N = 100,000$
556 (Appendix C, Fig. C.3). This puts our illustrative baseline case in the ballpark
557 of the US outbreak which [Worldometer](#) reported to have a cumulative case level
558 of 8.5% ($\approx 284.5 \times 10^6 / 332.9 \times 10^6$) after one year, but this is likely to be a
559 substantial underestimate of the actual number of cases [55].

560 To keep things simple in applying our model to a multistrain setting, we
561 assumed that the immune effects on infector shedding and mutation and repli-
562 cation within the infectee through a cross-immunity processes, can be tweaked
563 through a single constant $c \in [0, 1)$ (making the elements $c_{j\ell}$ of the matrix
564 C strain-dependent will obviously require considerable supporting data). One
565 scenario is to assume cross immunity only applies to nearest neighbors, i.e.:

$$\text{Nearest-neighbor } \mathbf{C}: \quad c_{j\ell} = \begin{cases} 1 & \text{if } \ell = j \\ c & \text{if } \ell \text{ differs from } j \text{ by one allele} \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

566 For other formulations, such as “Cascading \mathbf{C} ” or including escape mutations
567 in the mix see Eqs. A.7 and A.9 in Appendix A.

568 Additionally, we expect this type of evidence to become increasing available
569 with whole genome sequence of pathogen strains [56] collected during host shed-
570 ding. For simplicity’s sake, however, we assume that infectee with major strain
571 j will shed minor strains in the immediate neighborhood of j at comparative

572 rate $\zeta \in [0, 1)$ and be major strain-independent: i.e., we assume

$$\zeta_{j\ell} = \begin{cases} 1 & \text{if } \ell = j \\ \zeta & \text{if } \ell \text{ differs from } j \text{ by one allele} \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

573 Third, again in the absence of evidence to the contrary, we assume that
574 seasonal fluctuations are strain independent. In this case, for strain-independent
575 constants $\delta \in [0, 1)$ and $\theta \in [-\pi, \pi]$ that set seasonal fluctuations related to
576 environmental persistence, Eq. A.13 in the Appendix A simplifies for a single
577 constant $\bar{\eta}$ to:

$$\eta_{\ell}(t) = \bar{\eta} \left(1 + \delta \sin \left(\frac{2\pi t}{365} + \theta \right) \right) \quad \forall \ell \quad (3)$$

578 In our multistrain simulations, for purposes of illustration, we set both the
579 shedding parameter ζ in Eq. 2 above and μ in Eq. A.16 to be 0.001. Obviously, it
580 would be of considerable interest to know how the values of these two parameters
581 impact strain proliferation, but this is left to future studies when data are
582 available to guide simulations of realistic scenarios.

583 Acknowledgements and Funding Sources

584 This work was funded in part by NSF Grant 2032264 (PI: WMG).

References

- 585
- 586 [1] Anderson RM. The impact of vaccination on the epidemiology of infectious
587 diseases. In: *The Vaccine Book*. Elsevier; 2016. p. 3–31.
- 588 [2] Kohli M, Maschio M, Becker D, Weinstein MC. The potential public health
589 and economic value of a hypothetical COVID-19 vaccine in the United
590 States: Use of cost-effectiveness modeling to inform vaccination prioritiza-
591 tion. *Vaccine*. 2021;39(7):1157–1164.
- 592 [3] Eguia RT, Crawford KH, Stevens-Ayers T, Kelnhofer-Millevolte L,
593 Greninger AL, Englund JA, et al. A human coronavirus evolves antigeni-
594 cally to escape antibody immunity. *PLoS pathogens*. 2021;17(4):e1009453.
- 595 [4] Jeyanathan M, Afkhami S, Smaill F, Miller MS, Lichty BD, Xing Z. Im-
596 munological considerations for COVID-19 vaccine strategies. *Nature Re-
597 views Immunology*. 2020;20(10):615–632.
- 598 [5] Koopman JS, Simon CP, Getz WM, Salter R. Modeling the population
599 effects of epitope specific escape mutations in SARS-CoV-2 to guide vacci-
600 nation strategies. *medRxiv*. 2021;.
- 601 [6] Sariol A, Perlman S. Lessons for COVID-19 immunity from other coron-
602 avirus infections. *Immunity*. 2020;.
- 603 [7] Tai W, Wang Y, Fett CA, Zhao G, Li F, Perlman S, et al. Recombinant
604 receptor-binding domains of multiple Middle East respiratory syndrome
605 coronaviruses (MERS-CoVs) induce cross-neutralizing antibodies against
606 divergent human and camel MERS-CoVs and antibody escape mutants.
607 *Journal of virology*. 2017;91(1).
- 608 [8] Harvey WT, Carabelli AM, Jackson B, Gupta RK, Thomson EC, Harrison
609 EM, et al. SARS-CoV-2 variants, spike mutations and immune escape.
610 *Nature Reviews Microbiology*. 2021;p. 1–16.
- 611 [9] Saad-Roy CM, Levin SA, Metcalf CJE, Grenfell BT. Trajectory of in-
612 dividual immunity and vaccination required for SARS-CoV-2 community
613 immunity: A conceptual investigation. *Journal of the Royal Society Inter-
614 face*. 2021;18(175):20200683.
- 615 [10] Saad-Roy CM, Wagner CE, Baker RE, Morris SE, Farrar J, Graham AL,
616 et al. Immune life history, vaccination, and the dynamics of SARS-CoV-2
617 over the next 5 years. *Science*. 2020;370(6518):811–818.
- 618 [11] Willem L, Verelst F, Bilcke J, Hens N, Beutels P. Lessons from a decade of
619 individual-based models for infectious disease transmission: a systematic
620 review (2006-2015). *BMC infectious diseases*. 2017;17(1):612.
- 621 [12] Chowell G, Sattenspiel L, Bansal S, Viboud C. Mathematical models to
622 characterize early epidemic growth: A review. *Physics of life reviews*.
623 2016;18:66–97.

- 624 [13] Keskinocak P, Oruc BE, Baxter A, Asplund J, Serban N. The impact of
625 social distancing on COVID19 spread: State of Georgia case study. *Plos*
626 *one*. 2020;15(10):e0239798.
- 627 [14] Abueg M, Hinch R, Wu N, Liu L, Probert W, Wu A, et al. Modeling
628 the effect of exposure notification and non-pharmaceutical interventions
629 on COVID-19 transmission in Washington state. *NPJ digital medicine*.
630 2021;4(1):1–10.
- 631 [15] Chen PZ, Bobrovitz N, Premji Z, Koopmans M, Fisman DN, Gu FX. Het-
632 erogeneity in transmissibility and shedding SARS-CoV-2 via droplets and
633 aerosols. *eLife*. 2021 apr;10:e65774. Available from: [https://doi.org/](https://doi.org/10.7554/eLife.65774)
634 [10.7554/eLife.65774](https://doi.org/10.7554/eLife.65774).
- 635 [16] van Doremalen N, Bushmaker T, Morris DH, Holbrook MG, Gamble
636 A, Williamson BN, et al. Aerosol and Surface Stability of SARS-CoV-
637 2 as Compared with SARS-CoV-1. *New England Journal of Medicine*.
638 2020;382(16):1564–1567. Available from: [https://doi.org/10.1056/](https://doi.org/10.1056/NEJMc2004973)
639 [NEJMc2004973](https://doi.org/10.1056/NEJMc2004973).
- 640 [17] Maginnis MS. Virus–receptor interactions: the key to cellular invasion.
641 *Journal of molecular biology*. 2018;430(17):2590–2611.
- 642 [18] Challen R, Brooks-Pollock E, Read JM, Dyson L, Tsaneva-Atanasova K,
643 Danon L. Risk of mortality in patients infected with SARS-CoV-2 variant
644 of concern 202012/1: matched cohort study. *bmj*. 2021;372.
- 645 [19] Lavine JS, Bjornstad ON, Antia R. Immunological characteristics govern
646 the transition of COVID-19 to endemicity. *Science*. 2021;371(6530):741–
647 745.
- 648 [20] Alizon S, Luciani F, Regoes RR. Epidemiological and clinical consequences
649 of within-host evolution. *Trends in microbiology*. 2011;19(1):24–32.
- 650 [21] Fenichel EP, Castillo-Chavez C, Ceddia MG, Chowell G, Parra PAG, Hick-
651 ling GJ, et al. Adaptive human behavior in epidemiological models. *Proce-*
652 *edings of the National Academy of Sciences*. 2011;108(15):6306–6311.
- 653 [22] Gozzi N, Bajardi P, Perra N. The importance of non-pharmaceutical in-
654 terventions during the COVID-19 vaccine rollout. *medRxiv*. 2021; Avail-
655 able from: [https://www.medrxiv.org/content/early/2021/01/09/](https://www.medrxiv.org/content/early/2021/01/09/2021.01.09.21249480)
656 [2021.01.09.21249480](https://www.medrxiv.org/content/early/2021/01/09/2021.01.09.21249480).
- 657 [23] Kumar A, Dowling WE, Román RG, Chaudhari A, Gurry C, Le TT, et al.
658 Status Report on COVID-19 Vaccines Development. *Current Infectious*
659 *Disease Reports*. 2021;23(6):1–12.
- 660 [24] Chretien JP, George D, Shaman J, Chitale RA, McKenzie FE. In-
661 fluenza forecasting in human populations: a scoping review. *PloS one*.
662 2014;9(4):e94130.

- 663 [25] Getz WM, Gonzalez JP, Salter R, Bangura J, Carlson C, Coomber M,
664 et al. Tactics and strategies for managing Ebola outbreaks and the salience
665 of immunization. *Computational and mathematical methods in medicine*.
666 2015;2015.
- 667 [26] Getz WM, Salter R, Mgbara W. Adequacy of SEIR models when epidemics
668 have spatial structure: Ebola in Sierra Leone. *Philosophical Transactions
669 of the Royal Society B*. 2019;374(1775):20180282.
- 670 [27] Drake JM. Limits to forecasting precision for outbreaks of directly trans-
671 mitted diseases. *PLoS Med*. 2005;3(1):e3.
- 672 [28] Jefferies JM, Clarke SC, Webb JS, Kraaijeveld AR. Risk of red queen
673 dynamics in pneumococcal vaccine strategy. *Trends in microbiology*.
674 2011;19(8):377–381.
- 675 [29] Thomas LJ, Huang P, Yin F, Luo XI, Almquist ZW, Hipp JR, et al. Spa-
676 tial heterogeneity can lead to substantial local variations in COVID-19
677 timing and severity. *Proceedings of the National Academy of Sciences*.
678 2020;117(39):24180–24187.
- 679 [30] Getz WM, Salter R, Luisa Vissat L, Horvitz N. A versatile web app for
680 identifying the drivers of COVID-19 epidemics. *Journal of Translational
681 Medicine*. 2021;19(1):1–20.
- 682 [31] Da Silva RM, Mendes CF, Manchein C. Scrutinizing the heterogeneous
683 spreading of COVID-19 outbreak in large territorial countries. *Physical
684 Biology*. 2021;18(2):025002.
- 685 [32] Zhang J, Litvinova M, Liang Y, Wang Y, Wang W, Zhao S, et al. Changes in
686 contact patterns shape the dynamics of the COVID-19 outbreak in China.
687 *Science*. 2020;.
- 688 [33] Udwardia ZF, Raju RS. How to protect the protectors: 10 lessons to learn
689 for doctors fighting the COVID-19 coronavirus. *Medical Journal, Armed
690 Forces India*. 2020;.
- 691 [34] Van Damme W, Dahake R, van de Pas R, Vanham G, Assefa Y. COVID-19:
692 Does the infectious inoculum dose-response relationship contribute to un-
693 derstanding heterogeneity in disease severity and transmission dynamics?
694 *Medical hypotheses*. 2021;146:110431.
- 695 [35] Zhao Z, Li H, Wu X, Zhong Y, Zhang K, Zhang YP, et al. Moderate
696 mutation rate in the SARS coronavirus genome and its implications. *BMC
697 evolutionary biology*. 2004;4(1):1–9.
- 698 [36] Geoghegan JL, Holmes EC. The phylogenomics of evolving virus virulence.
699 *Nature Reviews Genetics*. 2018;19(12):756–769.
- 700 [37] Larsen LG, Eppinga MB, Passalacqua P, Getz WM, Rose KA, Liang
701 M. Appropriate complexity landscape modeling. *Earth-science reviews*.
702 2016;160:111–130.

- 703 [38] Yang P, Yang G, Qi J, Sheng B, Yang Y, Zhang S, et al. The effect of mul-
704 tiple interventions to balance healthcare demand for controlling COVID-19
705 outbreaks: a modelling study. *Scientific reports*. 2021;11(1):1–13.
- 706 [39] Baker RE, Park SW, Yang W, Vecchi GA, Metcalf CJE, Grenfell BT.
707 The impact of COVID-19 nonpharmaceutical interventions on the future
708 dynamics of endemic infections. *Proceedings of the National Academy of*
709 *Sciences*. 2020;117(48):30547–30553.
- 710 [40] Greaney AJ, Starr TN, Gilchuk P, Zost SJ, Binshtein E, Loes AN,
711 et al. Complete mapping of mutations to the SARS-CoV-2 spike receptor-
712 binding domain that escape antibody recognition. *Cell host & microbe*.
713 2021;29(1):44–57.
- 714 [41] Brüssow H. COVID-19: emergence and mutational diversification of SARS-
715 CoV-2. *Microbial Biotechnology*. 2021;14(3):756–768.
- 716 [42] Keeling MJ, Eames KT. Networks and epidemic models. *Journal of the*
717 *Royal Society Interface*. 2005;2(4):295–307.
- 718 [43] Roberts GO, Rosenthal JS, et al. General state space Markov chains and
719 MCMC algorithms. *Probability Surveys*. 2004;1:20–71.
- 720 [44] Lele SR, Dennis B, Lutscher F. Data cloning: easy maximum likelihood es-
721 timation for complex ecological models using Bayesian Markov chain Monte
722 Carlo methods. *Ecology letters*. 2007;10(7):551–563.
- 723 [45] Suthar MS, Zimmerman MG, Kauffman RC, Mantus G, Linderman SL,
724 Hudson WH, et al. Rapid generation of neutralizing antibody responses in
725 COVID-19 patients. *Cell Reports Medicine*. 2020;1(3):100040.
- 726 [46] Muruato AE, Fontes-Garfias CR, Ren P, Garcia-Blanco MA, Menach-
727 ery VD, Xie X, et al. A high-throughput neutralizing antibody assay
728 for COVID-19 diagnosis and vaccine evaluation. *Nature communications*.
729 2020;11(1):1–6.
- 730 [47] Tkachenko AV, Maslov S, Elbanna A, Wong GN, Weiner ZJ, Goldenfeld
731 N. Time-dependent heterogeneity leads to transient suppression of the
732 COVID-19 epidemic, not herd immunity. *Proceedings of the National*
733 *Academy of Sciences*. 2021;118(17).
- 734 [48] Koyama T, Weeraratne D, Snowdon JL, Parida L. Emergence of drift vari-
735 ants that may affect COVID-19 vaccine development and antibody treat-
736 ment. *Pathogens*. 2020;9(5):324.
- 737 [49] Getz WM, Lloyd-Smith JO. Basic methods for modeling the invasion and
738 spread of contagious diseases. *DIMACS Series in Discrete Mathematics*
739 *and Theoretical Computer Science*. 2006;71:87.
- 740 [50] McCallum H, Barlow N, Hone J. How should pathogen transmission be
741 modelled? *Trends in ecology & evolution*. 2001;16(6):295–300.

- 742 [51] Anderson RM, May RM, et al. Coevolution of hosts and parasites. *Parasitology*. 1982;85(Pt 2):411–426.
743
- 744 [52] Getz WM, Salter R, Lyons AJ, Sippl-Swezey N. Panmictic and clonal
745 evolution on a single patchy resource produces polymorphic foraging guilds.
746 *PloS one*. 2015;10(8):e0133732.
- 747 [53] Hussein M, Toraih E, Elshazli R, Fawzy M, Houghton A, Tatum D, et al.
748 Meta-analysis on serial intervals and reproductive rates for SARS-CoV-2.
749 *Annals of surgery*. 2021;273(3):416–423.
- 750 [54] Grimm V, Mengel F, Schmidt M. Extensions of the SEIR model for the
751 analysis of tailored social distancing and tracing approaches to cope with
752 COVID-19. *Scientific Reports*. 2021;11(1):1–16.
- 753 [55] Wu SL, Mertens AN, Crider YS, Nguyen A, Pokpongkiat NN, Djajadi S,
754 et al. Substantial underestimation of SARS-CoV-2 infection in the United
755 States. *Nature communications*. 2020;11(1):1–10.
- 756 [56] Munnink BBO, Nieuwenhuijse DF, Stein M, O’Toole Á, Haverkate M,
757 Mollers M, et al. Rapid SARS-CoV-2 whole-genome sequencing and analy-
758 sis for informed public health decision-making in the Netherlands. *Nature*
759 *medicine*. 2020;26(9):1405–1410.
- 760 [57] Kramer A, Schwebke I, Kampf G. How long do nosocomial pathogens per-
761 sist on inanimate surfaces? A systematic review. *BMC infectious diseases*.
762 2006;6(1):1–8.
- 763 [58] Liu X, Huang J, Li C, Zhao Y, Wang D, Huang Z, et al. The role of
764 seasonality in the spread of COVID-19 pandemic. *Environmental research*.
765 2021;195:110874.
- 766 [59] Valtonen A, Molleman F, Chapman CA, Carey JR, Ayres MP, Roininen
767 H. Tropical phenology: Bi-annual rhythms and interannual variation in an
768 Afrotropical butterfly assemblage. *Ecosphere*. 2013;4(3):1–28.
- 769 [60] Getz WM, Salter R, Muellerklein O, Yoon HS, Tallam K. Modeling epi-
770 demics: A primer and Numerus Model Builder implementation. *Epidemics*.
771 2018;25:9–19.
- 772 [61] Lythgoe KA, Hall M, Ferretti L, de Cesare M, MacIntyre-Cockett G,
773 Trebes A, et al. SARS-CoV-2 within-host diversity and transmission.
774 *Science*. 2021;372(6539). Available from: [https://science.sciencemag.
775 org/content/372/6539/eabg0821](https://science.sciencemag.org/content/372/6539/eabg0821).
- 776 [62] Jones JH. Notes on R0. California: Department of Anthropological Sciences.
777 2007;323.

