

1 **Containment of future waves of COVID-19: simulating the impact of different policies and**
2 **testing capacities for contact tracing, testing, and isolation**

3 Vincenzo G. Fiore^{1*}, Nicholas DeFelice², Benjamin S. Glicksberg^{3,4}, Ofer Perl¹, Anastasia
4 Shuster¹, Kaustubh Kulkarni¹, Madeline O'Brien¹, M. Andrea Pisauro⁵, Dongil Chung⁶, Xiaosi Gu¹

5 ¹ Icahn School of Medicine at Mount Sinai, Department of Psychiatry, 1 Gustave L. Levy Pl, New
6 York, NY 10029

7 ² Icahn School of Medicine at Mount Sinai, Department of Environmental Medicine and Public
8 Health, 1 Gustave L. Levy Pl, New York, NY 10029

9 ³ Icahn School of Medicine at Mount Sinai, Department of Genetics and Genomic Sciences, 1
10 Gustave L. Levy Pl, New York, NY 10029

11 ⁴ Icahn School of Medicine at Mount Sinai, Hasso Plattner Institute for Digital Health at Mount
12 Sinai, 770 Lexington Ave, 14th Floor, New York, NY 10065

13 ⁵ Wellcome Centre for Integrative Neuroimaging, University of Oxford, Department of
14 Experimental Psychology, New Radcliffe House, Walton St, Oxford OX2 6GG

15 ⁶ Ulsan National Institute of Science and Technology, Department of Human Factors Engineering,
16 50 UNIST-gil, Ulsan, South Korea

17

18 *Corresponding author:

19 Vincenzo G Fiore, Ph.D.

20 Icahn School of Medicine at Mount Sinai, 1 Gustave L. Levy Pl, New York, NY 10029

21 E-mail address: vincenzo.fiore@mssm.edu

22 **Abstract**

23 We used multi-agent simulations to estimate the testing capacity required to find and isolate a
24 number of infections sufficient to break the chain of transmission of SARS-CoV-2. Depending on
25 the mitigation policies in place, a daily capacity between 0.7 to 3.6 tests per thousand was required
26 to contain the disease. However, if contact tracing and testing efficacy dropped below 60% (e.g. due
27 to false negatives or reduced tracing capability), the number of infections kept growing
28 exponentially, irrespective of any testing capacity. Under these conditions, the population's
29 geographical distribution and travel behaviour could inform sampling policies to aid a successful
30 containment.

31

32 **Keywords**

33 Coronavirus, COVID-19, SARS-CoV-2, pandemic, second wave, multi-agent simulations

34

35 1. Introduction

36 In December of 2019 a cluster of cases of pneumonia was recorded among people associated with
37 the Huanan Seafood Wholesale Market in Wuhan, Hubei Province in China¹. They were infected
38 with a novel coronavirus, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). In a few
39 months, the virus spread rapidly around the world and with no vaccine or identified treatment it has
40 forced a growing number of countries to implement robust non-pharmacological policies of social
41 distancing, such as stay-at-home orders. As of early June 2020, the world health organization
42 (WHO) has reported that these mitigation strategies have been successful in containing new daily
43 infections, where applied¹. However, these measures have determined a high social and economic
44 cost, leading policy makers to consider different follow-up strategies for mitigation and
45 containment. The objective for the months to come has now shifted towards the need to allow as
46 many people as possible to go back to work safely, in order to reduce the economic impact of the
47 pandemic, while avoiding or at least mitigating a second wave of infections that could overwhelm
48 the healthcare system².

49 To this end, the WHO has suggested that enhanced capacity for contact tracing and testing is
50 necessary to continuously monitor the intensity and geographical spread of the virus, detect new
51 outbreaks at their onset, isolate (i.e. quarantine) new infections and prevent subclinical
52 transmission. Contact tracing, testing and isolating potential vectors of the disease is the key public
53 health process that has been used for decades to break the chain of transmission of a disease³⁻⁵.
54 This strategy aims at identifying those individuals who have come into contact with an infected
55 person in order to prevent new pre-symptomatic viral shedding⁶, promoting targeted isolation
56 whenever necessary. The high transmissibility of SARS-CoV-2⁷ and the understanding that the
57 onset of viral shedding precedes the manifestation of symptoms^{8,9} have already led policy makers
58 to vastly increase testing and contact tracing capacity. At the same time, they are also tasked with
59 striking a balance between the need to trace back the movements of anyone who tested positive for
60 the virus (and those of her contacts) and the concern that legitimate health-related tracking policies
61 may deteriorate into a form of mass surveillance. This trade-off comes in a time of crisis for
62 democracies across the world with the associated disenfranchisement of many citizens and mistrust
63 of government communication, a phenomenon that reduces voluntary compliance to tracking
64 methods. Furthermore, despite the general consensus that “more is better”, it remains to be
65 determined the order of magnitude of the resources required to appropriately identify, test and
66 isolate a number of pre-symptomatic infections sufficient to reduce the SARS-CoV-2 transmission
67 to a tolerable societal risk¹⁰. It is also unclear how these estimates may dynamically change as a
68 function of efficacy and reliability of testing and tracing^{11,12}, or depending on which non-

69 pharmacological policies are implemented, leaving policy-makers to rely on heuristic approaches
70 which may lead to severe and systematic errors.

71 Here we used a series of multi-agent simulations^{13,14} to highlight emergent dynamics in the
72 interaction between agents, environment, viral transmission and testing policies. The simulations
73 were aimed at estimating ecologically plausible capacities for contact tracing and testing that would
74 allow to identify and isolate a number of infections sufficient to break the chain of transmission of
75 the virus. In the first set of simulations, we used fifteen conditions in a 3 (disease incidence) x 5
76 (contact tracing and testing efficacy) design. Three different levels of incidence values of the
77 disease, e.g. due to different non-pharmacological mitigation policies in place, determined three
78 growth rates in the number of daily infections. Five different rates of contact tracing and testing
79 efficacy determined the number of infected subjects that were found or missed by the testing
80 process, e.g. due false negatives or untraced contacts. These simulations highlighted a systemic
81 failure with the process of contact tracing and testing, which arises when its efficacy drops below a
82 threshold that varies as a function of the disease incidence.