778 **Figures and Tables**

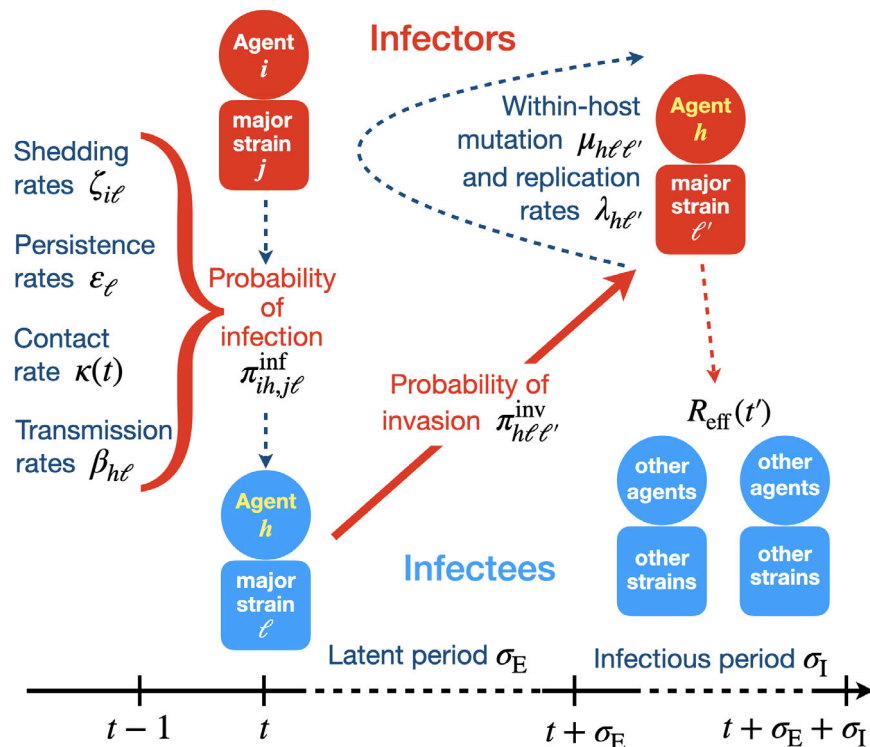


Figure 1: An overview of the processes included in our SEIVD-IBM model. The probability $\pi_{ih,j\ell}^{\text{inf}}$ of A_h being infected primarily with pathogen ℓ in terms of receiving an effective dose from agent A_i , (η) is computed in terms of a concatenation of shedding rates ($\zeta_{i\ell}$), environmental persistence rates (ε_ℓ), and host transmission ($\beta_{h\ell}$) processes (Eq. A.15) and includes both waning and cross immunity factors. The probability $\pi_{h\ell\ell'}^{\text{inv}}$ that the dominant strain emerging in host A_h is strain ℓ' given initial infection with strain ℓ is computed in terms of within-host mutation and within-host replication process (Eq. A.16) and also includes both waning and cross immunity factors. These two probabilities are then used to compute the overall probability $\pi_{ih,j\ell'}$ (Eq. A.17) that infector i , infected with major strain j , infects infectee h with major strain ℓ' . The quantity $R_{\text{eff}}(t')$ is the expected number of individuals each infectious agent is expected to infect around time $t' \in [t + \sigma_E, t + \sigma_E + \sigma_I]$, where $R_0 = R_{\text{eff}}(0)$ is estimated for our model using Eq. A.29. Equation references for the other mathematical functions provided in this figure are given in Table 2.

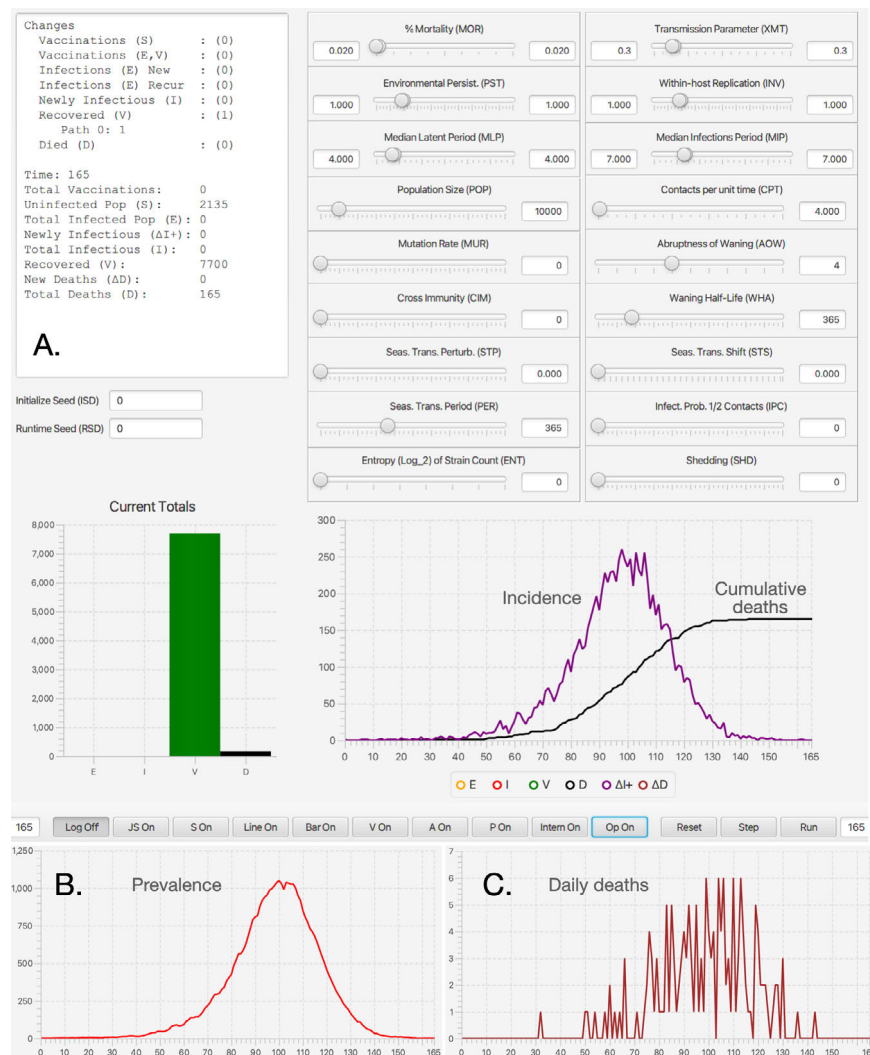


Figure 2: **A.** The dashboard of our Java Runtime-Alterable-Model Platform (J-RAMP) SEIRD (S=susceptibles, E=exposed, I=infectious, V=immune, D=dead) individual-based model (IBM) and simulations obtained using the parameters values depicted in the slider windows (also see Table 1). The top left window of this dashboard contains information on the final state of the population (in this case $S = 2135$ and $D = 165$ in a population of $N_0 = 10,000$), the bottom left bar graph of dashboard panel is the final values of E, I, V and D at epidemic cessation at time $t = 166$ (days) or the simulation run time, whichever comes first. Dashboard also shows a graph of incidence and cumulative deaths (purple and black: selected using colored buttons below the graph). The bottom ribbon of the dash board has a series of radio buttons that respectively open a Log, a JavaScript (JS), and a Scripting (S) window, Line and Bar graph windows (for multistrain runs), as well as windows for controlling vaccination strategies (V), listing realtime agent information (A), pathogen parameter values (P), monitoring probability computations (Intern), coding and controlling runtime alternative operations (Op), and three runtime buttons (Reset, Step, Run). **B.** Graphs of prevalence (cut out from main panel when only the red button is on) and **C.** daily deaths (crimson button) are pasted below the dashboard.

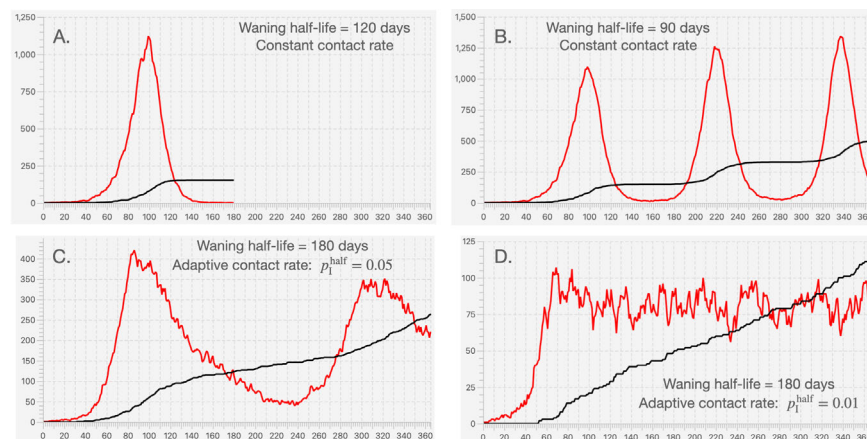


Figure 3: Plots of prevalence (I: red) and dead (D: black) individuals for simulations using the parameter values given in Table 1 with $N = 10,000$. The waning immunity half-life values are $t^{\text{half}} = 90, 120$ or 180 , as labeled. In the top two panels the contact rate is constant, while in the bottom two panels it is adaptive with $p_I^{\text{half}} = 0.05$ or 0.01 as labeled (Eq. A.10). Note that the adaptive case approaches the constant case as $p_I^{\text{half}} \rightarrow \infty$ and that the vertical axis in the four panels each have different scale, though the horizontal time axes is the same in all four case (365 days).

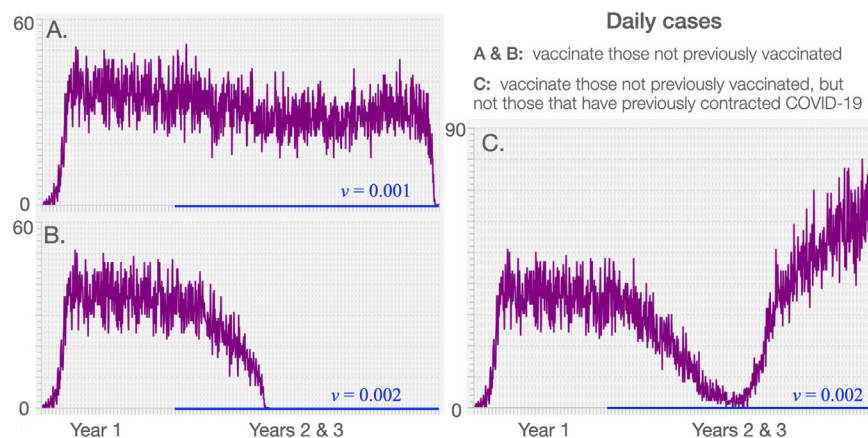


Figure 4: Incidence (ΔI^+ : purple) is plotted over 3 years for the baseline run (parameters given in Table 1 with $N = 100,000$, $p_I^{\text{half}} = 0.002$, and $t_v^{\text{half}} = 365$) for the cases where vaccination rates $v(t)$ (indicated by blue lines) are applied during the second and third years only to individuals not previously vaccinated selected at random (A. $v(t) = 0.001$, B. $v(t) = 0.002$) and to individuals not previously vaccinated and who also have not previously been infected with SARS-CoV-2 (C. $v(t) = 0.002$).

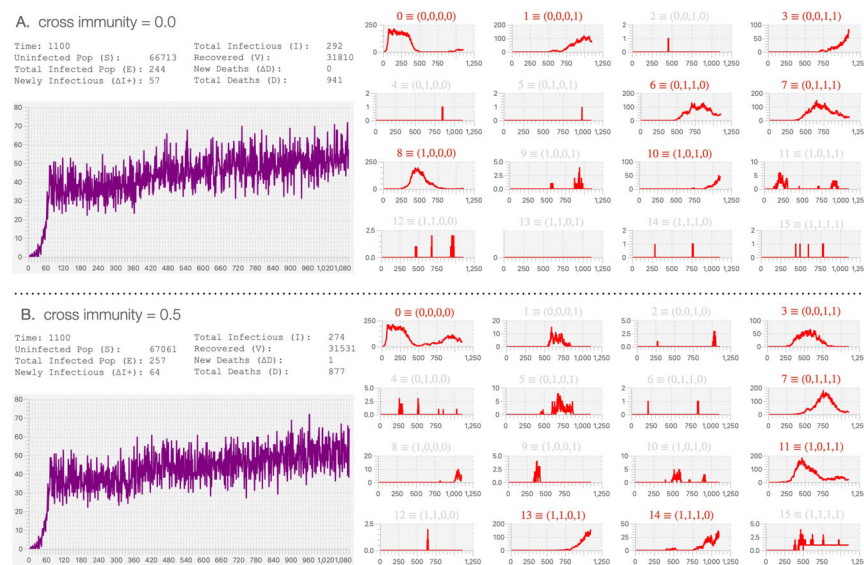


Figure 5: The total daily incidence ($\Delta I+$: purple) and strain-specific prevalence (I : red) individuals for a 16-strain epidemic burning through a population of size $N = 100,000$ are plotted over a 3-year period for the cases where cross-immunity is absent (A. $c = 0.0$) or intermediate among nearest neighbors differing by one allele (B. $c = 0.5$). The remaining parameters are listed in Table 1 and for all strains include the same host transmission ($\beta = 0.3$) and virulence ($p_\alpha = 0.02$), waning half-life $t^{\text{half}} = 365$, minor strain shedding $\zeta = 0.001$ and within host mutation ($\mu = 0.001$) parameter values. Note that each panel has its own vertical scale and that the horizontal scales are set by the J-RAMP and are 0-1100 days for the incidence curves and 0-1250 days for the prevalence curves, although all curves are drawn to 1100 days. Dominant or codominant strains are labeled in red (number and equivalent allele representation), while minor strains are labeled in light grey.

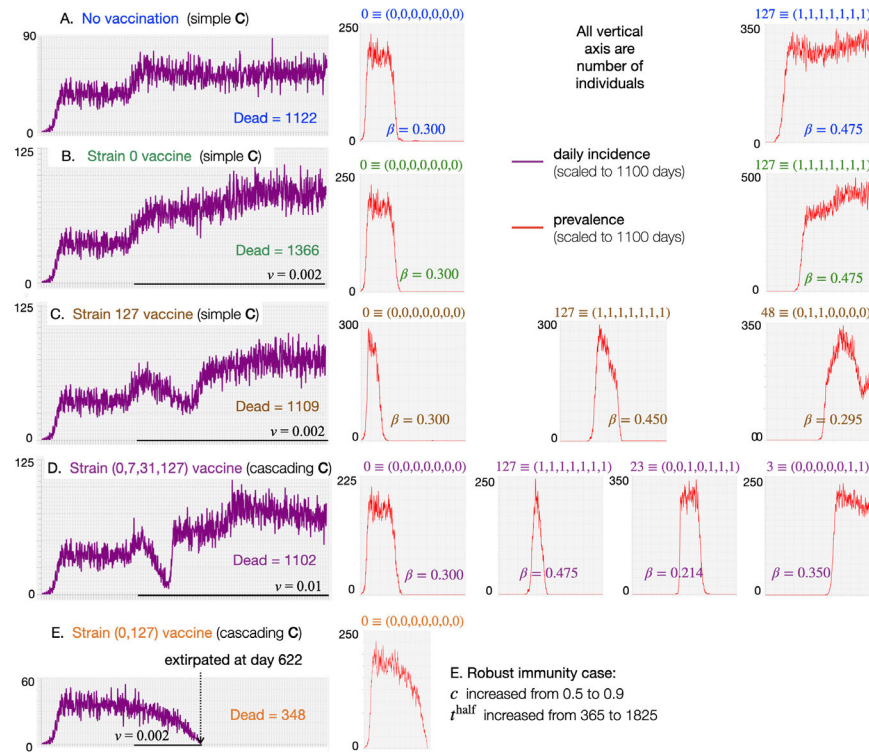


Figure 6: Daily incidence (deep purple curves) and prevalence of individual strains (red curves; with strain number labels plotted in order of emergence over time) that exceeded a maximum prevalence of 12 individuals at any time are plotted over a three year period ($t \in [0, 1100]$ days), firstly, for three cases (A.-C. in a system with nearest-neighbor only cross-immunity (labeled “simple”, see Eq. 1) and, secondly, in two cases (D. and F.) in a system with cascading cross-immunity (Eq. A.7). In the first three cases we have: A., no vaccination (blue labels); an B. & C., a vaccination rate of $v = 0.002$ applied in years 2 and 3 ($t \in [365, 110]$) to individuals who have not previously been vaccinated, respectively for the cases of univalent vaccines (strain 0, green labels and strain 127, brown labels). In the fourth case, D., a quadravalent vaccine (strain 0, 7, 31 & 127; plum labels) is applied at a vaccination rate of $v = 0.01$ (i.e., 1%) at random, irrespective whether or not individuals selected have been previously vaccinated (or ill with COVID-19). In the fifth case, E (“robust immunity case”), a bivalent vaccine (strains 0 & 127; orange labels) is applied, as in cases B. & C., at a vaccination rate of $v = 0.002$ (i.e., 1%) at random to individuals selected that have been previously vaccinated. In this latter case, however immune waning has been reduced by a factor of 5 ($t^{\text{half}} = 1825$ days) with cascading cross immunity increased from $c = 0.5$ to $c = 0.9$. The dashboards for all these cases, and miniature prevalence plots of all 128 strains for each of these four cases can be found in Figs. C.5-C.9 in the Appendix. The parameter values are the same as in Fig. 5B, apart from the immune parameters being greatly strengthened in the “robust immunity case” E., and in all cases strain specific transmission parameter values β_j ($j = 0, \dots, 127$) that were randomly assigned values between 0 and 0.3, except for the increasingly dominant strain sequence $\beta_0 = 0.3$ ($0 \equiv (0, 0, 0, 0, 0, 0, 0)$), $\beta_1 = 0.325$ ($1 \equiv (0, 0, 0, 0, 0, 0, 1)$), $\beta_3 = 0.35$ ($\equiv (0, 0, 0, 0, 0, 1, 1)$), ..., $\beta_{127} = 0.475$ ($127 \equiv (1, 1, 1, 1, 1, 1, 1)$) (particular values of β_j are provided for the different strains depicted here; see Table in Fig. C.4 for all values of β_j , $j = 0, \dots, 127$). Note that the vertical axis are all on the same scale height for incidence, but are individually scaled for prevalence, while the time axis are all on the same scale for the prevalence curves which differs from the incidence curves, which among themselves are all on the same time scale.

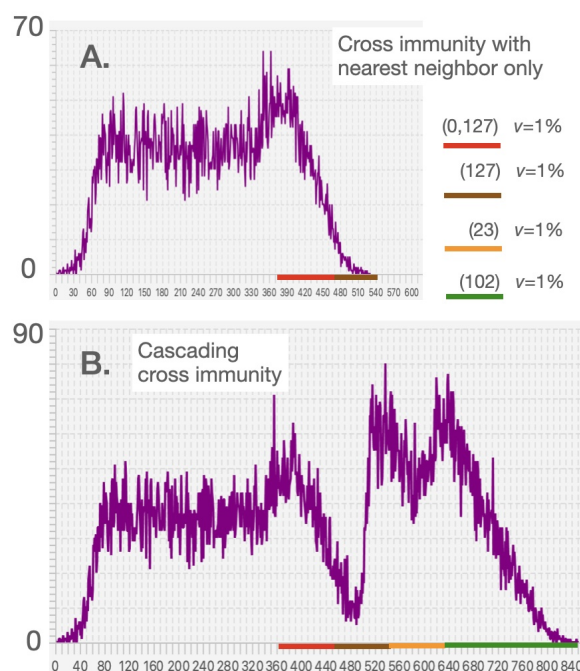


Figure 7: The daily incidences (vertical axis: number of individuals) are plotted over time (horizontal axis: number of days; plots have been squashed to ensure they are all on the same time scale) for two cases: **A.** adaptive vaccination for nearest-neighbor only cross immunity (Eq. 1) and **B.** adaptive vaccination for cascading cross immunity case (Eq. A.7), and should be compared with the non-vaccination case depicted in Fig. 6A and the non-adaptive vaccination case depicted in Fig. 6D. As in this latter non-adaptive case, vaccination is performed on non-infectious individuals, selected at random (i.e., re-vaccinations can occur) at a rate of 1% per day, but now the dominant strain is targeted, as well as a second strain if its prevalence exceeds 10 cases, with vaccine valency switches every 90 days. The realized adaptive vaccine valencies for these two runs are indicated by colored bars on the horizontal time axis (red, valency is (0,127); brown, valency is (127); orange, valency is (23); green, valency is (102)).

Table 1: Parameter values used to simulate single and multistrain outbreaks

Parameter	Symbol	Value	Source/Comment
Single strain simulations			
Time unit	t	daily	Goldilocks value*
Nominal pop size	N_0	1 to 10×10^4	Goldilocks value [‡]
Basic contact rate	κ_0	4 per day	effective contacts [†]
Transmission	β	0.3	ensures $R_0 \approx 3^+$
Latent Period	σ_E	4 days	median time in E [¶]
Infectious period	σ_I	7 days	median time in I [§]
Immunity half-life	t^{half}	1/4 to 1 year	per run specs.
Disease-induced mort.**	p_α	2% of cases	mortality rate is $\alpha^\#$
Adaptive contact param.*†	p_I^{half}	0, 0.002, 0.05	reduces κ from κ_0
Seasonal fluctuation param.	δ	0	seasons ignored* ⁺
Multi strain simulations (single strain parameter values for all strains if not mentioned below)			
Mutation rate* [‡]	μ	0.001* [#]	See Eq. A.16
Strain number	$j = 0, \dots, 2^J - 1$	J is 1 to 8	i.e., 2 to 1024 strains
Cross immunity	c_{jm}	0 & 0.5	See Eq. 1
Pathogen shedding	\bar{c}_{jm}	0.001* [#]	See Eq. A.12
Environmental persistence	$\bar{\eta}_j$	1 for all j	See Eq. A.13
Transmission	$\bar{\beta}_j$	[0,0.475]	per run specs.
Within-host replication rate	λ_j	1 for all j	See Eq. A.16
Disease-induced mort.**	p_{α_j}	[0,0.05]	per run specs.

*Weekly unit produced excessive variation (see SOI)

[‡]Trade-off between excessive demographic stochasticity and simulation run time

[†]Sufficiently long and close to make transmission possible (also see Eq. A.29)

⁺See Eq. A.29

[¶]Reciprocal of γ in continuous time computation of R_0 per Eq. A.29

[§]Reciprocal of ρ in continuous time computation of R_0 per Eq. A.29

^{||}See Eq. A.6: note $w(t)$ switches from 1 to 0 as immunity goes from complete to absent

[#]This is “virulence” parameter of continuous-time SEIR models

**If $\alpha \ll 1$ then $p_{\alpha_j} = 1 - e^{-\alpha} \approx \alpha$

^{††}See Eq. A.10. Mostly 0.002 is used. Setting $p_I^{\text{half}} = 0$ implements $\kappa(t) = \kappa_0$, though $\kappa(t) \rightarrow \kappa_0$ as $p_I^{\text{half}} \rightarrow \infty$

⁺⁺Implies values of k and θ in Eq. A.13 are irrelevant

^{‡‡}Strain independent—strain dependence requires more elaborate model

^{##}Quantifies the pass-on rate of mutations rather than the mutational rate of a locus or gene

^{***}If $\alpha_j \ll 1$ then $p_{\alpha_j} = 1 - e^{-\alpha_j} \approx \alpha_j$

Table 2: Variables, indices and functions used in the model

Symbols	Variables and indices	Equation
Variables		
$N_S(t), N_A(t), N_D(t)$	size of sets S , A and D	Eq. A.1
J, j, m, ℓ and ℓ'	strain entropy and indices $(0, \dots, 2^J - 1)$	Eq. A.2
$N_I(t), N_{I_j}(t)$	number of infectious: total and strain j at time t	
A_i, A_h	specific agents $i, h = 1, \dots, N_A(t)$ in set A	Eq. A.4
Functions		
$\omega_{ij}(t)$	waning immunity of A_i w.r.t. strain j	Eq. A.6
$\kappa(t)$	adaptive contact rate	Eq. A.10
$\phi_{ij}(t)$	immunity modifier	Eq. A.11
$\zeta_{ij}(t)$	shedding rate of strain j by infector A_i	Eq. A.12
$\eta_\ell(t)$	environmental persistence	Eq. A.13
$\beta_{h\ell}(t)$	strain transmission to infectee A_h	Eq. A.14
$\pi_{ih,j\ell}^{\text{inf}}(t)$	probability A_i infects A_h	Eq. A.15
$\pi_{h\ell\ell'}^{\text{inv}}(t)$	probability ℓ' is major strain when ℓ invades	Eq. A.16
$\pi_{ih,j\ell'}(t)$	probability ℓ' is major strain in A_h when j is major strain in A_i	Eq. A.17

1 APPENDICES

2 A Model Construction

3 Here we formulate an individual-based or agent-based (ABM) SEIR epidemio-
4 logical model to include host immunological waning and pathogen genetic drift
5 with variation across strain transmissibility and virulence. [1]

6 Assumptions, definitions, and states

7 The population consists of a well-mixed pool of N_0 individuals that is homoge-
8 neous except for the fact that some are uninfected (denoted S), some currently
9 infected (E: exposed and not yet infectious; I infectious and asymptomatic or
10 symptomatic) or have been infected and are now either dead (D) or recov-
11 ered/vaccinated with some level of immunity (V) to one or more of 2^J pathogen
12 strains. This immunity wanes over time and its current level, augmented by
13 specified levels of strain cross-immunity, factored into an agent specific time-
14 dependent strain-resistance function that impacts the shedding of mutant strains
15 by infectors and the within-host replication rates of mutant strains in infectees.

16 At the start of the epidemic, all individuals are assumed to encounter, on
17 average, $\kappa_0 > 0$ other individuals during each time period $[t, t + 1]$, but this
18 “effective contacts” rate adaptively decreases with increasing prevalence of the
19 disease due to the implementation of non-pharmaceutical interventions (social
20 distancing, hand washing, mask wearing, and other hygienic precautions). In
21 our selection of epidemiological parameter values, a unit of time is taken to be a
22 24-hour day. Other scalings of time would then require appropriately adjusted
23 epidemiological parameter values. Refined versions of the model could include
24 age-related parameter values and contact rates, as well as contact tracing, quar-
25 antining, and isolation of infected individuals; but these will not be considered
26 here.

27 Initially, at model time $t = 0$, all individuals are considered SARS-CoV-
28 2 naïve susceptible apart from one individual who is considered to have just
29 entered the infectious stage, infected by a pathogen designated as pathogen
30 strain 0 (wildtype). Throughout the model simulation, the N_0 agents in the
31 population are partitioned into three disjoint sets: the set of SARS-CoV-2 naïve
32 individuals, $\mathbf{S}(t)$, containing $N_S(t)$ (the susceptibles); the set of identified agents,
33 $\mathbf{A}(t)$, containing $N_A(t)$ individuals who are either currently infected (time t)
34 with a particular strain of SARS-CoV-2, or have some level of waning immunity
35 to one or more strains of SARS-CoV-2; and the set of dead individuals $\mathbf{D}(t)$,
36 currently of size $N_D(t)$. Only the individuals in $\mathbf{A}(t)$ are uniquely identified as
37 they become infected for the first time and make the transition from set $\mathbf{S}(t)$ to
38 set $\mathbf{A}(t)$, where they are sequentially labeled using the index $i = 1, \dots, N_A(t)$.
39 The single infected individual at time zero will be designated Agent 1 (also
40 known as patient zero and denoted by A_1). Thus at time t it follows that

$$N_S(t) + N_A(t) + N_D(t) = N_0 \quad (\text{a constant}) \quad (\text{A.1})$$

41 We note that individuals in set $\mathbf{A}(t)$ can be in a disease state E or I with
42 respect to pathogen j , but simultaneously can be in multiple immune states if

43 they have been infected with more than one pathogen strain in the past. We also
44 note that the distinction between symptomatic and asymptomatic individuals
45 in state I will not be considered here; and only need be incorporated if testing,
46 quarantining, and treatment processes are included in the model.

47 The total number of pathogen strains is set by a parameter $J > 0$, where each
48 pathogen is represented by a J -bit binary number. Thus, there are 2^J possible
49 strains indexed by $j = 0, 1, 2, \dots, 2^J - 1$ where j is the decimal equivalent that
50 corresponds to a given binary string. The initial strain, $j = 0$ is the binary
51 string of J zeros.

52 Sets of stochastic epidemic events (i.e., transitions from classes S to E, E to
53 I, I to V or D) are implemented at consecutive integer points in time (one set
54 of events for each point in time). Events will only be considered on individuals
55 that have been infected by at least one of the pathogens at some time after $t = 0$
56 (this means that initially the epidemic computation proceeds rather rapidly, but
57 becomes more computationally intensive for each time step as time proceeds).

58 Pathogen set

59 At the start of the simulation ($t = 0$), the set of potential pathogens indexed
60 by $j = 0, \dots, 2^J - 1$ is generated along with its associated environmental persis-
61 tence ($\bar{\eta}_j$), transmission ($\bar{\beta}_j$), within host replication (λ_j) and disease-induced
62 mortality rate (probability of dying from the disease p_{α_j}) parameters. These
63 may be specified or drawn from underlying distributions (e.g., the uniform dis-
64 tributions $\beta \sim \text{Uniform}[\beta_{\min}, \beta_{\max}]$ and so on). Also, our model includes two
65 $2^J \times 2^J$ matrices of constants that are associated with pathogen mutations dur-
66 ing strain shedding (elements ζ_{jm}) and cross-immunity (elements c_{jm}) processes
67 and thus involve but are conditioned on either the major strain that an infector
68 is harboring or on immunological state of the agents involved. These are the
69 shedding and cross-immunity matrices with elements $j, m = 1, \dots, 2^J - 1$. Thus
70 we generate the following list of parameters associated with our 2^J pathogen
71 strains:

$$\begin{aligned} \text{Pathogen list} &= & \text{(A.2)} \\ & \{(\bar{\eta}_j, \bar{\beta}_j, \lambda_j, p_{\alpha_j}; \zeta_{jm} \text{ and } c_{jm} \text{ for } m = 0, \dots, 2^J - 1) \mid j = 0, \dots, 2^J - 1\} \end{aligned}$$

72 Agent states

73 In accordance with the above set of assumptions, each agent has the following
74 basic disease states at time t , where disease states in agent A_i are referenced by
75 the time $\tau_j > 0$ at which the most recent infection with strain j has occurred
76 (an individual may be re-infected after immunity to the strain has waned to
77 relatively low levels):

- 78 1. $S(t)$: An individual who at time t has not been infected with any strain
79 of the pathogen up to time t . All these individuals belong to set $\mathbf{S}(t)$
- 80 2. $E_j(t, \tau_{ij})$: An agent A_i who was infected with strain j at time τ_{ij} , but has
81 not yet become infectious (this is an individual in the latent stage that
82 lasts for σ_E units of time). All these individuals belong to set $\mathbf{A}(t)$

- 83 3. $I_j(t, \tau_{ij})$: An agent A_i who is currently infectious with strain j , after being
 84 infected with strain j at time τ_{ij} (this is the infectious stage that lasts for
 85 σ_I units of time). All these individuals belong to set $\mathbf{I}(t) \subseteq \mathbf{A}(t)$
- 86 4. $V_j(t, \tau_{ij})$: An agent A_i who was infectious with strain j , having been
 87 infected with strain j at time τ_{ij} , but is now non-infectious with regard to
 88 this strain—that is, recovered with some immunity to strain j , as well as
 89 some cross immunity to strains closely related to j . All these individuals
 90 belong to set $\mathbf{A}(t)$
- 91 5. $\mathbf{D}(t)$: An individual at time t who has died after being exposed to and
 92 become infectious with some strain of the pathogen. In a refined version
 93 of the model, a record will be kept of the time of death and the strain that
 94 caused death. All these individuals belong to set $\mathbf{D}(t)$.

95 Since an agent A_i may be infected over time by more than one strain j , its
 96 complete epidemiological state is represented by a list

$$A_i(t) = \{\text{state w.r.t. pathogen } 0, \dots, \text{state w.r.t. pathogen } 2^J - 1\} \quad (\text{A.3})$$

If a living agent does not fall into any of the categories 2 – 4 with respect to pathogen j , we denote its epidemiological state at position j as \emptyset (the empty set). Consequently, if an agent A is susceptible at time t (i.e., an element of $S(t)$), then we write

$$A_0(t) = \{\emptyset, \dots, \emptyset\} \in \mathbf{S}$$

97 However, while such individuals are omitted from the A list (hence we did not
 98 subscript the agent A above), they may be recognized as “virtual members”
 99 with this implicit state. Some other examples are:

- If $A_i(t)$ is infected, but not yet infectious, with pathogen strain j at time t but has not been infected with any other pathogen in its past history, then

$$A_i(t) = \{\emptyset, \dots, \emptyset, E_j(t, \tau_j), \emptyset, \dots, \emptyset\}$$

- On the other hand if A_i recovered from an infection with pathogen 0 at time τ_0 , and is now in infectious with pathogen j at time t , having become infected with this pathogen at time τ_j then we write

$$A_i(t) = \{V_0(t, \tau_0), \emptyset, \dots, \emptyset, I_j(t, \tau_j), \emptyset, \dots, \emptyset\}$$

100 As we shall see, an agent history may contain at most one instance of either E_j
 101 or I_j , while possibly containing multiple instances of V_j .

102 Agent and index sets

103 At the start of each time period, we update the set of identified agents \mathbf{A} by
 104 adding susceptibles that became infected with pathogens during the previous
 105 time period and removing agents that died during the previous time period.
 106 Thus if \mathbb{I}_A is the index set for non-empty elements of \mathbf{A} , with new indices added

107 for newly infected susceptibles and indices removed for individuals that died,
 108 then by definition:

$$\mathbf{A}(t+1) = \{A_i(t+1) | i \in \mathbb{I}_A(t+1)\} \quad (\text{A.4})$$

109 where the number of indices in the updated set $\mathbb{I}_A(t+1)$ is $N_A(t+1)$ and the
 110 updated number of dead is $N_D(t+1)$ at time $t+1$.