83 In the second set of simulations, we investigated whether different sampling strategies for testing
84 could be used to aid of the contact tracing process, to find infected agents missed by this process
85 and signal the presence of new outbreaks. For these simulations, we used the lowest value for
86 contact tracing and testing efficacy (20%), jointly with the medium incidence setting (25% increase
87 in daily infections), across three different conditions of population density, geographical
88 distribution and simplified travel habits for the population. These simulations showed a possible
89 solution to overcome the systemic failure reported for low efficacy contact tracing and testing that
90 relies on population-level analysis of geographical distribution and travel behaviour, thus mitigating
91 mass surveillance concerns.

92

93 **2. Results**

94 **2.1 Estimating the optimal contact tracing capacity.**

95 Our first set of simulations covered 60 days across 50 scenarios (i.e. 50 random seeds) and 15
96 environmental conditions, in a 3 (disease incidence) x 5 (contact tracing and testing efficacy)
97 design. Each scenario assumed an initial number of ~50 infected agents¹⁵, uniformly distributed in a
98 population of 100,000 simulated agents. Three tested levels of incidence determined a number of
99 new infections equivalent to a daily incidence growth of 35%, 25% or 15% of the number all non-
100 isolated infected agents: i.e. in the absence of any containment strategy the number of infections

101 doubled approximately every 2.5, 3 or 6 days, mimicking different rates of exponential growth of
102 infections reported across countries in early 2020 for COVID-19¹. To simulate the impact of
103 missed contacts or false negative tests, five rates of contact tracing and testing efficacy were
104 examined, controlling the percentage (100%, 80%, 60%, 40% or 20%) of infected agents that would
105 be found positive, among those who had been in contact with and infected by any agent who tested
106 positive. Finally, we assumed virus shedding started three days prior to showing symptoms⁹. All
107 symptomatic infected agents were assumed to be isolated (i.e. quarantined at home or in a hospital)
108 and therefore could not further contribute to virus transmission after symptom onset. Based on
109 preliminary reports, we estimated that 20% of infected agents would not show symptoms severe
110 enough to induce self-isolation or hospitalization^{16,17}. Therefore, in our simulations, these
111 asymptomatic agents kept propagating the disease until fully recovered⁹ or found positive in a
112 targeted test (e.g. they were found with contact tracing), in which case they would then be
113 considered isolated at home.

114 For eight of the fifteen simulated conditions we found a parameterisation that resulted in the
115 suppression of the virus transmission (effective reproduction < 1 ; Table 1, Figure 1). These
116 conditions were characterised by high ($\geq 60\%$) contact tracing and testing efficacy, and a testing
117 capacity between 0.7 (low incidence) and 3.6 (high incidence) per thousand agents. Under four of
118 the remaining conditions (Table 1), the simulations indicated a testing capacity varying between 0.7
119 (low incidence) and 4.5 (high incidence) could contain the virus transmission, but the number of
120 agents tested and isolated was not sufficient to reduce the daily number of new infections to zero.
121 Instead, the daily number of infected agents remained stable or slightly decreased on average across
122 the simulated scenarios (effective reproduction ≈ 1 ; Figure 1a,e). Similarly, for the remaining three
123 conditions characterised by low contact tracing and testing efficacy (20% and 40%) and medium or
124 high incidence (number of infection growth rate: 35% and 25%), the exponential growth of
125 infections could not be contained, independent of the capacity available (effective reproduction > 1 ;
126 Table 1, Figure 1c,e).

127 These capacity values were found in a heuristic research (see methods) and they were set as the
128 lowest sufficient to either stabilising or reducing the number of infected agents per day (Table 1).
129 Importantly, for those conditions showing that the process of contact tracing, testing and isolation
130 was insufficient to suppress or halt the virus transmission, the testing capacity was set above the
131 daily number of tests actually performed. Thus, the failure in containing the spread of the disease in
132 presence of low efficacy is systemic: it is not necessarily due to the availability of tests *per se*, as a
133 further increase in daily capacity did not affect this result (Figure 1d,f). Instead, the high percentage
134 of missed contacts enhanced a predator-prey dynamic (i.e. Lotka-Volterra non-linearity¹⁸; Figure

135 2), where the predators (the tests) lost track of too many preys (the infections), and remained idle
136 (part of the daily test capacity remained unused). As a consequence, the disease spread undetected,
137 thus leading to a minimally mitigated “wave” of infections.

138 **2.2 Solutions for the low efficacy contact tracing and testing systemic failure.**

139 The straightforward solution to this failure is to increase the efficacy of contact tracing, for instance
140 by enhancing various forms of movement surveillance. Our findings indicated that 60% of contact
141 tracing and testing efficacy allowed for tracing, testing and isolating a sufficient number of
142 infections to contain or suppress the virus transmission, irrespective of the simulated incidence.
143 However, we decided to assess also alternative routes for mitigating the virus transmission,
144 considering that the idle daily testing capacity could be used to find “preys” missed by the contact
145 tracing “predators”, thus providing new traces to follow. To this end, we simulated and compared
146 the effects of multiple testing policies in support of contact tracing (see Methods for details), in a 3
147 (geographical distribution) x 2 (distributions of travel behaviours) design. We simulated the effects
148 of using these mixed policies in three geographical maps of population density (New York
149 metropolitan area, Southeast Italy, and the Midlands in UK, Figure 2), combined with different
150 distributions of travelling behaviours for the agents. We devised three cohorts of travel behaviour
151 (Figure 2a): one marked agents moving in a small size sector (e.g. within one city or one borough),
152 a second for a medium size sector (e.g. comprising two boroughs or two separate towns, depending
153 on the map) and a third of agents freely moving in the entire map. We tested two distributions of
154 agents among the three cohorts: a “quasi uniform” distribution where the three cohorts consisted of
155 40%, 30% and 30% of the simulated agents; a “skewed” distribution, where the small size sector
156 cohort comprised 80% of the population, while 10% of the agents were included in the remaining
157 two. Each member of any of the three cohorts had the potential to shed the virus anywhere in the
158 simulated environment within the limits of their travel range. For these simulations, we kept
159 constant both the incidence (25% daily growth in the number of non-isolated infections) and the
160 contact tracing efficacy (20%), as a proof of concept. This second simulation set covered 60 days,
161 across the same 50 scenarios controlled by random seeds used for the first set.

162 First, the simulations showed the policy of contact tracing, testing and isolation mitigated virus
163 transmission in similar ways across geographical and travel behaviour distribution (Figure 3). The
164 mean number of infected agents recorded at day 60 across the 50 simulated scenarios was
165 104.74 ± 32.04 and 104.94 ± 31.97 , respectively, for the New York metropolitan area, with the quasi
166 uniform (NY₁) and skewed travel distribution (NY₂). Similarly, for southeast Italy, we found
167 100.18 ± 31.76 and 101.46 ± 34.81 infected agents in association with the same two distributions of

168 travel behaviours ($seIT_1$ and $seIT_2$, respectively). Finally, for the map of Midlands, UK, we found
169 109.32 ± 30.36 and 101.4 ± 34.44 infected agents (Mid_1 and Mid_2 , respectively). A 3x2 within
170 scenario repeated measures ANOVA revealed no significant effect for the geographical distribution
171 (Sphericity assumed, $F(2,98)=.87$, $p=.42$), distribution of travel behaviours ($F(1,49)=.62$, $p=.43$) or
172 interaction effect ($F(2,98)=.86$, $p=.42$).

173 Second, we found that the unused capacity available for contact tracing and testing could be
174 employed in further testing policies, marking an improvement in terms of the isolation of infections
175 and thus reducing virus transmission. For instance, contact tracing coupled with random sampling
176 from the entire map significantly reduced the number of infections recorded by the last day of
177 simulation across all scenarios (NY_1 : 86.62 ± 32.04 , NY_2 : 79.08 ± 31.97 , $seIT_1$: 80.6 ± 31.76 , $seIT_2$:
178 88 ± 34.81 , Mid_1 : 86.36 ± 28.83 , Mid_2 : 79.74 ± 28.33 for the 3x2 conditions), when compared with the
179 same scenarios under the use of contact tracing, testing and isolation, alone. A 2x3x2 within
180 scenarios repeated measures ANOVA reported a significant effect for the testing policy (Sphericity
181 assumed, $F(1,49)=115.59$, $p<.0001$) and no significant effect for any other factor (geography:
182 $F(2,98)=.19$, $p=.83$, travel behaviour: $F(1,49)=1.65$, $p=.2$) or interaction of factors
183 (policy*geography: $F(2,98)=1.04$, $p=.36$, policy*travel behaviour: $F(1,49)=.0$, $p=.97$,
184 geography*travel behaviour: $F(2,98)=2.58$, $p=.08$, policy*geography*travel behaviour:
185 $F(2,98)=1.03$, $p=.36$; Figure 3).

186 Finally, we found geography- and behaviour-specific policies that further improved the containment
187 of the disease beyond the added benefit of random sample testing. In particular, we tested different
188 sampling methods where we limited the targeted area for the sampling to either a small or a medium
189 size cell (equivalent to the small and medium size travel range sectors, Figure 2a) and we sampled
190 the population giving different probability weights to different travel cohorts. We found that travel-
191 weighted sampling of the population localised in a small cell around the most recent outbreaks,
192 determined on a day-by-day basis, would successfully aid contact tracing and testing (Figure 3). We
193 found that the optimal weights of these outbreak-centred sampling policies varied as a function of
194 the population distribution on the maps and the distribution of travel behaviours. In t-test
195 comparisons between these mixed policies and the joint use of contact tracing and random sampling
196 over the entire population, the mixed policies were found to significantly reduce the number of
197 infected agents by day 60 of the simulation, for NY_1 (73.3 ± 26.69 : $t(49)=3.29$, $p=.002$; Figure 3a),
198 NY_2 (60.34 ± 29.29 : $t(49)=4.12$, $p=.0001$; Figure 3b), $seIT_2$ (69.78 ± 30.51 : $t(49)=4.31$, $p<.0001$;
199 Figure 3d), Mid_1 (72.52 ± 27.94 : $t(49)=2.86$, $p<.006$; Figure 3e) and Mid_2 (48.68 ± 27.8 : $t(49)= 8.49$,
200 $p<.0001$ Figure 3e). We did not find a geography- and behaviour-specific policy -among those
201 tested- that led to significant improvements in comparison with the combined contact tracing and

202 map-wide random sampling for seIT₁ (the best option led to a mean of 76.62±31.48: t(49)= .7,
203 p=.48; Figure 3c). For the seIT₁, we also did not find a mixed policy that would result in halting or
204 marking a decline in the number of infections by the end of the simulations. Intuitively, where
205 mixed testing policies are in place, an increase in testing capacity does result in an increase in the
206 number of infected agents found and isolated, due to the increased sampling. Our simulations
207 confirmed this expectation as we increased the testing capacity for the condition seIT₁, from 3 to 4
208 per thousand agents, determining a successful containment of the disease (supplementary Figure 1).

209

210 3. Discussion

211 The sweeping stay-at-home orders that have been put in place across the world to contain the
212 transmission of the novel coronavirus SARS-CoV-2 were urgently needed to avoid overwhelming
213 the healthcare systems, but they have also had a high social and economic negative impact. As
214 multiple countries are planning to re-open or are in the process of re-opening industries and
215 services, policy-makers are developing strategies that are meant to prevent a second wave of
216 infections². Here, we estimated the testing capacity required to identify and isolate a sufficient
217 number of infected subjects to break the chains of transmission of SARS-CoV-2, therefore allowing
218 to contain the impact of the disease and avoid the most severe containment measures. For these
219 estimations, we used multi-agent simulations in a soft artificial life approach¹³, in place of well-
220 known statistic and mechanistic models^{2,10,19,20}, so to highlight emergent properties and dynamics
221 resulting from the interaction between simulated agents, environment and containment policies.

222 We found that several variables affect the estimations of the testing capacity. Some of these
223 variables are related to the disease and are yet to be fully understood, e.g. the percentage of
224 asymptomatic or paucisymptomatic among the infected^{16,17,21,22}, or the timing for the viral shedding
225^{8,9}. Other variables pertain to external factors and should be considered when planning for testing
226 policy development. To account for this uncertainty, we simulated different levels of disease
227 incidence, resulting in different rates of growth of the number of infections, putatively simulating
228 the effects of different mitigation policies in place (social-distancing, face masks, etc.). Second, we
229 simulated fifty different starting conditions in terms of the number of imported infections and
230 location of infected agents at day 1 (i.e. at the start of the new policy implementation). Finally, we
231 simulated the efficacy of contact tracing and testing to account for the reliability of the tests^{11,12} and
232 the ability of a country or region to trace movements and contacts of new found infections. Despite
233 these substantial simplifications, our simulations indicated a few key insights that can guide policy
234 development.