111 For mathematical convenience all susceptibles S will also be referred to as
 112 A_0 ; i.e., there are $N_S(t)$ individuals referenced by A_0 at time t . It will be useful
 113 to partition the set $\mathbf{A}(t)$ itself into three subsets at time t by identifying the sets
 114 $\mathbf{E}(t)$ and $\mathbf{I}(t)$ which respectively contain all agents that are currently in a state
 115 $E_j(t)$ or a state $I_j(t)$ at time t for some $j = 0, \dots, 2^J - 1$. We note the intersection
 116 of these two sets is empty—i.e., $\mathbf{E}(t) \cap \mathbf{I}(t) = \emptyset$ —as will become apparent below
 117 from the transmission process rules set up below. We will use the notation

$$\mathbf{A}_S(t) = \mathbf{A}(t) \setminus (\mathbf{E}(t) \cup \mathbf{I}(t)) \quad (\text{A.5})$$

118 to denote the set of agents in $\mathbf{A}(t)$ but not in $\mathbf{E}(t)$ or $\mathbf{I}(t)$.

We also identify the set of infectious agents with infectious strain j . If A^j
 denotes an agent whose epidemiological state contains an entry $I_j(t, -)$, then

$$\mathbf{I}_j(t) = \{A_{i_1}^j(t), A_{i_2}^j(t), \dots, A_{i_{N_{I_j}(t)}}^j(t)\},$$

where the number of such agents is denoted by $N_{I_j}(t)$, and its index set by

$$\mathbb{I}_{I_j}(t) = \{i_{1_j}, \dots, i_{N_{I_j}(t)}\}.$$

119 Epidemiological processes

120 Immunity

121 In compartmental SIRS and SEIRS models, a concept of waning immunity and
 122 its impact on epidemics is associated with the rates at which individuals in class
 123 R revert back to class S. In agent-based SIRS and SEIRS models, we have the
 124 opportunity to consider the immunological history of individuals and, hence,
 125 can take a more refined approach to the complex process of how pathogens in
 126 an infector A_i are passed on to an infectee A_h . Here we model this as a prob-
 127 ability generated from a concatenation of rates that include pathogen shedding
 128 by A_i , the survival of pathogens in the environment, whether contained in feces,
 129 urine, sweat, mucosal secretions or water droplets excreted by an infector, and
 130 a process whereby pathogens gain access to a host (entering through wounds,
 131 mucosal membranes or other membranes in the pulmonary or alimentary sys-
 132 tems). We then characterize pathogen within-host strain replication rates in
 133 terms of pathogen mutational and reproductive processes. The final outcome
 134 in our model is either host recovery with some immunity or host death. We
 135 also consider the induction of host immunity through vaccination and make the
 136 assumption that waning immunity is the same, whether it stems from natural
 137 infection or vaccination. Of course, these may be modelled in different ways
 138 should data become available to make this distinction an important modeling

139 consideration.

140

141 Waning immunity. Recall that we use A_0 to denote an anonymous (generic)
 142 member of \mathbf{S} and that A_i for $i > 0$ refers to a specific individual with an
 143 associated state list/vector. If some specific A_i is in immune state V_j having
 144 been infected with this strain at time τ_{ij} , we assume that the level of relative
 145 susceptibility of agent A_i to reinfection by strain j is given by (noting that the
 146 existence of the value τ_{ij} implies that infection of individual i by strain j at
 147 time τ_{ij} ensures that the V_j is no longer “null”)

$$\omega_{ij}(t, \tau_{ij}) = \begin{cases} 0 & \text{if } V_j \text{ is null} \\ \frac{1}{1 + ((t - \tau_{ij} - \sigma_I - \sigma_E)/t_j^{\text{half}})^\sigma} & \text{if } t \geq \tau_{ij} + \sigma_I + \sigma_E \\ 1 & \text{if } t < \tau_{ij} + \sigma_I + \sigma_E \end{cases} \quad (\text{A.6})$$

148 We note the following: 1.) the first case implies that τ_{ij} has yet to be defined; 2.)
 149 the second case is equivalent to the statement that $\tau_{ij} \geq 0$ now exists for strain
 150 j , since this occurs at time $t = \tau_{ij}$ (through the invocation of state $E_j(t, \tau_{ij})$);
 151 3.) $\omega_{ij}(t, \tau)$ ranges from 1 (i.e. full “on”) at $t = \tau + \sigma_I + \sigma_E$ and decays to
 152 0 as $t > \tau + \sigma_I + \sigma_E \rightarrow \infty$; 4.) agent i cannot be reinfected with its current
 153 major strain or with any other strain while it is currently itself in any state E_j
 154 or I_j for any $j = 0, \dots, 2^J - 1$; 5.) the larger the value of σ the steeper or more
 155 abrupt the switch is from full immunity (equal to 1) at time τ through 1/2 at
 156 time t_j^{half} to approach 0 as $t \rightarrow \infty$ (we set $\sigma = 4$ as providing an intermediate
 157 level of abruptness).

158

159 Cross immunity with escape mutations. A somewhat more general implemen-
 160 tation of cross-immunity may be based on the number of alleles by which strains
 161 j and ℓ differ. If they differ in k positions, then the level of cross immunity can
 162 be set to c^k . In this case

$$\text{Cascading C:} \quad c_{j\ell} = \begin{cases} 1 & \text{if } \ell = j \\ c^k & \text{if } \ell \text{ differs from } j \text{ by } k \text{ alleles and } \ell \end{cases} \quad (\text{A.7})$$

163 In addition to this, one may also designate certain changes in certain alleles as
 164 “escape mutants” with respect to the progenitor strain. For these, the level of
 165 cross-immunity would be set to zero. Under these assumptions, a generalization
 166 of Eq. A.7 would be

$$\text{Escape mutation:} \quad c_{j\ell} = \begin{cases} 1 & \text{if } \ell = j \\ c^k & \text{if } \ell \text{ differs from } j \text{ by } k \text{ alleles} \\ & \text{and } \ell \text{ is not an escape mutation} \\ 0 & \text{if } \ell \text{ is an escape mutation} \end{cases} \quad (\text{A.8})$$

167 In particular, for the case $J = 7$, if strain $(1, x, x, x, x, x, x)$ is an escape
 168 mutation with respect to some level of immunity to strains $(0, x, x, x, x, x, x)$,
 169 then Eq. A.8 becomes

$$c_{j\ell} = \begin{cases} 1 & \text{if } \ell = j \\ 0 & \text{if } j \leq 63 \text{ and } \ell > 63 \\ c^k & \text{otherwise, where } \ell \text{ differs from } j \text{ by } k \text{ alleles} \end{cases} \quad (\text{A.9})$$

170 **Vaccination.** A vaccine may be designed to give immunity to one or more
 171 particular identified strain j . Vaccination strategies include vaccinating at a
 172 fixed rate $v(t)$ (percent of individuals vaccinated at each time period) over a
 173 fixed period that begins at t_v^{on} and ends at t_v^{off} and can focus on drawing only
 174 on: i) individuals in the set \mathbf{S} , ii) any non-infectious individual in \mathbf{S} or \mathbf{A} , or iii)
 175 any non-infectious, not previously vaccinated individual in \mathbf{S} or \mathbf{A} . The vaccine
 176 itself can be designed as follows:

- 177 • *Dominant strain vaccination at time τ_{vac} .* An individual \mathbf{S} or \mathbf{A}_i vacci-
 178 nated with the dominant strain, say j , at time $\tau_{\text{vac}} \in [t_v^{\text{on}}, t_v^{\text{off}}]$ serves to
 179 add the disease state $V_j(t, \tau_{\text{vac}})$ to that individual’s list. If the individual
 180 is already in state $V_j(t, \tau')$ at time $t > \tau_{\text{vac}}$, then its status is updated so
 181 that at time $t > \tau_{\text{vac}}$ it is now $V_j(t, \tau_{\text{vac}})$ rather than $V_j(t, \tau')$
- 182 • *Multi-strain vaccination at time t_{vac} .* An individual \mathbf{S} or \mathbf{A} vaccinated with
 183 a multi strain concoction at time $\tau \in [t_v^{\text{on}}, t_v^{\text{off}}]$, say with strain j_1, \dots, j_ν ,
 184 will have their disease status updated with regard to all these strains, as
 185 in the dominant strain case.

186 Infectious contacts

187 Infectious individuals are assumed to make $\hat{\kappa}$ *effective contacts* each time period;
 188 where effective contacts are those that are sufficiently close and of a sufficiently
 189 long duration to constitute a “risk of transmission.” This rate is either a constant
 190 κ_0 , or in stochastic implementations drawn from a Poisson distribution
 191 with mean κ_0 , or in adaptive formulations (e.g., under social distancing be-
 192 haviour) is a function of the severity of the ongoing outbreak. We also assume at
 193 time t that under a random contact process, proportion $\frac{\hat{\kappa}(t)N_S(t)}{N_0 - N_D(t)}$ and $\frac{\hat{\kappa}(t)N_A(t)}{N_0 - N_D(t)}$
 194 of these contacts will respectively be with susceptible and with uniquely identi-
 195 fied agents, although only $\frac{\hat{\kappa}(t)(N_A(t) - N_E(t) - N_I(t))}{N_0 - N_D(t)}$ of those will be susceptible to
 196 infection with a new strain or reinfection with the same strain.

197 In the adaptive case, we assume $\kappa(t)$ decreases from κ_0 as the proportion
 198 of infectious individuals, $N_I(t)/(N_0 - N_D(t))$, increases such that $\kappa(t) = \kappa_0/2$
 199 when $N_I(t)/(N_0 - N_D(t)) = p_I^{\text{half}}$. For convenience of implementation, however,
 200 we define the following “switching” (as apposed to hyperbolic) function

$$\kappa(t) = \begin{cases} \frac{\kappa_0}{1 + \left(\frac{N_I(t)}{N_0 - N_D(t)} / p_I^{\text{half}}\right)^2} & \text{when } p_I^{\text{half}} > 0 \\ \kappa_0 & \text{when } p_I^{\text{half}} = 0 \end{cases} \quad (\text{A.10})$$

201 even though, from a continuity point of view, the top part of this expression
 202 implies that $\kappa(t) \rightarrow 0$ at $p_I^{\text{half}} \rightarrow \infty$.

203 Probability of infection

204 In deriving a probability $\pi_{ih,j\ell}^{\text{inf}}$ of an agent \mathbf{A}_h being infected with strain ℓ
 205 by and agent \mathbf{A}_i who is infectious with major strain j , we concatenate (i.e.,
 206 multiply together) several process, each of which involves nominal constants.
 207 Thus, in all but one of these processes, the scaling of these constants can be
 208 normalized and given a relative set of values across strains though one set of

209 constants though relative, will ultimately all be scaled by fitting the model to
 210 real data. In our treatment below, constants associated with shedding and per-
 211 sistence will be scaled while those associated with within-host replication will be
 212 kept unscaled to be ultimately fitted to data. In particular, the parameters $\bar{\beta}_j$
 213 associated with pathogen transmission (i.e., from contact to the start of within
 214 host replication—see Fig. 1) will be scaled by fitting to epidemiological data,
 215 while the relative values for the different strains regarding pathogen shedding
 216 and environmental persistence can be fitted to experimental data collect to set
 217 values of these processes when considered on their own.

218
 219 Pathogen shedding. We assume that shedding is affected by the immune state
 220 of the infector A_i and thus posit the shedding rates below for this individual
 221 when its major infectious strain is j . In general, we have a matrix of shedding
 222 rates $\bar{\zeta}_{j\ell}$ before accounting for immunity and cross immunity that is specific to
 223 agent A_i . Immunity and cross-immunity act to reduce shedding rates through
 224 functions $\phi_{ij}(t) \in [0, 1]$ that are computed in terms of A_i 's waning functions
 225 ω_{im} with respect to strain m and a matrix of cross-immunity values c_{jm} that
 226 have been normalized so that $c_{jj} = 1$ for $j = 0, \dots, 2^J - 1$ and $c_{jm} \in [0, 1]$ for
 227 $j, m = 0, \dots, 2^J - 1$. Specifically, we define agent-specific *immunity modifying*
 228 *functions*

$$\phi_{ij}(t) = \prod_{m=0}^{2^J-1} (1 - c_{mj}\omega_{im}(t)) \quad (\text{A.11})$$

229 and assume that the shedding rates can be expressed as

$$\zeta_{ij\ell}(t) = \bar{\zeta}_{j\ell}\phi_{i\ell} \quad \ell = 0, \dots, 2^J - 1 \quad (\text{A.12})$$

230 Environmental persistence. The persistence of pathogens in the environment
 231 are known to be impacted by humidity, temperature, airflow, and the surface
 232 properties of fomites [57]. This, and other factors relating the effects of weather
 233 on contact rates and efficacy, may result in overall pathogen transmission having
 234 a seasonal component to it [58]. In particular, viral persistence indoors may
 235 be much greater than outdoors, with a greater proportion of indoor contacts
 236 taking place during cold or wet weather. Thus the most appropriate place
 237 to introduce seasonal effects into epidemic processes is through contact rates
 238 and environmental persistence cycling over time with a period of one year (or
 239 even half-a-year if two comparatively spaced rainy seasons occur, as in some
 240 in tropical locations [59]) and an amplitude obtained by fitting parameters to
 241 data. Thus, in our model, we introduce constants $\bar{\eta}_\ell$, $\delta \in (0, 1)$, k (appropriately
 242 scaled, depending on the units of time) and θ and assume that

$$\eta_\ell(t) = \bar{\eta}_\ell \left(1 + \delta \sin \left(\frac{2\pi t}{k} + \theta \right) \right) \quad \ell = 0, \dots, 2^J - 1 \quad (\text{A.13})$$

243 The case $\delta = 0$ corresponds to constant values $\eta_\ell(t) = \bar{\eta}_\ell$ for all t , while if
 244 $\delta = 1$ we get the largest possible fluctuation between 0 and $2\bar{\eta}_\ell$. The constant k
 245 relates to the time units so we get one cycle per year, and θ shifts the cycle to
 246 set the points in time at which the maximum and minimum values of $\eta_\ell(t)$ occur.

247

248 Strain transmission. In the context of a standardized dose (which will be
 249 modified by multiplying the strain *effective contact* and transmission by both
 250 pathogen shedding and environmental persistence functions), the differential
 251 rates of strain transmission, which we denote by $\beta_{h\ell}$, will depend on a constant
 252 strain transmission rate parameter $\bar{\beta}_\ell$ modified by a function that represents
 253 the immune state of the infectee at time t : viz., recalling Eq. A.11

$$\beta_{h\ell}(t) = \bar{\beta}_\ell \phi_{h\ell}(t), \quad \ell = 0, \dots, 2^J - 1 \quad (\text{A.14})$$

254 Probability of infection. Using a competing rates formulation [60] to compute
 255 the probability of infection as a concatenation of the process of infector shedding
 256 (ζ), environmental persistence (η) and transmission rates (β), we obtain

$$\pi_{ih,j\ell}^{\text{inf}}(t) = \frac{\zeta_{ij\ell}(t)\eta_\ell(t)\beta_{h\ell}(t)}{\sum_{m=0}^{2^J-1} \zeta_{ijm}(t)\eta_m(t)\beta_{hm}(t)} \left(1 - e^{-\zeta_{ij\ell}(t)\eta_\ell(t)\beta_{h\ell}(t)}\right), \quad \ell = 0, \dots, 2^J - 1 \quad (\text{A.15})$$

257 Within-host processes

258 If after receiving an initial infectious dose of pathogen, an individual is infected
 259 primarily with strain ℓ , then we expect this strain to dominate unless intrinsic
 260 mutational processes are high (which is not the case for COVID-19) or the
 261 individual has some immunity to this dominant strain. In the latter case the
 262 situation is ripe for an “escape mutation,” that is one that evades the immune
 263 system, to arise.

264 If we nominally set the relative rate at which an individual invaded by strain
 265 ℓ has an infection dominated by strain ℓ (i.e., ℓ in the terminology of [61] is
 266 the major strain of the infection) to be $(1 - \mu)$, then the probability that one
 267 of the other strains $\ell' \neq \ell$ (in the case of COVID we assume that $\mu > 0$ is
 268 very close to 0—e.g. of order 10^{-3} to 10^{-6} —while for viruses lacking error
 269 correcting machinery it can be considerably larger and of the order 10^{-1}). We
 270 can partition the latter probability according to a set of comparative strain
 271 within-host replication rates $\lambda_{\ell'}$, each moderated by its immune state function
 272 $\phi_{h\ell'}$ and a normalizing factor $\frac{1}{\sum_{m \neq \ell} \lambda_m \phi_{hm}}$ to obtain

$$\pi_{h\ell\ell'}^{\text{inv}}(t) = \begin{cases} 1 - \mu & \text{for } \ell' = \ell \\ \mu \left(\frac{\lambda_{\ell'} \phi_{h\ell'}}{\sum_{m \neq \ell} \lambda_m \phi_{hm}} \right) & \text{for } \ell' \neq \ell \end{cases} \quad (\text{A.16})$$

273 Pathogen progression equations

274 Probability that infector A_i with major strain j will result in infectee A_h express
 275 ℓ' as its major strain is

$$\pi_{ih,j\ell'}(t) = \sum_{\ell=0}^{2^J-1} \pi_{ih,j\ell}^{\text{inf}}(t) \pi_{h\ell\ell'}^{\text{inv}}(t), \quad \ell = 0, \dots, 2^J - 1 \quad (\text{A.17})$$

276 **Single strain case**

277 In the single strain case ($J = 0$), the waning immunity equation Eq. A.6 reduces
 278 to (dropping the redundant index $j = 0$, and noting that the existence of a value
 279 τ_i implies A_i has been infected at time τ_i in the past)

$$\omega_i(t) = \begin{cases} 0 & \text{if } A_i \text{ has never been infected} \\ \frac{1}{1 + ((t - \tau_i - \sigma_I - \sigma_E)/t^{\text{half}})^\sigma} & \text{if } t \geq \tau_i + \sigma_I + \sigma_E \\ 1 & \text{if } t < \tau_i + \sigma_I + \sigma_E \end{cases} \quad (\text{A.18})$$

280 and the modifying immunity functions ϕ_{ij} (Eq. A.11) collapse to 1, which implies
 281 that the pathogen shedding functions ζ_{il} (Eq. A.12) collapse to 1. Without loss
 282 of generality, we can also assume that single strain value $\eta = 1$ in Eq. A.13,
 283 which implies that the probability of infection (Eq. A.15) reduces to

$$\pi_{ih}^{\text{inf}}(t) = 1 - e^{-\bar{\beta}(1-\omega_h(t))} \quad (\text{A.19})$$

284 Further, since in the single strain case there are no mutations to consider, it
 285 follows from Eq. A.16 that $\pi_{h\ell'}^{\text{inv}}(t) = 1$ for all h and we finally have that $\pi_{ih}(t) =$
 286 $\pi_{ih}^{\text{inf}}(t) = 1 - e^{-\bar{\beta}(1-\omega_h(t))}$ (Eq. A.17) for all h .

287 **Simulation algorithm**

288 **1. Parameters selected at the start of a simulation**

- 289 (a) N_0 : Number of individuals in the population. Assumed to be fixed
 290 over time (i.e., the population is closed), but partitioned into sets **S**,
 291 **A** and **D** with respectively $N_S(t)$, $N_A(t)$ and $N_D(t)$ individuals in
 292 each set and satisfying Eq. A.1.
- 293 (b) J : The \log_2 of the number of possible strains indexed by $j = 0, \dots, 2^J -$
 294 1
- 295 (c) $\bar{\beta}_j$: Strain dependent transmission parameters (the process between
 296 contact and the start of strain replication and nominally equivalent
 297 to transmission in SEIR models—see Fig. 1) for pathogen strain j
- 298 (d) t_j^{half} : The time it takes for immunity to strain j to have waned by
 299 half.
- 300 (e) σ_{E_j} : The time it takes from initial infection for an infected individual
 301 to become more likely to become infectious than remain infected
 302 without being infectious.
- 303 (f) σ_{I_j} : The additional time it takes beyond σ_{E_j} for an infectious indi-
 304 vidual to more likely transition beyond being infectious than remain
 305 infectious.
- 306 (g) p_{α_j} : The proportion of individuals leaving the infectious category
 307 that die, which implies that $1 - p_{\alpha_j}$ is the proportion that become
 308 immune.

309 **2. Initialization**

- 310 (a) Set up pathogen list (see Eq. A.2)
 311 (b) Initialize the simulation by setting $t = 0$ and creating the agent list
 312 $\mathbf{A}(0)$ one infectious and $N_0 - 1$ susceptible agents.

313 **3. Time t : vaccination loop.**

- 314 (a) Carry out the vaccination process before going into the rest of the
 315 loops with the updated S and A sets after the vaccinations.

316 **4. Time t : contact loop.** Set up contacts for the current round of en-
 317 counters at time t (i.e., the inner agent-driven contact loop within the
 318 outer time-driven loop) and tag for outer loop update of disease status, as
 319 follows:

320 (a) *Numbers in various sets and associated index sets.* Identify the num-
 321 ber of individuals $N_S(t)$, $N_A(t)$ and $N_D(t)$ in sets \mathbf{S} , \mathbf{A} and \mathbf{D} at
 322 time t respectively, as well as the number of exposed (but not yet
 323 infectious) agents $N_E(t)$, infectious agents $N_I(t)$ and identified non-
 324 infected agents $N_{AS}(t) = N_A(t) - N_I(t) - N_E(t)$. Break down the
 325 infectious agents tally into the number of agents N_{I_j} infectious with
 326 strain $j = 0, 1, \dots, 2^J - 1$. We will also need the index sets \mathbb{I}_{AS} and
 327 $\mathbb{I}_{I_j}(t)$, $j = 0, \dots, 2^J - 1$ at time t .

(b) *Infectious contacts with each group.* The rate at which any individual
 contacts other individuals per unit time is given by the contact rate
 parameter $\kappa > 0$. Assuming random contact events over one unit
 of time, the actual number of individuals that agent A_i contacts at
 time t is then given by

$$\hat{\kappa}_i(t) \sim \text{POISSON}[\text{mean} = \kappa(t)]$$

328 Of these, proportions

$$\pi_{iS} = \frac{N_S(t)}{N_0 - N_D(t)} \quad (\text{A.20})$$

329 and

$$\pi_{iA} = \frac{N_A(t) - N_I(t) - N_E(t)}{N_0 - N_D(t)} \quad (\text{A.21})$$

330 are expected to come from susceptibles in the sets $\mathbf{S}(t)$ and $\mathbf{A}_S(t)$
 331 (see Eq. A.5) respectively. Thus the actual number of contacts in set
 332 $\mathbf{S}(t)$, $\mathbf{A}_S(t)$, and $\mathbf{E}(t) \cup \mathbf{I}(t)$ are

$$\left(\hat{N}_i^S(t), \hat{N}_i^{AS}(t), \hat{N}_i^E(t) + \hat{N}_i^I(t) \right) = \text{Multinomial}[\hat{\kappa}_i; \pi_{iS}, \pi_{iA}, 1 - \pi_{iS} - \pi_{iA}] \quad (\text{A.22})$$

333 We note that only $\hat{N}_i^S(t)$ and $\hat{N}_i^{AS}(t)$ are of interest because indi-
 334 vidual in states E and I cannot be reinfected. Also, we make the
 335 assumption below that the first infection that an individual in set \mathbf{A}
 336 contracts in this contact loop, is the one that counts (i.e., there will

337 be no simultaneously infections with multiple strains). Finally, since
 338 contacting individuals is tantamount to sampling with replacement,
 339 the number of unique contacts (i.e., all multiple contacts are counted
 340 as a single contact) that agent A_i has with individuals in the set \mathbf{S} is
 341 $\hat{N}_i^S(t)$ reduced by excluding multiple contacts (which under a random
 342 contact model is a negative exponential correction) to obtain

$$N_i^{S*}(t) = \min \left\{ \hat{N}_i^S(t), \text{Binomial} \left[\hat{\kappa}_i(t), e^{-\hat{\kappa}_i(t)/\hat{N}_i^S(t)} \right] \right\} \quad (\text{A.23})$$

343 Thus if $\hat{\kappa}_i(t) \ll \hat{N}_i^S(t)$, $N_i^{S*}(t)$ is expected to be very close to the
 344 upper value $\hat{\kappa}_i(t)$. On the other hand, if $\hat{\kappa}_i(t) \approx \hat{N}_i^S(t)$, then $N_i^{S*}(t)$ is
 345 expected to be around $\hat{\kappa}_i(t)/e \approx 0.37\hat{\kappa}_i(t)$. Additionally, after dealing
 346 with each agent i reduce in the size of $N_S(t)$ to take account of those
 347 agents that had been infected by agent A_i and had now entered the
 348 ranks of the set \mathbf{A} .

- (c) *Identify all infectious agents and their pathogens strains.* Among all agents in the set $\mathbf{A}(t)$ (Eq. A.4), identify those that have an infectious strain I_j for some $j = 0, \dots, 2^J - 1$. Thus, if the number of infectious agents with infectious strain j is $N_{I_j}(t)$ then consider the set

$$\mathbf{I}_j(t) = \{A_{i_1}^j(t), A_{i_2}^j(t), \dots, A_{N_{I_j}(t)}^j(t)\}$$

with index set

$$\mathbb{I}_j(t) = \{i_{1_j}, \dots, i_{N_{I_j}(t)}\}$$

349 Initially, most of these sets will be empty, but will fill in over time.

- 350 (d) *Susceptible contacts.* The probability that an agent A_i with a strain
 351 j major infection infects a susceptible (nominally denoted by individ-
 352 uals of type A_0) who then becomes infectious with dominant strain ℓ'
 353 is given by the probability $\pi_{i0,j\ell'}$ computed in Eq. A.17, which itself
 354 relies on expressions Eq. A.10-A.16. The actual number of individ-
 355 uals in the set \mathbf{S} will make effective contact with one more infectious
 356 individuals is $N_i^{S*}(t)$ obtained using Eq. A.23. Thus, from a multi-
 357 nomial drawing, we can now generate the number of newly exposed
 358 individuals, $N_{0\ell'}^{E+}(t+1)$ (the “+” is used to denote these are newly
 359 added and the “0” that they are coming from the set \mathbf{S}), with major
 360 strain ℓ' at time $t+1$, have been infected by agent A_i with major
 361 pathogen strain j on the time interval $[t, t+1)$:

$$\begin{aligned} (\hat{N}_{00}^{E+}(t), \dots, \hat{N}_{02^J-1}^{E+}(t)) &\sim \\ &\text{Multinomial} [N_i^{S*}(t); \pi_{i0,j0}(t), \dots, \pi_{i0,j2^J-1}(t)] \end{aligned} \quad (\text{A.24})$$

362 These individuals will be used to update list of currently infected
 363 individuals in the sets \mathbf{A}_{E_j} , $j = 0, \dots, 2^J - 1$ at time $t+1$, which is
 364 computed in the outer loop computation, as presented below. We also
 365 note that the probabilities in the above multinomial add to less than
 366 1, so that at the end of the drawing a proportion of the individuals
 367 $N_i^{S*}(t)$ remain uninfected.

368 (e) *Agent contacts.* The number of agents $\hat{N}_i^{\text{As}}(t) \in \mathbb{I}_{\mathcal{A} \setminus (\text{E} \cup \text{I})}$ that come
 369 into contact with agent A_i over the interval $(t, t + 1)$ is given by
 370 Eq. A.22. This number is drawn from the set $\mathbb{I}_{\mathcal{A} \setminus (\text{E} \cup \text{I})}$ with replace-
 371 ment and the following multinomial computation is used to determine
 372 how to update agent A_h at time $t + 1$ when coming into contact with
 373 agent A_i on the interval $(t, t + 1)$ using the probabilities of transmis-
 374 sion given in Eq. A.17. Specifically, agent A_h will become infected
 375 with major strain ℓ' at time $t + 1$ is determined by the multinomial
 376 drawing

$$A_h \in \mathbf{A}_{E_j} \text{ for some } j \sim \text{Multinomial} [1; \pi_{ih,j0}(t), \dots, \pi_{ih,j} 2^{J-1}(t)] \quad (\text{A.25})$$

377 We note here that since the agents $A_h, h \in \mathbb{I}_{\mathcal{A} \setminus (\text{E} \cup \text{I})}$ are drawn with
 378 replacement as the computation proceeds and the agents $A_i, i \in \mathbb{I}$ are
 379 cycled through, if a previously drawn A_h is drawn again, but has already
 380 been infected in the current round then we ignore the latest event, but keep
 381 the previous infection event intact. To obviate bias in this procedure, we
 382 need cycle through the agents $A_i, i \in \mathbb{I}$ at random rather than in numerical
 383 order.

384 **5. Time t : disease progression loop.**

385 (a) *Individuals in \mathbf{A}_E at time t .* An individual $A_i \in \mathbf{A}_E$ at time t and in
 386 state $E_j(t, \tau_i), j = 0, \dots, 2^J - 1$, becomes either an individual in state
 387 $E_j(t + 1, \tau_i)$ with probability

$$\pi_{E_j}(t) = \frac{1}{1 + \left(\frac{t - \tau_i}{\sigma_{E_j}}\right)^4} \quad (\text{A.26})$$

388 or transfers to state $I_j(t + 1, \tau_i)$ with probability $(1 - \pi_{E_j}(t))$ thereby
 389 entering class \mathbf{A}_I at time $t + 1$.

390 (b) *Individuals in \mathbf{A}_I at time t .* An individual $A_i \in \mathbf{A}_I$ at time t and in
 391 state $I_j(t, \tau_i), j = 0, \dots, 2^J - 1$, becomes either an individual in state
 392 $I_j(t + 1, \tau_i)$ with probability

$$\pi_{I_j}(t) = \frac{1}{1 + \left(\frac{t - \tau_i}{\sigma_{E_j} + \sigma_{I_j}}\right)^4} \quad (\text{A.27})$$

393 or leaves the set $I_j(t + 1, \tau_i)$ with probability $(1 - \pi_{I_j}(t))$. In this
 394 latter case, the individual either dies with probability p_{α_j} or enters
 395 the state $V_j(t + 1, \tau_i)$ at time $t + 1$ with probability $1 - p_{\alpha_j}$. The total
 396 number of individuals dying over the interval $[t, t + 1)$ is noted as
 397 having a value $\Delta N_D(t)$.

398 **6. Time $t + 1$: outer loop update.** The outer loop records all the events
 399 that took place in the contact and disease progression loops and updates
 400 the agents state at the next time step. It also updates all other states as
 401 follows.

- (a) *Individuals in \mathbf{A}_S at time t .* For the $N_S(t)$ individuals in \mathbf{A}_S at time t , we have $N_{0j}^{E+}(t)$ enter set $\mathbf{A}_{E_j}(t+1)$ and we update

$$N_S(t+1) = N_S(t) - \sum_{j=0}^{2^J-1} N_{0j}^{E+}(t)$$

where Eq. A.23 ensures that $N_S(t+1) \geq 0$

- (b) *Individuals in \mathbf{A}_S that are infected again over $[t, t+1)$.* These individuals can become reinfected as calculated in the contact loop. Those that become reinfected with strain j , $j = 0, \dots, 2^J - 1$ enter state $\mathbf{E}_j(t+1, t+1)$ at time $t+1$.
- (c) *Updating the immunity of individuals in \mathbf{A}_S .* Every individual within \mathbf{A}_S at time t must have its immunity status updated so that for $j = 0, \dots, 2^J - 1$, if A_i is in state $V_j(t, \tau_{ij})$ at time t then it transfers to state $V_j(t+1, \tau_{ij})$ at time $t+1$, even if reinfected, as in b.) above.
- (d) *Transfer from \mathbf{S} to \mathbf{A} .* The $\hat{N}_{0j}^E(t)$ computed in Eq. A.24 become newly listed members of the set \mathbf{A} by entering state $\mathbf{E}_j(t+1, t+1)$, $j = 0, \dots, 2^J - 1$. This involves updating the equations for $N_S(t)$ and $N_A(t)$, including taking account of the number of individuals $\Delta N_D(t)$ that died from the disease in the immediate time period, i.e.:

$$\begin{aligned} N_S(t+1) &= N_S(t) - \sum_{j=0}^{2^J-1} \hat{N}_{0j}^E(t) \\ N_A(t+1) &= N_A(t) + \sum_{j=0}^{2^J-1} \hat{N}_{0j}^E(t) - \Delta N_D(t) \\ N_D(t+1) &= N_D(t) + \Delta N_D(t) \end{aligned} \quad (\text{A.28})$$

- (e) Along with input parameter values $t_{\text{vac.on}} \geq 0$, $t_{\text{vac.off}}$ and $p_v \in [0, 0.1]$, we also need to specify the valency of the vaccination by selecting 1 to 4 numbers that take on values $0, \dots, 2^J - 1$ (if more valencies are needed than 4, then the platform needs to be modified accordingly). We also need specify whether N_{select} will just be individuals in the set $\mathbf{S}(t)$ ($N_{\text{select}} = N_S$) or will be any individual other than those in the set $\mathbf{A}_I(t)$ ($N_{\text{select}} = N_S + N_A - N_I$).

In Algorithm 1 we summarise the steps of the simulation algorithm, as described in this section. On the right we report the name and numbering of the subsections while in the for loops we list the various steps respecting the item letters. Note that technical steps not explicitly described in the text (e.g. store updates, store set progression) do not present letters or numbers. The time set is defined with T while to describe temporal progression of set \mathbf{S} , \mathbf{A} and \mathbf{D} we use the symbols \mathcal{S} , \mathcal{A} , \mathcal{D} respectively.

Estimation of R_0 .