235 First, we found that at low levels of efficacy (i.e. low reliability of tests and/or low ability to trace
236 contacts), a high capacity of testing dedicated only to contact tracing, testing and isolation remains
237 partially unused, despite a growing number of infections. This is a striking result considering that in
238 most countries indeed failed in uncovering the real dimension of the disease prevalence, as clearly
239 demonstrated by the wide differences among case fatality rates across countries (4.7% in Germany,
240 5.9% in USA, 14% in UK or 14.3% in Italy), including those that were successful in containing the
241 disease (e.g. 0.5% in Iceland, 1.9% in New Zealand, 2.3% South Korea)¹. This result may also
242 explain why several countries prepared for a significant amount of tests per day and a diffused
243 network for contact tracing could not use their available capacity and failed to contain COVID-19,
244 eventually reporting a first wave of infections (e.g. as indicated for Germany²³). The virus
245 transmission may be intrinsically associated with reduced contact tracing efficacy, for instance due
246 to the relevance of indirect forms of propagation (e.g. droplets left on a handle in public transport
247 caught by a passenger hours or even days later^{24,25}). In this case, our simulations suggest that
248 improvements in contact tracing and testing efficacy, e.g. due to increased reliability of tests or in
249 implementing tracking methods for the population, are required to exceed a threshold of 60%.
250 Beyond this value, jointly with the safe isolation of all symptomatic subjects and those who tested
251 positive, we found a steady decline in the number of infections across all simulated levels of
252 incidence. Importantly, this result is consistent with an estimation recently provided in a
253 mechanistic model¹⁰, demonstrating robustness of the finding across theoretical constructs.

254 Second, our data suggest that improvements in the containment of the disease can be achieved with
255 mixed testing policies. These policies combined contact tracing with independent testing of selected
256 samples of people, so aiding the monitoring of new outbreaks and feeding missed contacts to the
257 main process of contact tracing. Importantly, these mixed policies were designed to use the entire
258 testing capacity that remained after exhausting the needs of the process of contact tracing and
259 testing. Therefore, an increase in capacity in presence of these mixed policies improved the
260 containment of the virus transmission, due to the increased ability to find and isolate new infected
261 agents independent of existing traces. Finally, while the process of contact tracing and testing is
262 agnostic to both the geographic distribution and the population-level behaviours, the simulations
263 showed that optimal aiding policies are shaped by the features of the environment and cohort-level
264 behaviour²⁶. Thus, the use of these mixed policy would reduce the necessity for mass surveillance,
265 relying instead on anonymized, population-level, information.

266 The current study has a few limitations due to the simplifications that have been incorporated in the
267 simulations. Some of these simplifications have been motivated by the fact that the virus itself is
268 still very well under investigation and is therefore associated with multiple open questions. Further

269 refining of our knowledge of the transmission mechanisms, the viral shedding or the
270 symptomatology can affect the estimations of the testing capacities. Conversely, the described
271 systemic failure under conditions of low contact tracing and testing efficacy is driven by the well-
272 established presence of asymptomatic carriers, jointly with pre-symptomatic viral shedding. Other
273 simplifications, concerning for instance the behaviour of the agents, are motivated by the need to
274 execute a broad investigation across multiple conditions in a reasonably short time. Policy makers
275 could use our findings as a proof of concept while focussing on a single map to include more
276 realistic population-level behaviour (e.g. commuters might move long distances on a map, but
277 follow predictable paths every day). This would allow to simulate context specific effects of
278 tailored policies to aid contact tracing and testing, increasing the predictive power of the findings.

279

280 **4. Methods**

281 **4.1 Key parameters for the simulated scenarios**

282 All simulations started with a healthy population of 100,000 agents, distributed on a map depending
283 on the population density of the area analysed (Figure 2). For day 1 only, each agent had a .05%
284 probability of becoming infected, resulting in a randomly generated number of infected agents
285 (49.3 ± 7.3 across scenarios) and random geographical distributions of these agents at the beginning
286 of each simulation. These differences in the initial conditions led to diverging scenarios in terms of
287 the number of infected people, active outbreaks and the difficulty of containment, approximately
288 replicating the estimated numbers two weeks to ten days prior to establishing the lockdown
289 measures in France, in March¹⁵. Before the start of the simulation, each agent was pre-assigned to
290 one of five symptomatology categories, used only if the agent became infected. The five categories
291 included: asymptomatic (20%), symptomatic but not requiring hospitalization (65%), symptomatic
292 and requiring hospitalization (10%), symptomatic and requiring intensive care (4%) and
293 symptomatic and requiring intensive care, but will not survive (1%; Figure 4a)^{16,17,27}. Studies and
294 reports do not yet agree on the relative percentages of these categories, due to the differences
295 among regions and countries in the methods for testing and monitoring the infections in the
296 populations and requirements for hospitalization. Thus, to contain the effects that changes in the
297 distribution of symptomatic agents would have on the estimations concerning testing capacities and
298 aiding testing policies, we assumed that all agents were isolated as soon as they showed symptoms.
299 This simplification made the percentage of asymptomatic individuals in the simulation of particular
300 importance for the final estimations. The value of 20% was determined using a weighted mean
301 between two key studies reporting the percentage of asymptomatic infected subjects in the Town of

302 Vo' ¹⁷ and in the cruise ship Diamond Princess ¹⁶. The number of days required to develop
303 symptoms, if any (Figure 4b), the number of days to reach full recovery, with a bimodal distribution
304 due to the shorter time of recovery for the asymptomatic ⁹ (Figure 4c), and the time spent in
305 intensive care units (Figure 4d) were also predetermined ^{28,29}. Finally each agent was assigned to
306 one of three possible “travel cohorts”, defining the range of movement (and range of virus
307 transmission) of each agent: the entire map, a medium size sector or a small size sector (Figure 2a).
308 We simulated two conditions in terms of cohort distribution for the large, medium and small sector
309 travel behaviour, as follows: 5%-5%-90% (skewed distribution) or 30%-30%-40% (quasi uniform
310 distribution). These have been chosen to test the effects the different policies have under
311 significantly different population-wise behaviours. For all conditions, we simulated 50 different
312 scenarios. These were controlled using numbered random seeds, to allow for within-scenario
313 comparisons.