In a continuous time, SEIR model, when κ is folded into an all encompassing frequency-dependent transmission rate parameter $\beta_\kappa = \beta\kappa > 0$ (i.e.,

Algorithm 1: Summary of simulation algorithm

```

input  $N_0, J, \bar{\beta}_j, t_j^{\text{half}}, \sigma_{E_j}, \sigma_{I_j}, p_{\alpha_j}, j = 0, \dots, 2^J - 1$  // 1)
  parameters
   $\{(\bar{\eta}_j, \bar{\beta}_j, \lambda_j, p_{\alpha_j}; \zeta_{jm} \text{ and } c_{jm} \text{ for } m = 0, \dots, 2^J - 1)\}$  // 2a) pathogen
   $| j = 0, \dots, 2^J - 1\}$  // list
   $t = 0, N_I(0) = 1, N_S(0) = N_0 - 1$  // 2b) initialization
  for  $t$  in  $T$  do
    if  $N_I > 0$  then
      for agent in  $\mathbf{S} \cup (\mathbf{A} \setminus \mathbf{I})$  do // 3) vaccination loop
        3a) Vaccination process
        Update  $N_S$  and  $N_A$ 
      for agent in  $\mathbf{I}$  do // 4) contact loop
        4a) Numbers in various sets and associated index sets
        4b) Infectious contacts with each group
        4c) Identify all infectious agents and their pathogens strains
        4d) Susceptible contacts
        4e) Agent contacts
        Store updates
      for agent in  $\mathbf{E}$  do // 5a) disease progression loop
        5a) Individuals in  $\mathbf{A}_E$  at time  $t$ 
        Store updates
      for agent in  $\mathbf{I}$  do // 5b) disease progression loop
        5b) Individuals in  $\mathbf{A}_I$  at time  $t$ 
        Store updates
      Updates from loops 4), 5a) and 5b)
      6a) Individuals in  $\mathbf{A}_S$  at time  $t$  // 6) updates in outer loop
      6b) Individuals in  $\mathbf{A}_S$  that are infected again over  $[t, t + 1)$ 
      6c) Updating the immunity of individuals in  $\mathbf{A}_S$ 
      6d) Transfer from  $\mathbf{S}$  to  $\mathbf{A}$ 
      6e) Specify the valency of the vaccination
       $\mathcal{S}[t] \leftarrow \mathbf{S}(t)$  // store set progression
       $\mathcal{A}[t] \leftarrow \mathbf{A}(t)$ 
       $\mathcal{D}[t] \leftarrow \mathbf{D}(t)$ 
    return  $\mathcal{S}, \mathcal{A}, \mathcal{D}$ 

```

433 total is transmission given by $\frac{\beta SI}{N}$ [49, 50]), $\gamma > 0$ is the transition rate
434 from E to I and $\rho > 0$ is recovery rate from I to R, then R_0 corresponds
435 to [62]

$$R_0 = \frac{\beta \kappa}{\gamma + \rho} \quad (\text{A.29})$$

436 B J-RAMP Details

437 General description

438 The J-RAMP design augments a desktop simulation platform with several novel
439 features that increase flexibility and expressiveness, and promote experimenta-
440 tion and interoperability with other platforms. These include an API (“appli-
441 cation programming interface”) fully supporting remote operation and direct
442 retrieval of data for external processing on other platforms, such as Python,
443 Javascript or the R statistical platform. The API can also be accessed by an
444 onboard scripting interface that uses the Nashorn Javascript engine.

445 Additionally, using a novel design, elements of the internal algorithm are
446 exposed for possible reprogramming in a secure fashion that will not damage
447 the overall system. These runtime alternative modules (RAMs) may also be
448 controlled from the API to facilitate selective algorithm redefinition during the
449 run of the simulation.

450 Use of the J-RAMP features require some experience with scripting and/or
451 Java coding, however the resulting modifications to the algorithm can be of great
452 significance. The RAM platform is implemented to support program redefinition
453 with no risk to damaging the underlying code base. It should be accessible to
454 anyone with moderate scripting experience.

455 A major goal of the J-RAMP project is to pre-package these functionalities so
456 that they can be readily deployed as part of simulation system design. This goal
457 has been partially realized with respect to the RAM platform: annotations can
458 be added to the simulator’s source code that direct the automatic generation of
459 Java code to integrate into the simulations’s source and provide the functionality.

460 The following discussion assumes some familiarity with script or program
461 development.

462 Runtime alternative modules

463 Figure B.1 shows the RAM redefinition frame. The available RAMs appear as
464 radio buttons along the bottom of the frame. Each RAM is a set of options
465 for defining a relatively short Java method implementing some key aspect of
466 the simulation algorithm. For example, included in this simulation are the
467 implementation for cross immunity given in Eq. 1; the implementation for β
468 given in Eq. A.14; and the implementation for ϕ given in Eq. A.11; etc. Each
469 RAM initially contains only a single option, Option 0, the default, internally
470 defined implementation. Option 0 cannot be edited and appears for reference
471 purposes only.

472 Additional options may be added to each RAM containing code redefining
473 the method. Two editor panes and one console pane are stacked in the frame

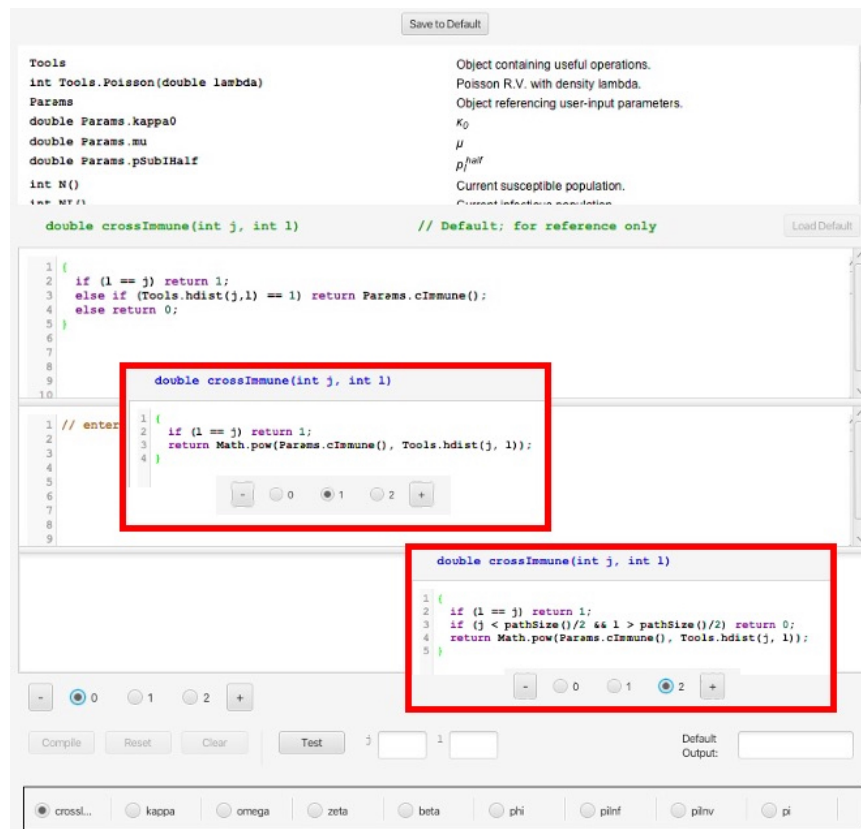


Figure B.1: RAM frame shows the implementation of the nearest-neighbor cross-immunity formulation \mathbf{C} of Eq. 1 as the default. In the red-bordered insets are the cascading cross-immunity with escape mutation formulations of Eqs. A.7 and A.9. Note the radio-buttons are set to 0 in the main figure and to 1 and 2 in the insets. Also note the “+” button which allows for an unlimited number of alternatives to be set up using consecutive integer numbers for the new radio-buttons that appear and pertain to the selection of each alternative. Note at the bottom the list of functions that can be altered at runtime. The “load default” button on the upper left hand side allows the user, when starting a new alternative, to insert the default code (which is immutable in radio-button 0) as a starting point. The frame also documents a list of terms in the upper panel that can be used to build any function.

474 and display the code and output of the RAM. These panes show the content
475 associated with the currently selected RAM and option. The top editor pane
476 contains the code for the method being redefined. The second editor pane
477 contains definitions of any new help functions required by the definition in the top
478 pane. The console pane contains messages and output that are useful during the
479 development of the option. For convenience, a “Load Default” button initializes
480 the editor to an editable version of the Option 0 default to use as a starting
481 point.

482 Figure B.1 shows the Option 0 default definition for the cross immunity

483 matrix function, as described by Eq. 1. Clicking the “+” button produced two
484 new options, which appear in the insets. These option implement the alternative
485 cascading cross immunity schemes presented in Eqs. A.7 and A.9. Option 2 code
486 appears below:

```
487     double crossImmune(int j, int k)  {  
488         if (l == j) return 1;  
489         if (j < pathSize()/2 && l > pathSize()/2) return 0;  
490         return Math.pow(Params.cImmune(), Tools.hdist(j, l));  
491     }
```

492 Note that we have substituted the function `pathSize` for a hard-coded value
493 of 63. `pathSize` returns the number of pathogens, allowing us to use this
494 formulation for any choice of entropy. Documentation for `pathSize` is at the
495 top of the window in the list of available help functions and parameters. There
496 is also more extensive documentation in a separate user guide (see Fig. B.2).

497 The platform duplicates a mini-development environment for building alter-
498 native definitions. Once code has been entered the “Compile” button checks
499 the legality of the code and makes it available for use at runtime. Legally com-
500 piled code will produce a “Compilation Successful” message. Errors will appear
501 with line numbers if they occur. Once the code is legal, the “Test” button can
502 be used with actual parameters entered into the small text fields to determine
503 correctness of the code. It is also possible to include `print` and `println` state-
504 ments in the code during development to further check correctness. Output
505 from print statements will appear in the bottom console window. The entire
506 RAM set can be saved and will reappear during subsequent launches of the
507 simulator platform.

508 To use an alternate RAM definition at runtime simply select the desired
509 option. (Selected options will be restored from a saved RAM set during subse-
510 quent launches.) The system will compile any uncompiled code the first time it
511 is accessed. If an error occurs during a runtime compilation an alert will notify
512 the user that the system is returning to the default definition of that RAM. At
513 no time is the internal logic of the program overridden.

514 Finally, RAM option selection is part of the API described in the next sec-
515 tion. This means that a script may run a simulation selecting different options
516 at different points in time, using logic that considers the state of the model. For
517 example, such an adaptive protocol might be appropriate for determining the
518 contact rate κ .

519 Application programming interface

520 The API is a simple bytecode¹ called BPL (Blackbox Programming Language)
521 that addresses all available user interactions with the simulator. Instructions
522 fall into three categories: parameter assignment and retrieval; simulator oper-
523 ation; and data retrieval. A complete list of instructions is shown in Fig. B.3.
524 Instructions are comprised of opcodes (e.g., `reset`, `step`, `get`) followed by 0
525 or more arguments. Every BPL operation returns a result, even if empty, for

¹a *bytecode* is computer source code that is processed immediately by a program, usually referred to as an interpreter or virtual machine.

J-RAMP SEIVD IBM A-Mod Guide

rms wgetz

May 2021

ω (Equation 9)

$$\omega_{ij}(t) = \begin{cases} 0 & \text{if } V_j \text{ is null} \\ \frac{1}{1 + ((t - \tau_{ij} - \sigma_I - \sigma_E)/t_j^{\text{half}})^\sigma} & \text{if } t \geq \tau_{ij} + \sigma_I + \sigma_E \\ 1 & \text{if } t < \tau_{ij} + \sigma_I + \sigma_E \end{cases}$$

Code

```
double omega(int i, int j)
{
    if (i == 0) return 0.0; // Special case; A_0 never infected
    InfectedAgentState astate = getStateOf(i, j);
    if (astate == null || !(astate instanceof V)) return 0.0;
    double tau = tau(i,j);
    double sigI = sigmaI(j);
    double sigE = sigmaE(j);
    double timeSinceInfect = tau + sigI + sigE;
    if (time() <= timeSinceInfect) return 1.0;
    return 1/(1 + Math.pow((time() - timeSinceInfect)/tHalf(j), Params.sigma()));
}
```

Notes

<pre>astate = getStateOf(i,j) astate == null !(astate == V) tau(i,j) sigmaI(j) sigmaE(j) tHalf(j) time() Params.sigma()</pre>	<pre>astate is assigned to a Java InfectedAgentState object for agent i with respect to pathogen j. InfectedAgentState objects can either be state E, I, V or null (the latter for never-infected or dead agents.) astate is something other than state V. = tau_ij. = sigma_I. = sigma_E. = tau_j^half. = t. = sigma.</pre>
--	--

Figure B.2: A description from our User Guide of the immunity waning function ω .

526 synchronization purposes. A string consisting of a sequence of opcodes and
 527 arguments may be submitted to the BPL interpreter, an example of which is
 528 shown in the notes in Fig. B.3.

529

530 **Parameter assignment and retrieval.** Every user-configurable element (in-
 531 cluding random number generator seeds) is addressed from BPL using a unique
 532 three-letter “airport code” (see Table. B1). Additionally, pathogens are ad-
 533 dressed by their id number (0 to $2^J - 1$) and agent states using identifiers S, E,
 534 I, V, DI+ and DD (the latter two represent ΔI and ΔD , respectively). RAM
 535 options are addressed in `setOption` and `getOption` using the name of the RAM
 536 (e.g., “crossImmune”). Get and set operations can be used on each of these with
 537 the exception of ENT (strain entropy), which is read-only.

538

539 **Simulator operation.** Simulation runs begin by executing the BPL `reset`
 540 instruction, followed by `step`, `run_for` or `run`. The BPL interpreter operates
 541 synchronously with the simulator by waiting to process subsequent commands
 542 during a simulation run. Operational instructions can be interspersed with pa-
 543 rameter set/get or data retrieval to use in runtime decision-making. Note that
 544 the `reset` operation restores the simulator to its state at the time of the last
 545 `reset`, so that no parameter changes made during a run are persistent.

546
 547 **Data retrieval** Operations to obtain the current population in each state, and
 548 to retrieve the runtime population history of each state and pathogen are also
 549 included. These can be easily transformed into R data frames, for example, for
 550 further analysis.

551
 552 Scripting can be deployed using either one of the two on-board script in-
 553 terpreter interfaces, or remotely from another platform using drivers provided
 554 with the simulator. The remote drivers use an operating system channel called
 555 *named pipes* or *fifo*s that are part of all MacOS, Windows, Linux, and other
 556 UnixTM-based systems.

557 On our main dashboard, we provide two scripting windows that are opened
 558 using the “S On” and “JS On” buttons (see button second and third from left at
 559 bottom of Fig. 2A). The former allows the user to write simulation driver scripts
 560 directly as command strings. (The commands listed in Fig. B.3 are accessed
 561 by pressing the “Command Reference” button in the “S On” window.) This
 562 window is used primarily to test and monitor scripts intended to be deployed on
 563 a remote platform. The JS window contains a Nashorn Javascript interpreter
 564 enhanced to accept and execute BPL operations. Scripts can developed, saved,
 565 and used to drive the simulator from this interpreter. For example, Fig. B.4, lists
 566 the code used to implement the adaptive vaccination programs. The SEIV object
 567 referenced in this code contains methods corresponding to the BPL operations
 568 detailed in Fig. B.3.

Table B1: Airport codes for parameter and variables used in the model algorithm

Name	Math	Code	Name	Math	Code
<i>Epidemic codes</i>					
% mortality	p_α	MOR	transmission	β	XMT
env. persist.	η	PST	within-host replication	λ	INV
median latent period	σ_E	MLP	median infectious period	σ_I	MIP
population size	N_0	POP	contact rate	$\bar{\kappa}$	CPT
mutation rate	μ	MUR	abruptness of waning	σ	AOW
cross immunity	c	CIM	waning half-life	t^{half}	WHA
seas. trans. perturb	δ	STP	seas. trans. shift	θ	STS
seas. trans. period	k	PER	adap. contact param.	p_I^{half}	IPC
strain entropy	J	ENT	shedding rate	ζ	SHD
<i>Vaccination codes</i>					
enable vaccination		VEN	vaccine on/off	$t_v^{\text{on}}/t_v^{\text{off}}$	VOO
strain valency type	VA#, #=1,...,4		selection composition		VCP
vaccinate susc. only		VSU	vaccinate non-infectious		NVI
vaccinate non-vacc.		VNV			

BPL Reference

Commands

reset	Resets simulation with initial slider/switch values.
restart	Resets simulation with current slider/switch values.
run_for T	Run the simulation from the current time for <i>T</i> steps.
run	Run the simulation for the number of steps set by the last run_for command.
step	Single-step the simulation one time unit.
get time	Returns current simulation time.
getLimit XXX	Returns lower and upper values of limit control labeled <i>XXX</i> on separate lines.
getPath XXX id	Returns value of parameter <i>XXX</i> in pathogen <i>id</i> .
get XXX	Returns value of non-limit control labeled <i>XXX</i> .
setLimit XXX lo hi	Sets the limit control <i>XXX</i> to limits <i>lo hi</i> .
setPath XXX id Y	Sets value of parameter <i>XXX</i> in pathogen <i>id</i> to <i>Y</i> .
set XXX Y	Sets value of non-limit control <i>XXX</i> to <i>Y</i> .
pathPop id Z	<i>id</i> is a pathogen id and <i>Z</i> must be one of the agent state identifiers (S, E, I, V, DI+, DD). Returns the current population in state <i>Z</i> with respect to pathogen <i>id</i> .
totPop Z	<i>Z</i> must be one of the agent state identifiers (S, E, I, V, DI+, DD). Returns the total population in state <i>Z</i> .
timeTrail	Returns the array of timesteps for use in a table with other trail data.
pathTrail id Z	<i>id</i> is a pathogen id and <i>Z</i> must be one of the agent state identifiers (S, E, I, V, DI+, DD). The population array for that pathogen and for that state over time is returned.
popTrail Z	<i>Z</i> must be one of the agent state identifiers (S, E, I, V, DI+, DD). The array of the entire population over time for that state is returned.
sleep T	NMBPL processor sleeps for <i>T</i> milliseconds. (May be used for synchronization during longer simulations.)

Notes

- Commands without explicit return do not return a value
- **get** and **set** use the 3-letter "airport codes" associated with each control
- Commands may be sequenced; e.g.

```
run_for 50 set XMT 1.5 run trail PATH I
```

will run for 50 timesteps, set the transmission parameter to 1.5, run for another 50, and return the 100 population values of I.

- Returned values are separated by newlines. When command sequences are sent remotely, left and right brackets (i.e. `\[` and `\]`) followed by newlines enclose the sequence of returned values, even if the sequence is empty; e.g.,

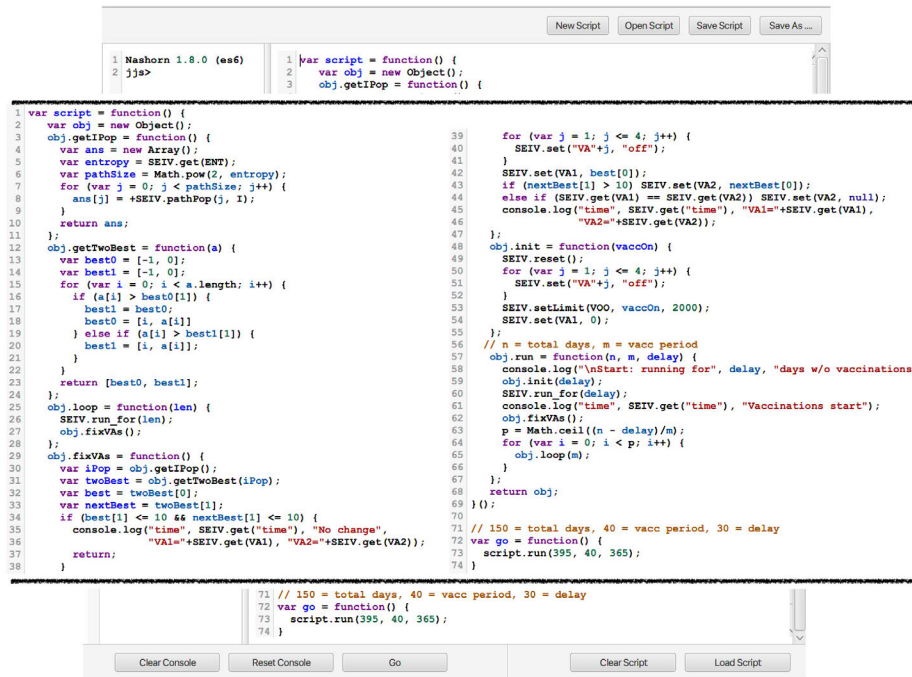
```
[\n50\n100\n1000\n]
```

- When using **trail**: if embedded in a command sequence, this must be the only command that returns a value (since the array returned by each of them will be embedded in the set of answers returned by the command sequence.)
- Use the console to test command sequences before using them remotely.
- Changes to the FIFOs, if saved to Preferences, will reappear in subsequent uses. Multiple instances of Numerus can run concurrently using different FIFOs.
- **Restart** button restarts the FIFO configuration.

Figure B.3: The list of Blackbox Programming Language (BPL) commands that can be used to write a simulation driver script, using the three-letter "airport codes" listed in Table A1 to access the parameters and variables in our coded algorithm. This list of commands can be accessed using the "Command Reference" button at the bottom of the Scripting Window.

569 R Integration

570 As previously mentioned, the API supports remote control of the simulator
571 from independent platforms using the operating system's indigenous *fifo* channels.
572 Of particular interest is integration with the R statistical programming



```
1 Nashorn 1.8.0 (es6)
2 jjs>
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
```

Figure B.4: The JS scripting window accessed by selecting the “JS On” button in the main dashboard (see button second from left at bottom of Fig. 2A). The script shown here was used to execute the adaptive vaccination strategy discussed in the main text.

573 environment. An R-package called “seiv” acts as a driver by synchronously is-
574 suing BPL command strings and waiting for results. Consequently, a simulation
575 can be driven entirely from within the R platform, treating the simulator as a
576 “virtual package”.

577 Fig. B.5 shows the code used to run the simulator multiple times with dif-
578 ferent random number generator seeds. Following each run, the time history of
579 the population in the I, DI+ and DD states is extracted directly to an R data
580 frame (without the need to save, for example, in a comma-separated list). At
581 the end of the run sequence the data frame is used to build the plots shown in
582 Fig. C.2.

583 R could be used in a more direct way by analyzing data at various points
584 throughout a single run and adjusting parameters programmatically, similar to
585 the adaptive vaccination strategy carried out in Javascript, only taking advan-
586 tage of the R environment’s powerful toolkit.

```
1 library("seiv")
2 library(ggplot2)
3 library(reshape)
4
5 execSEIVRemote <- function(seed=0, sampleSize=165, ...) {
6   who = list(...)
7   if (length(who) == 0) who = list("I", "DI+", "DD")
8   seivdispatch(paste("reset set RSD", seed, "run_for", sampleSize))
9   time <- 0:(sampleSize)
10  ans <- data.frame("Time"=time)
11  top <- sampleSize+1
12  for (z in who) {
13    nxt <- seivdispatch(paste("popTrail", z))
14    if (length(nxt) < top) nxt = c(nxt, integer(top-length(nxt)))
15    ans = cbind(ans, nxt)
16    names(ans)[length(names(ans))] = z
17  }
18  ans
19 }
20
21 multiSEIV <- function(runs=3, sampleSize=165) {
22   ans = list()
23   time <- 0:(sampleSize)
24   for (m in 1:3) ans[[m]] = data.frame("Time" = time)
25   for (n in 1:runs) {
26     print(n)
27     dat <- execSEIVRemote(n-1, sampleSize, "I", "DI+", "DD")
28     who <- c("I", "DI+", "DD")
29     for (m in 1:3) {
30       ans[[m]] = cbind(ans[[m]], dat[,c(m-1)])
31       names(ans[[m]])[length(names(ans[[m]])] + 1) = paste(who[m], "Run", n)
32     }
33   }
34   ans
35 }
36
37 buildMSFrame <- function(l) {
38   means = numeric()
39   stds = numeric()
40   for (m in 1:nrow(l)) {
41     smpl = as.numeric(l[m,2:ncol(l)])
42     means[m] = mean(smpl)
43     stds[m] = sqrt(var(smpl))
44   }
45   data.frame("Time"=l[,1], "μ"=means, "σ"=stds,
46             "μ-σ"=means-stds, check.names=FALSE)
47 }
48
49 buildGraph <- function(l, title) {
50   msFrame <- buildMSFrame(l)
51   molten = melt(msFrame, id.vars="Time")
52   ggplot(molten, aes(x=Time, y=value, col=variable)) +
53     geom_line() +
54     xlab("Day") +
55     ylab("Population") +
56     ggtitle(title)
57 }
58
59 go <- function(runs=30, sampleSize=165) {
60   data <- multiSEIV(runs, sampleSize)
61   list (
62     buildGraph(data[[1]], "Prevalence"),
63     buildGraph(data[[2]], "Incidence"),
64     buildGraph(data[[3]], "Daily Deaths")
65   )
66 }
```

R driver used to integrate SEIV-IBM J-RAMP into runs to generate plots

Figure B.5: Our SEIVD-IBM can be treated as an R-package called “seiv” and run as such in conjunction with other packages, such as ggplot2 and reshape to conduct multiple simulations and then carry out data and statistical analyses of the simulation results. The code shown here was used to produce the plots illustrated in Fig. C.2.

587 **C Additional Figures**

	N=10,000					N=30,000				
Run seed	S fnl	V fnl	Total	D	t fnl	S fnl	V fnl	Total	D	t fnl
0	2135	7700	165	166		6574	22964	462	212	
1	2234	7522	144	171		6431	23067	502	193	
2	2252	7602	146	158		6516	23020	464	175	
3	2174	7668	158	188		6568	22958	474	220	
4	2154	7684	162	162		6345	23214	441	191	
5	2127	7706	167	165		6520	22998	482	196	
6	2166	7662	172	159		6604	22950	446	198	
7	2280	7573	147	171		6282	23250	468	197	
8	2254	7613	133	161		6339	23184	477	187	
9	2153	7681	166	169		6529	23050	421	201	
Mean	2193	7641	156	167		6471	23066	464	197	
SD	56	61	13	9		114	112	23	13	
Mean %	0.219	0.764	0.0156			0.216	0.769	0.0155		
SD %	0.006	0.006	0.0013			0.004	0.004	0.0008		

Figure C.1: The final sizes of the uninfected ($S(t_{\text{fnl}})$ individuals), infected ($V(t_{\text{fnl}})$ individuals) and dead ($D(t_{\text{fnl}})$ individuals) classes and the length of the epidemic (t_{fnl} days) are provided here, together with summary statistics, for 10 simulations (runtime seed ranges from 0 to 9) in each case when the initial population sizes are $N_0 = 10,000$ and $N_0 = 30000$ individuals respectively. The parameter values used are otherwise the same as those used to produce the simulation depicted in Fig. 2 in the main text.

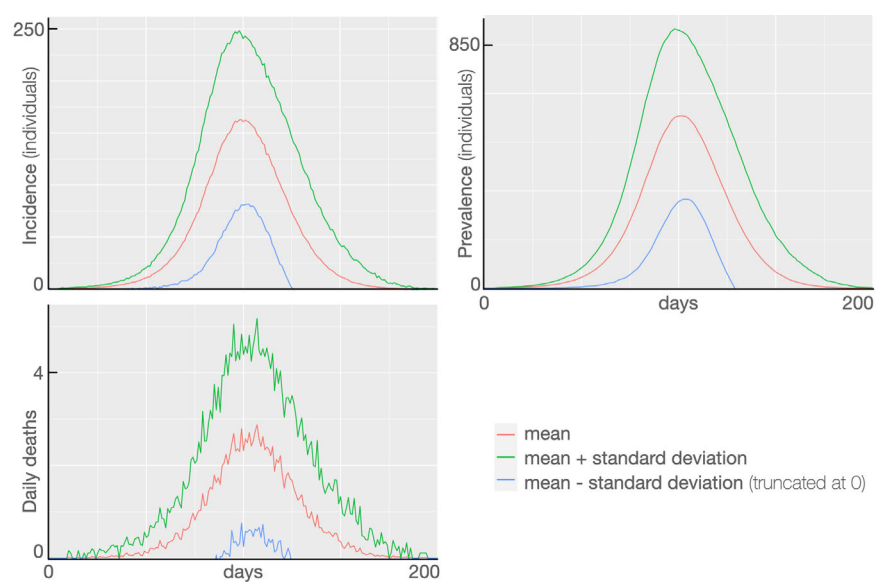


Figure C.2: Plots of the mean (red) \pm 1 standard deviation green and blue, the latter truncated at 0) of prevalence, incidence, and daily deaths for 100 runs (runtime seeds = 0,...,99) using the same parameter set used to produce the individual run (runtime seed = 0) depicted in Fig. 2 in the main text (also see Fig. C.1). These figures were produced by running our SEIVD-IBM J-RAMP as an R-package as described in Appendix B, with the code documented in Fig. B.5.

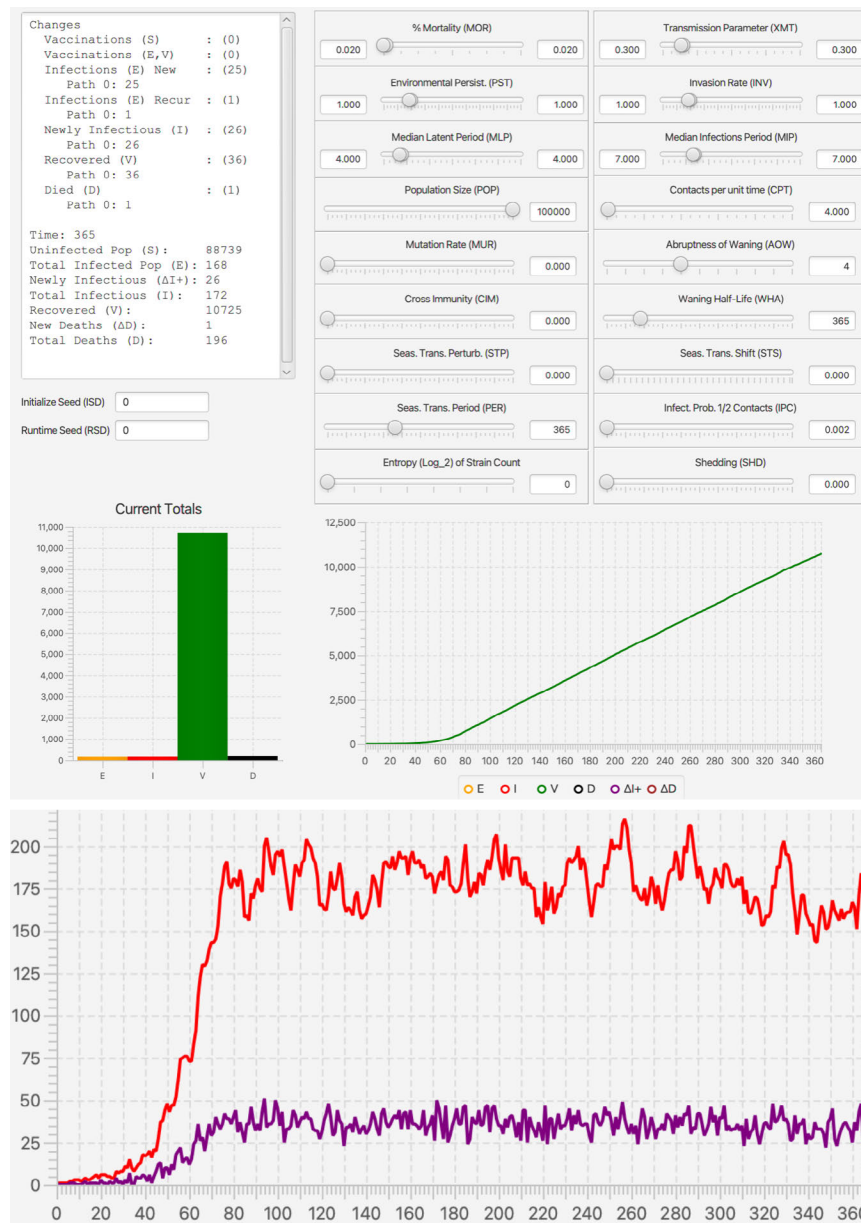


Figure C.3: If, together with our basic set of parameters (Table 1), we set $p_1^{\text{half}} = 0.005$ then the total number of individuals that have been infected (disease class V) at the end of one year is $10027/100,000 \approx 10\%$.

Strain	beta	Strain	beta	Strain	beta	Strain	beta	Strain	beta	Strain	beta	Strain	beta
0	0.300	16	0.149	32	0.246	48	0.295	64	0.235	80	0.152	96	0.196
1	0.325	17	0.097	33	0.244	49	0.076	65	0.121	81	0.127	97	0.172
2	0.007	18	0.117	34	0.011	50	0.289	66	0.249	82	0.138	98	0.177
3	0.350	19	0.089	35	0.029	51	0.155	67	0.079	83	0.019	99	0.165
4	0.252	20	0.222	36	0.162	52	0.043	68	0.133	84	0.258	100	0.037
5	0.217	21	0.292	37	0.214	53	0.110	69	0.127	85	0.181	101	0.087
6	0.105	22	0.222	38	0.102	54	0.263	70	0.136	86	0.042	102	0.230
7	0.375	23	0.214	39	0.036	55	0.049	71	0.184	87	0.284	103	0.096
8	0.214	24	0.205	40	0.010	56	0.037	72	0.166	88	0.218	104	0.219
9	0.256	25	0.100	41	0.173	57	0.108	73	0.015	89	0.249	105	0.058
10	0.217	26	0.089	42	0.093	58	0.201	74	0.048	90	0.025	106	0.033
11	0.151	27	0.026	43	0.256	59	0.041	75	0.131	91	0.052	107	0.219
12	0.009	28	0.296	44	0.236	60	0.159	76	0.013	92	0.160	108	0.151
13	0.077	29	0.065	45	0.038	61	0.053	77	0.207	93	0.121	109	0.005
14	0.026	30	0.019	46	0.186	62	0.186	78	0.019	94	0.234	110	0.110
15	0.400	31	0.425	47	0.078	63	0.450	79	0.149	95	0.165	111	0.214
												127	0.475

Figure C.4: This is the set of strain transmission parameter values $\beta_j, j = 0, \dots, 127$ used in our 128 strain simulation. Note strains 0, 1, 3, 7, 15, 63, 127 (light blue cells) are part of a dominating sequence of values increasing steadily in steps of 0.025 from $\beta_0 = 0.30$ to $\beta_{127} = 0.475$ along the one-step-neighborhood pathway represented by the allele structures $(0,0,0,0,0,0), (0,0,0,0,0,1), (0,0,0,0,0,1,1), \dots, (1,1,1,1,1,1,1)$. All the remaining β values were generated at random on the interval $[0, 0.3]$.