314 **4.2 Simulation of disease transmission and contact tracing**

315 To avoid overwhelmingly demanding computational resources, we did not simulate an ecological
316 behaviour for the artificial agents, as the artificial agents did not create contacts while navigating
317 the space or creating crowds (e.g. in the limited space of simulated mass transportation or building).
318 Instead, we developed a semi-static transmission mechanism that was aimed at prioritising the
319 simulation of growth in the number of infections. To this end, the transmission of the virus is
320 simulated by extracting a percentage of infected agents daily (15%, 25% or 35%, depending on the
321 simulated condition of incidence) who have not been isolated and starting from 3 days before
322 showing symptoms ⁹. Each of these agents could then propagate the infection to one contact per
323 day, randomly selected among those in the range of travel. If the randomly selected contact was
324 healthy, it was immediately infected, starting the countdown for symptom manifestation (if any).
325 Conversely, if the contacted agent had been already infected at any point in the past, the
326 propagation of the infection was null. Despite its simplicity, this mechanism replicates the dynamic
327 of growth of number of infections until reaching herd immunity, in keeping with current estimations
328 for the reproduction number R_0 of COVID-19 ⁷.

329 To simulate contact tracing, we established a day-to-day pool of all agents displaying symptoms or
330 found positive that had not been already used for contact tracing at an earlier point in time. These
331 were randomly selected, removing them from future pools and adding one positive test to the count
332 of the day. Then, the agent under examination was used to trace the agent (if any) that had been the
333 origin of its infection and the agents (if any) that it had infected at any time during the simulation.
334 Each of these contacts could be either “found” or “missed”, depending on the probability assigned

335 for contact tracing efficacy. This mechanism was meant to simulate the fact that a contact may be
336 missed because it is not tested at all (missed trace) or due to a false negative. For instance, a contact
337 tracing and testing efficacy of 20% may be due to a combined ability to trace 25% of contacts, with
338 80% test reliability.

339 The simulations allowed for perfect record keeping of the actual contacts of each infections, so we
340 implemented a system that could simulate the number of negative tests per each positive one. This
341 was performed dynamically to represent the different challenges in finding positive contacts,
342 depending on the percentage of infected in the entire population. For each positive contact, the
343 simulation automatically added a number of negative tests (that contributed to reach the daily
344 maximum capacity) equivalent to the ratio of untested infected over healthy agents for the entire
345 population, up to a maximum of 20 negative tests per each positive one.

346 Finally, to determine the optimal testing capacity for the process of contact tracing and testing, we
347 followed a simple heuristic. We initiated the simulations of the fifty scenarios with a value of 0.1
348 tests for thousand agents, across all conditions. If any of the fifty scenarios resulted in a number of
349 pre-symptomatic or asymptomatic infections above zero by day 60, the capacity was increased by
350 0.1 for that condition, restarting the process. For the conditions showing the number of pre-
351 symptomatic or asymptomatic infections could not converge to zero, irrespective of the testing
352 capacity, we increased this value so that the daily number of tests performed remained below
353 capacity for at least 50 days of simulated time, across all scenarios.

354 **4.4 Simulation of testing policies**

355 We tested several policies to aid contact tracing: two of these were used as controls and were
356 characterised by either contact tracing and testing, alone, or by the same process aided by random
357 sampling in the entire map. The remaining policies consisted in variations of weighted sampling
358 within a small or medium cell, replicating the dimension of the small and medium sector for the
359 travel behaviour. The cells were centred on the coordinates of highest concentration of new
360 infections, as recorded the day prior to the sampling. For instance, the optimal policy found for the
361 condition NY₂ consisted in sampling, within a small sector centred on the latest outbreak, with
362 weights of 60%-20%-20% for the three cohorts of travel behaviour (short, medium, long travel
363 range), whereas the optimal weights found for seIT₂ or Mid₁ were 20%-40%-40% and 80%-10%-
364 10%, respectively (Figure 3). Agents were extracted one by one as long as the capacity left unused
365 by contact tracing and testing allowed it. The agents found positive would then be isolated (i.e. they
366 could not contribute to the future transmission of the virus) and would be included in the pool of
367 agents to be traced, starting from the subsequent simulated day.

368 **4.5 Code specifics and availability**

369 The code was optimised for MATLAB r2019b (MathWorks, Natick, MA) , and it allows loading
370 any black and white dot-map of population density to test the effects of the different policies under
371 realistic conditions of population density and geographic distribution. The maps used in the
372 described case studies have been acquired using screenshots from the dot-maps provided by the
373 Cooper Center of the University of Virginia
374 (<https://demographics.virginia.edu/DotMap/index.html>), for the metropolitan area of New York,
375 and by Urban Data Visualisation by Duncan Smith, CASA UCL
376 (<https://luminocity3d.org/WorldPopDen>) for the south-east of Italy and the Midlands in UK. These
377 regions have been chosen to illustrate differences in testing policies can emerge when comparing
378 significantly different population distributions and geographical features. The maps were manually
379 converted (with the free software Gimp, v2.8) into grey scale JPEG -1000x1000 pixel, 300dpi- so to
380 have brightest part of the picture representing the highest population density. The resulting images
381 were automatically converted in the script into a matrix of probabilities that matched the grey
382 scale/distribution in the source dot-map: at the beginning of the simulation agents were randomly
383 assigned a position in the map according to these probabilities. This system allows to freely change
384 the map or the number of agents in the simulation, as the script automatically adjusts the
385 distribution to the new parameters. The entire codebase for these simulations is freely available
386 here: [LINK to be provided upon acceptance]. Ideally, we hope it can be used as a starting point for
387 more realistic simulations that would increase the ecological validity of the described estimations.

388

389 **Acknowledgments**

390 XG is supported by National Institute on Drug Abuse [grant numbers: R01DA043695,
391 R21DA049243], National Institute of Mental Health [R21MH120789], The Swartz Foundation, and
392 The Realm Foundation. NDF is supported by National Aeronautics and Space Administration
393 [ECOSTRES18-0046], National Institute Environmental health sciences [R24ES028522,
394 UH3OD023337]. DC is supported by the National Research Foundation of Korea [NRF-
395 2018R1D1A1B07043582]. The funders had no role in study design, data analysis, decision to
396 publish, or preparation of the manuscript.