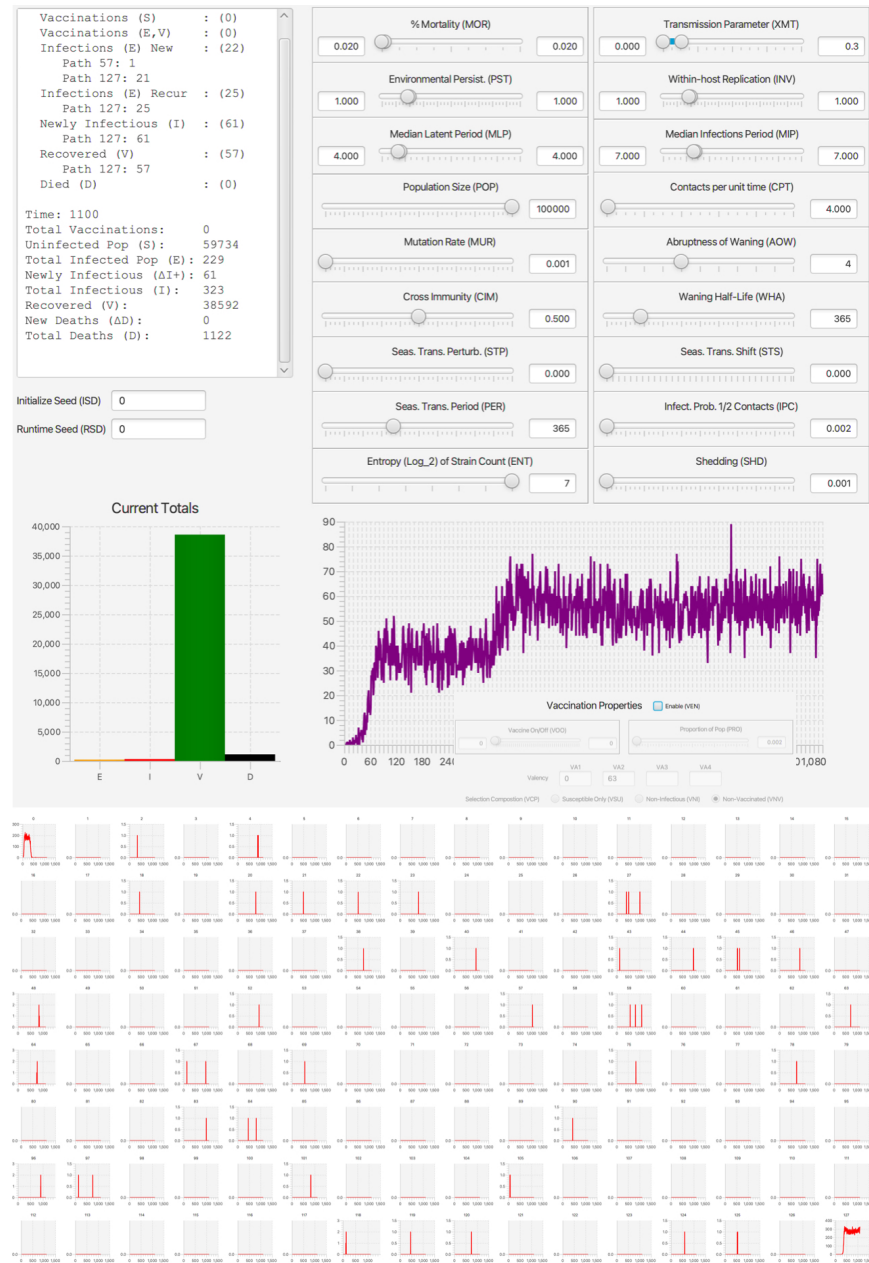


Figure C.5: This is a baseline 3-year no vaccination run for our 128-strain vaccination study with strain-specific transmission values depicted in Fig. C.4 and other parameter values as indicated on the dashboard (also see Table 1). See Fig. 6 in main text for enlarged plots of predominant strains.

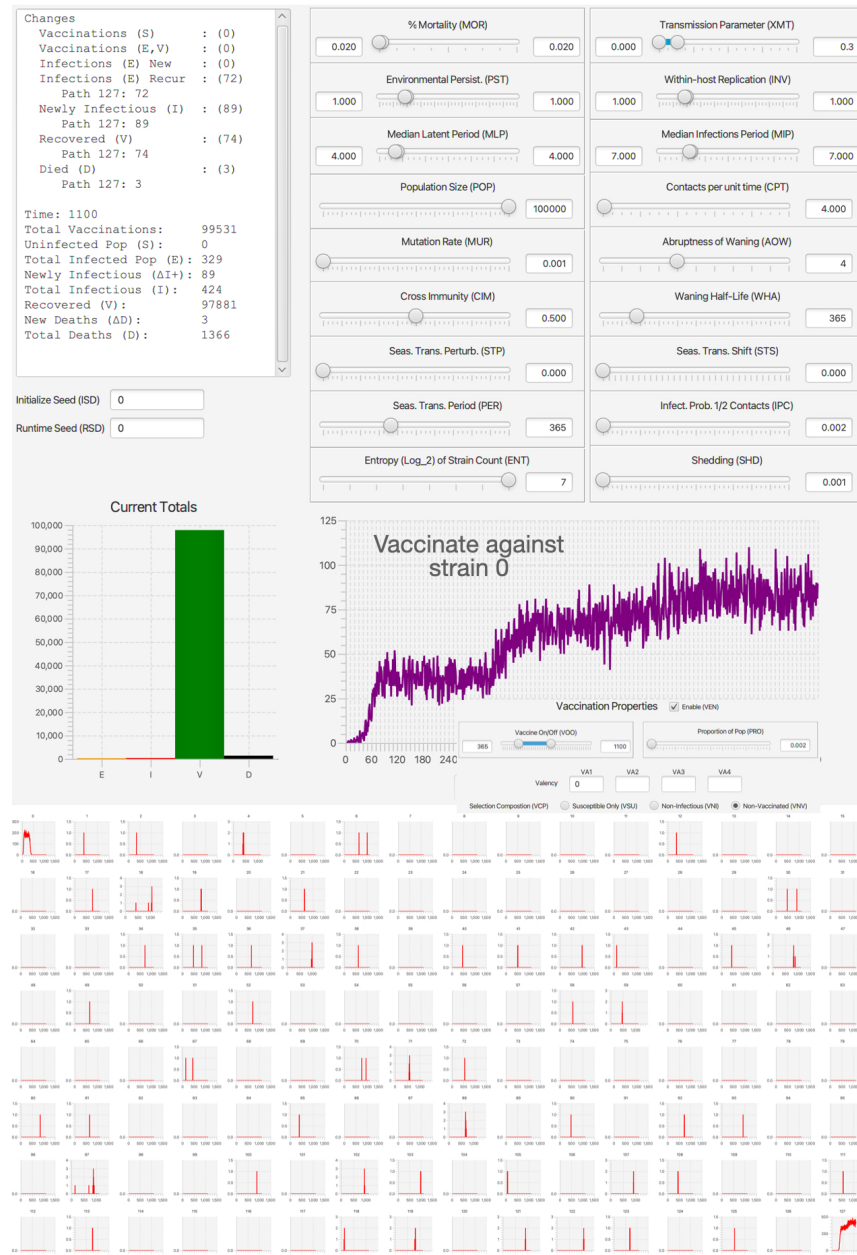


Figure C.6: This is vaccination in years 2 and 3 against strain 0 at the rate of 0.2% of the population per day in our 128-strain vaccination study with strain-specific transmission values depicted in Fig. C.4 and other parameter values as indicated on the dashboard (also see Table 1). See Fig. 6 in main text for enlarged plots of predominant strains.

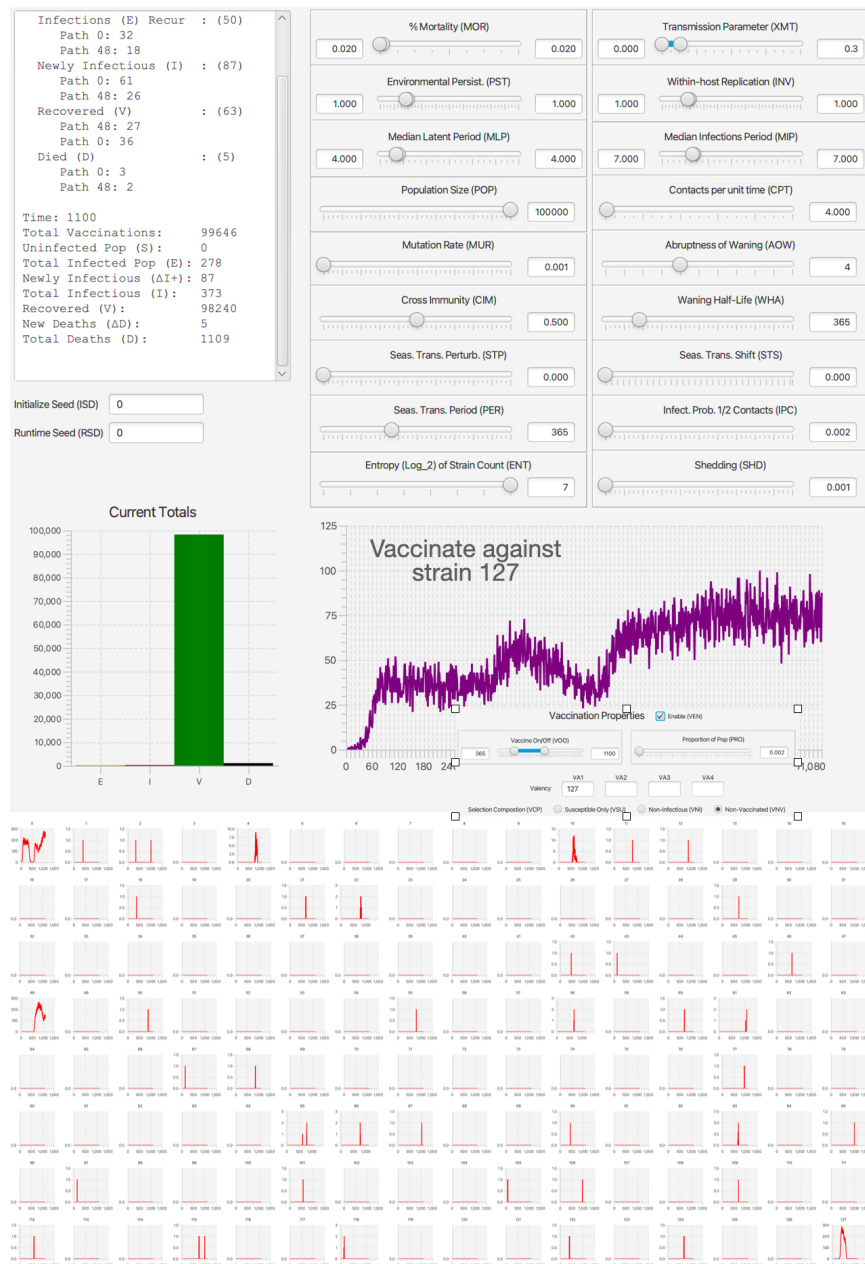


Figure C.7: This is vaccination in years 2 and 3 against strain 127 at the rate of 0.2% of the population per day in our 128-strain vaccination study with strain-specific transmission values depicted in Fig. C.4 and other parameter values as indicated on the dashboard (also see Table 1). See Fig. 6 in main text for enlarged plots of predominant strains.

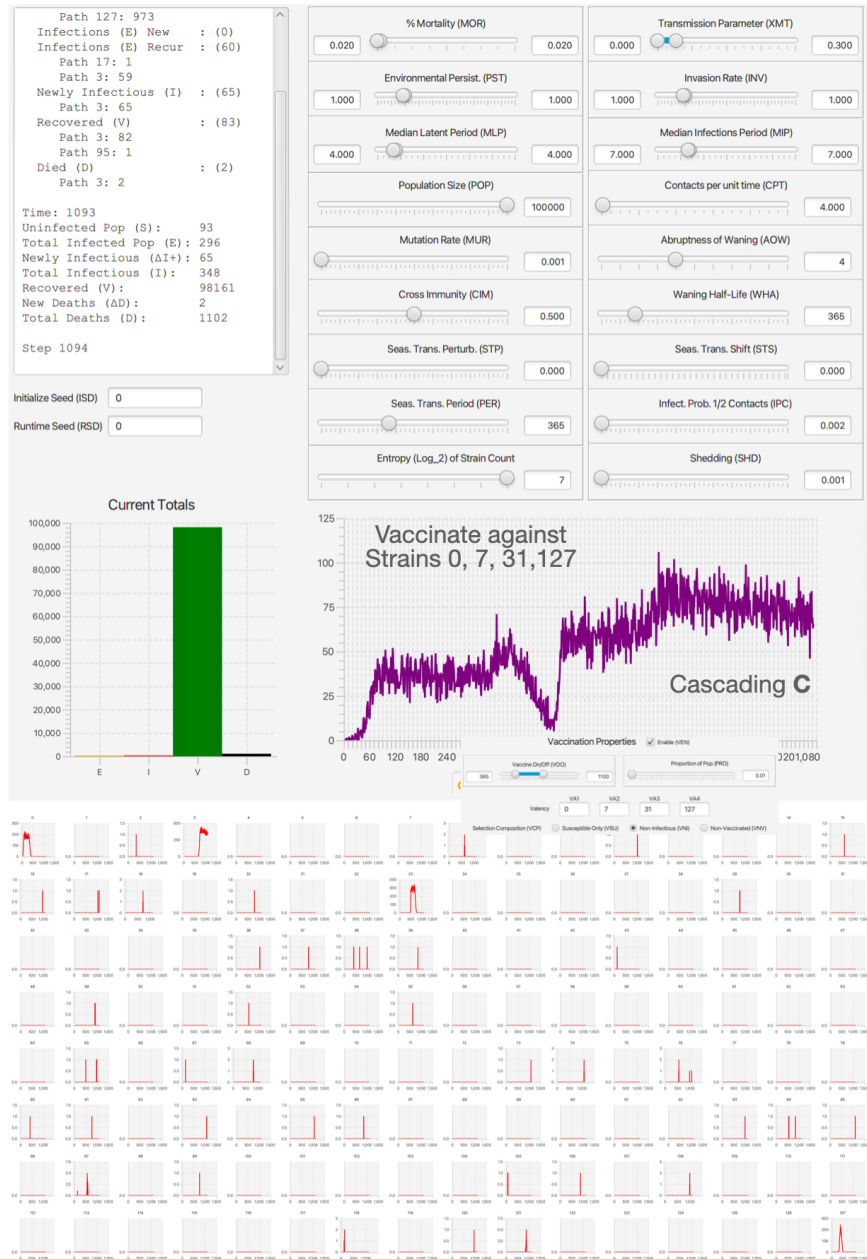


Figure C.8: This is a multivalent vaccination in years 2 and 3 against strains 0, 7, 31 and 127 at the rate of 1.0% of the population per day in our 128-strain vaccination study with strain-specific transmission values depicted in Fig. C.4 and other parameter values as indicated on the dashboard (also see Table 1), except that the cross immunity matrix C is now defined by Eq. A.7. See Fig. 6 in main text for enlarged plots of predominant strains.

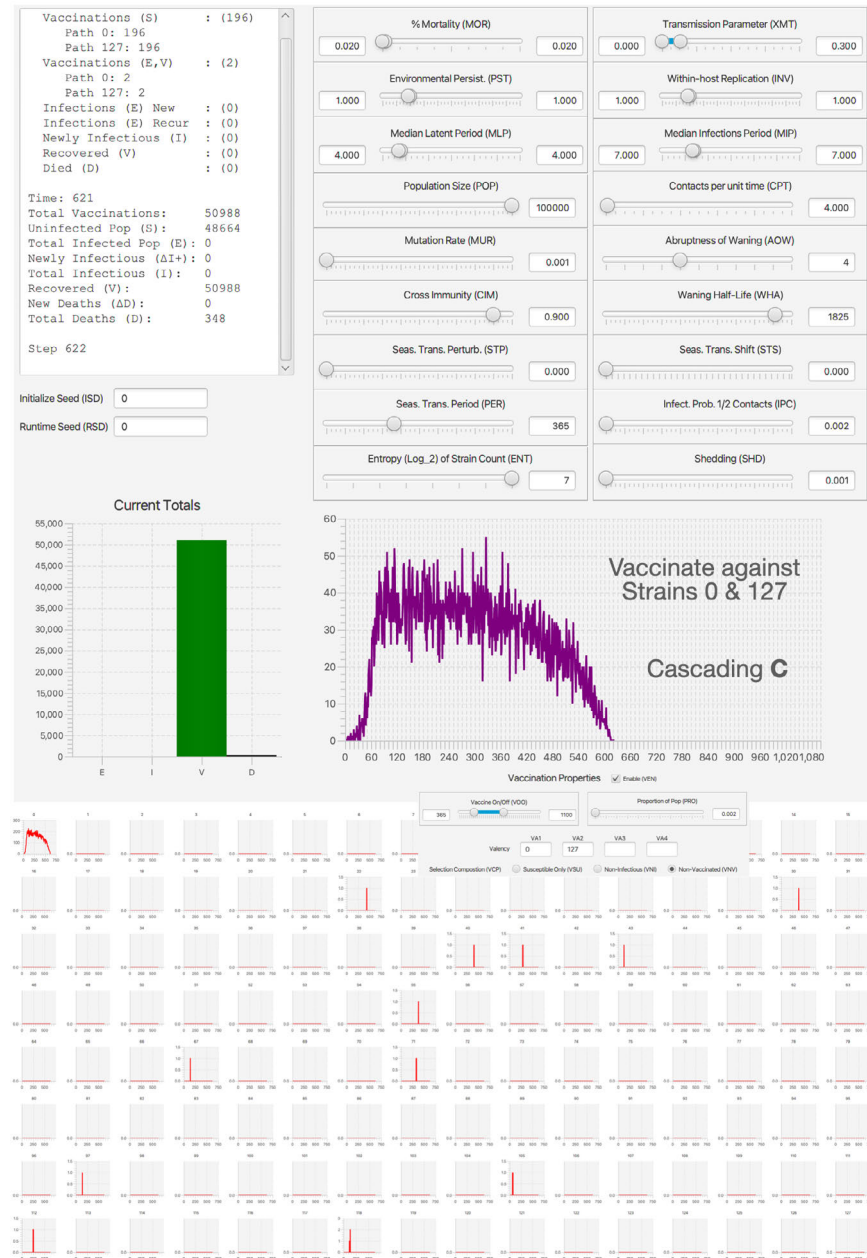


Figure C.9: This is a bivalent vaccination in years 2 and 3 against strains 0 and 127 at the rate of 0.2% of the population per day in our 128-strain vaccination study with strain-specific transmission values depicted in Fig. C.4 and other parameter values as indicated on the dashboard (also see Table 1), except that the half-waning values has been increased fivefold $t^{\text{half}} = 1825$ days the cross immunity matrix \mathbf{C} is now defined by Eq. A.7 with $c = 0.9$. See Fig. 6 in main text for enlarged plots of the predominant strain 0.