397

398 **Author Contributions**

399 VGF designed the study, wrote the code, analysed the data, wrote the paper. NDeF, BSG, AP, DC
400 and XG contributed to study design, data analysis, paper writing and literature review. OP, AS, KK
401 and MO'B contributed to paper writing, literature review and graphical reporting of data.

402 **Competing Interests statement**

403 The authors report no conflict of interest.

404

405 **References**

- 406 1 World Health Organization - Coronavirus disease (COVID-19) Situation Report – 133.
407 (2020).
- 408 2 Kissler, S. M., Tedijanto, C., Goldstein, E., Grad, Y. H. & Lipsitch, M. Projecting the
409 transmission dynamics of SARS-CoV-2 through the postpandemic period. *Science*,
410 doi:10.1126/science.abb5793 (2020).
- 411 3 Hellewell, J. *et al.* Feasibility of controlling COVID-19 outbreaks by isolation of cases and
412 contacts. *The Lancet. Global health* **8**, e488-e496, doi:10.1016/S2214-109X(20)30074-7
413 (2020).
- 414 4 Swanson, K. C. *et al.* Contact tracing performance during the Ebola epidemic in Liberia,
415 2014-2015. *PLoS neglected tropical diseases* **12**, e0006762,
416 doi:10.1371/journal.pntd.0006762 (2018).
- 417 5 Kang, M. *et al.* Contact Tracing for Imported Case of Middle East Respiratory Syndrome,
418 China, 2015. *Emerging infectious diseases* **22**, 1644-1646, doi:10.3201/eid2209.152116
419 (2016).
- 420 6 He, X. *et al.* Temporal dynamics in viral shedding and transmissibility of COVID-19.
421 *Nature medicine*, doi:10.1038/s41591-020-0869-5 (2020).
- 422 7 Flavia Riccardo, M. A., Xanthi Andrianou, Antonino Bella, Martina Del Manso, Massimo
423 Fabiani, Stefania Bellino, Stefano Boros, Alberto Mateo Urdiales, Valentina Marziano,
424 Maria Cristina Rota, Antonietta Filia, Fortunato (Paolo) D'Ancona, Andrea Siddu, Ornella
425 Punzo, Filippo Trentini, Giorgio Guzzetta, Piero Poletti, Paola Stefanelli, Maria Rita
426 Castrucci, Alessandra Ciervo, Corrado Di Benedetto, Marco Tallon, Andrea Piccioli, Silvio
427 Brusaferrò, Giovanni Rezza, Stefano Merler, Patrizio Pezzotti, COVID-19 working group.
428 Epidemiological characteristics of COVID-19 cases in Italy and estimates of the
429 reproductive numbers one month into the epidemic. *medRxiv 2020.04.08.20056861*;
430 doi:<https://doi.org/10.1101/2020.04.08.20056861> (2020).

- 431 8 Liu, Y. *et al.* Viral dynamics in mild and severe cases of COVID-19. *The Lancet. Infectious*
432 *diseases*, doi:10.1016/S1473-3099(20)30232-2 (2020).
- 433 9 Wolfel, R. *et al.* Virological assessment of hospitalized patients with COVID-2019. *Nature*,
434 doi:10.1038/s41586-020-2196-x (2020).
- 435 10 Alyssa Bilinski, F. M., Joshua A Salomon. Contact tracing strategies for COVID-19
436 containment with attenuated physical distancing. *medRxiv 2020.05.05.20091280*,
437 doi:<https://doi.org/10.1101/2020.05.05.20091280> (2020).
- 438 11 Ingrid Arevalo-Rodriguez, D. B.-G., Daniel Simancas-Racines, Paula Zambrano-Achig,
439 Rosa del Campo, Agustin Ciapponi, Omar Sued, Laura Martinez-Garcia, Anne Rutjes,
440 Nicola Low, Jose A Perez-Molina, Javier Zamora. False-Negative Results of Initial RT-PCR
441 Assays for COVID-19: A Systematic Review. *medRxiv 2020.04.16.20066787*,
442 doi:<https://doi.org/10.1101/2020.04.16.20066787> (2020).
- 443 12 Sethuraman, N., Jeremiah, S. S. & Ryo, A. Interpreting Diagnostic Tests for SARS-CoV-2.
444 *JAMA*, doi:10.1001/jama.2020.8259 (2020).
- 445 13 Langton, C. G. *Artificial life : an overview*. (MIT Press, 1995).
- 446 14 (Sackler NAS Colloquium) *Adaptive Agents, Intelligence, and Emergent Human*
447 *Organization: Capturing Complexity through Agent-Based Modeling*. (The National
448 Academies Press, 2002).
- 449 15 Salje, H. *et al.* Estimating the burden of SARS-CoV-2 in France. *Science*,
450 doi:10.1126/science.abc3517 (2020).
- 451 16 Mizumoto, K., Kagaya, K., Zarebski, A. & Chowell, G. Estimating the asymptomatic
452 proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess
453 cruise ship, Yokohama, Japan, 2020. *Euro surveillance : bulletin Europeen sur les maladies*
454 *transmissibles = European communicable disease bulletin* **25**, doi:10.2807/1560-
455 7917.ES.2020.25.10.2000180 (2020).
- 456 17 Enrico Lavezzo, E. F., Constanze Ciavarella, Gina Cuomo-Dannenburg, Luisa Barzon,
457 Claudia Del Vecchio, Lucia Rossi, Riccardo Manganelli, Arianna Loregian, Nicolò Navarin,
458 Davide Abate, Manuela Sciro, Stefano Merigliano, Ettore Decanale, Maria Cristina
459 Vanuzzo, Francesca Saluzzo, Francesco Onelia, Monia Pacenti, Saverio Parisi, Giovanni
460 Carretta, Daniele Donato, Luciano Flor, Silvia Cocchio, Giulia Masi, Alessandro Sperduti,
461 Lorenzo Cattarino, Renato Salvador, Katy A.M. Gaythorpe, Imperial College London
462 COVID-19 Response Team, Alessandra R Brazzale, Stefano Toppo, Marta Trevisan,
463 Vincenzo Baldo, Christl A. Donnelly, Neil M. Ferguson, Ilaria Dorigatti, Andrea Crisanti.
464 Suppression of COVID-19 outbreak in the municipality of Vo, Italy. *medRxiv*
465 *2020.04.17.20053157*, doi:<https://doi.org/10.1101/2020.04.17.20053157> (2020).

- 466 18 Freedman, H. I. *Deterministic mathematical models in population ecology*. (M. Dekker,
467 1980).
- 468 19 Leung, K., Wu, J. T., Liu, D. & Leung, G. M. First-wave COVID-19 transmissibility and
469 severity in China outside Hubei after control measures, and second-wave scenario planning:
470 a modelling impact assessment. *Lancet* **395**, 1382-1393, doi:10.1016/S0140-
471 6736(20)30746-7 (2020).
- 472 20 Ferguson NM, L. D., Nedjati-Gilani G, Imai N, Ainslie K, Baguelin M, Bhatia S, Boonyasiri
473 A, Cucunubá Z, Cuomo-Dannenburg G, Dighe A. Impact of non-pharmaceutical
474 interventions (NPIs) to reduce COVID-19 mortality and healthcare demand. Imperial
475 College COVID-19 Response Team, London, March, 16. (2020).
- 476 21 Cereda D, T. M., Rovida F, Demicheli V, Ajelli M, Poletti P, Trentini F, Guzzetta G,
477 Marziano V, Barone A, Magoni M, Deandrea S, Diurno G, Lombardo M, Faccini M, Pan A,
478 Bruno R, Pariani E, Grasselli G, Piatti A, Gramegna M, Baldanti F, Melegaro A, Merler S.
479 The early phase of the COVID-19 outbreak in Lombardy, Italy. *arXiv* (2020).
- 480 22 Jacob B Aguilar, J. S. F., Lauren M. Westafer, Juan B. Gutierrez. Investigating the Impact of
481 Asymptomatic Carriers on COVID-19 Transmission. *medRxiv* 2020.03.18.20037994,
482 doi:<https://doi.org/10.1101/2020.03.18.20037994> (2020).
- 483 23 Bohmer, M. M. *et al.* Investigation of a COVID-19 outbreak in Germany resulting from a
484 single travel-associated primary case: a case series. *The Lancet. Infectious diseases*,
485 doi:10.1016/S1473-3099(20)30314-5 (2020).
- 486 24 Kampf, G., Todt, D., Pfaender, S. & Steinmann, E. Persistence of coronaviruses on
487 inanimate surfaces and their inactivation with biocidal agents. *The Journal of hospital
488 infection* **104**, 246-251, doi:10.1016/j.jhin.2020.01.022 (2020).
- 489 25 van Doremalen, N. *et al.* Aerosol and Surface Stability of SARS-CoV-2 as Compared with
490 SARS-CoV-1. *The New England journal of medicine* **382**, 1564-1567,
491 doi:10.1056/NEJMc2004973 (2020).
- 492 26 Andrew C Miller, N. J. F., Joseph A Lewnard, Nicholas P Jewell, Carlos Guestrin, Emily B
493 Fox. Mobility trends provide a leading indicator of changes in SARS-CoV-2 transmission.
494 *medRxiv* 2020.05.07.20094441, doi:<https://doi.org/10.1101/2020.05.07.20094441> (2020).
- 495 27 Russell, T. W. *et al.* Estimating the infection and case fatality ratio for coronavirus disease
496 (COVID-19) using age-adjusted data from the outbreak on the Diamond Princess cruise
497 ship, February 2020. *Euro surveillance : bulletin Europeen sur les maladies transmissibles
498 = European communicable disease bulletin* **25**, doi:10.2807/1560-
499 7917.ES.2020.25.12.2000256 (2020).

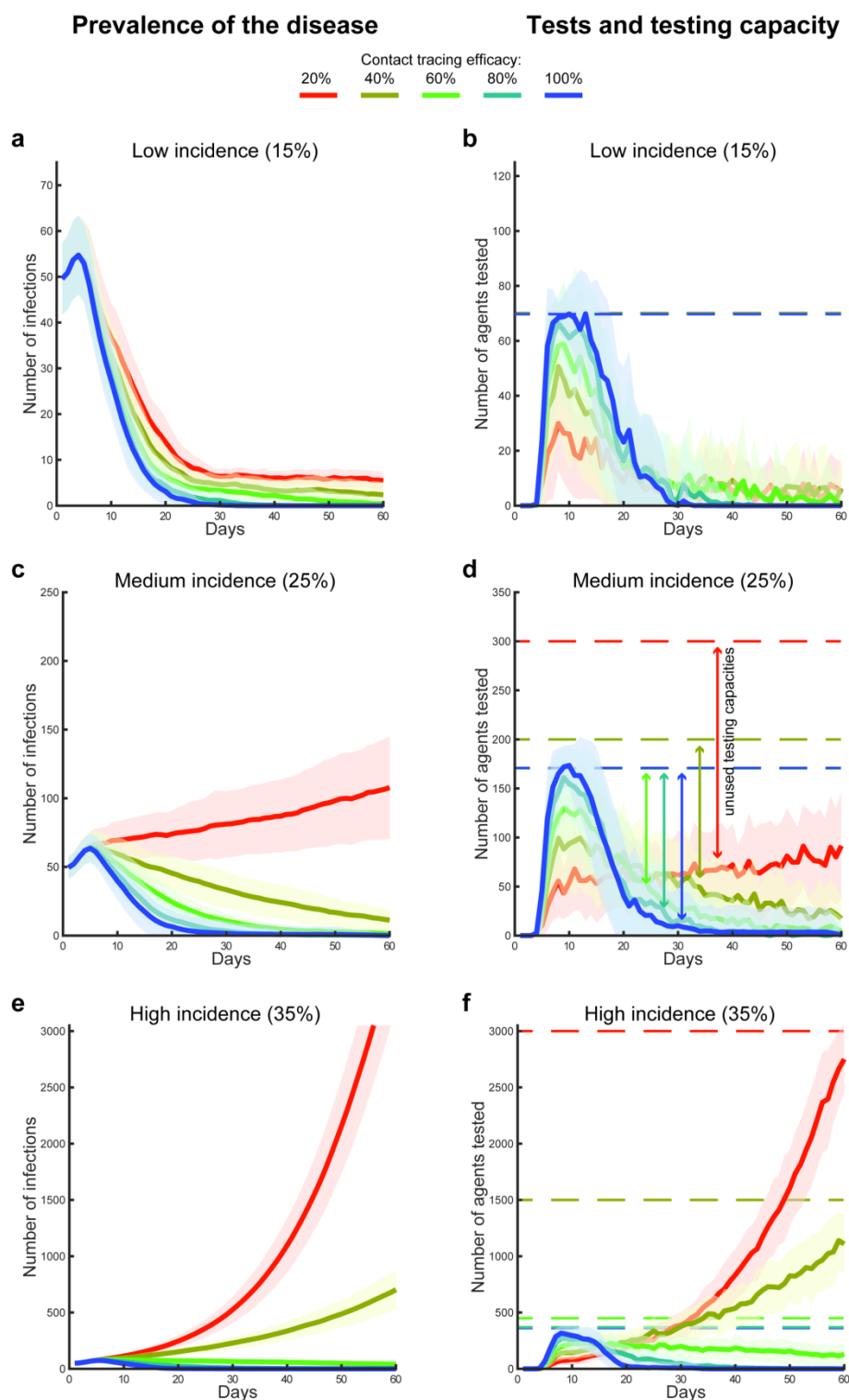
- 500 28 Lauer, S. A. *et al.* The Incubation Period of Coronavirus Disease 2019 (COVID-19) From
501 Publicly Reported Confirmed Cases: Estimation and Application. *Annals of internal*
502 *medicine*, doi:10.7326/M20-0504 (2020).
- 503 29 Intensive Care National Audit & Research Centre (ICNARC). ICNARC report on COVID-
504 19 in critical care (17 April 2020). Available from: [https://www.icnarc.org/Our-](https://www.icnarc.org/Our-Audit/Audits/Cmp/Reports)
505 [Audit/Audits/Cmp/Reports.](https://www.icnarc.org/Our-Audit/Audits/Cmp/Reports), (2020).

506

Table 1. Simulated testing capacities expressing the availability of tests per thousand agents. Background colour of an individual cell indicates reported conditions of disease containment: a white background indicates suppression of the transmission ($R_{\text{effective}} < 0$), a light grey background indicates containment but not suppression ($R_{\text{effective}} \approx 1$), and dark grey background indicates exponential growth ($R_{\text{effective}} > 1$).

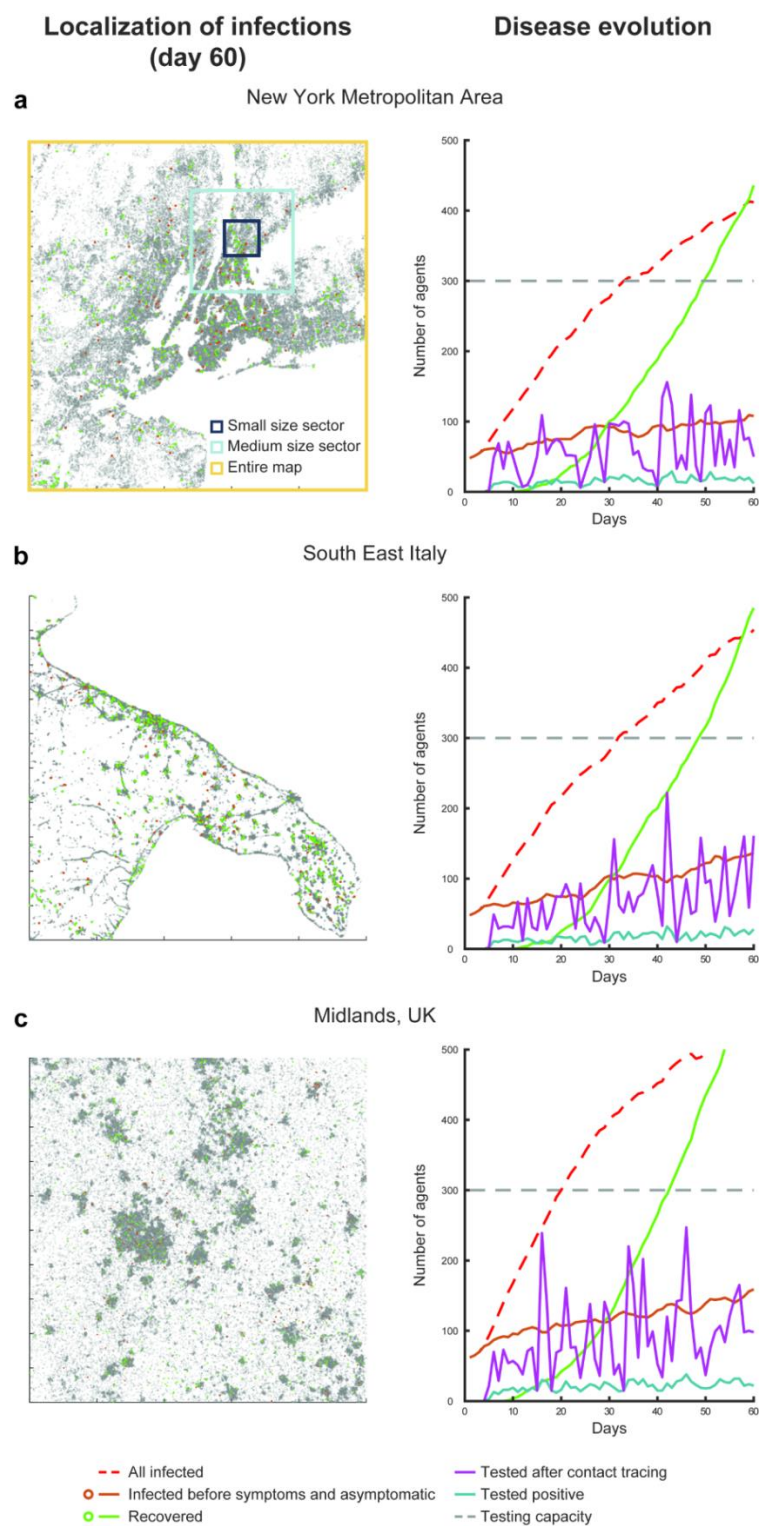
	Growth rate: 15%	Growth rate: 25%	Growth rate: 35%
100% contacts traced	0.7	1.7	3.6
80% contacts traced	0.7	1.7	3.6
60% contacts traced	0.7	1.7	4.5
40% contacts traced	0.7	2	15
20% contacts traced	0.7	3	30

507



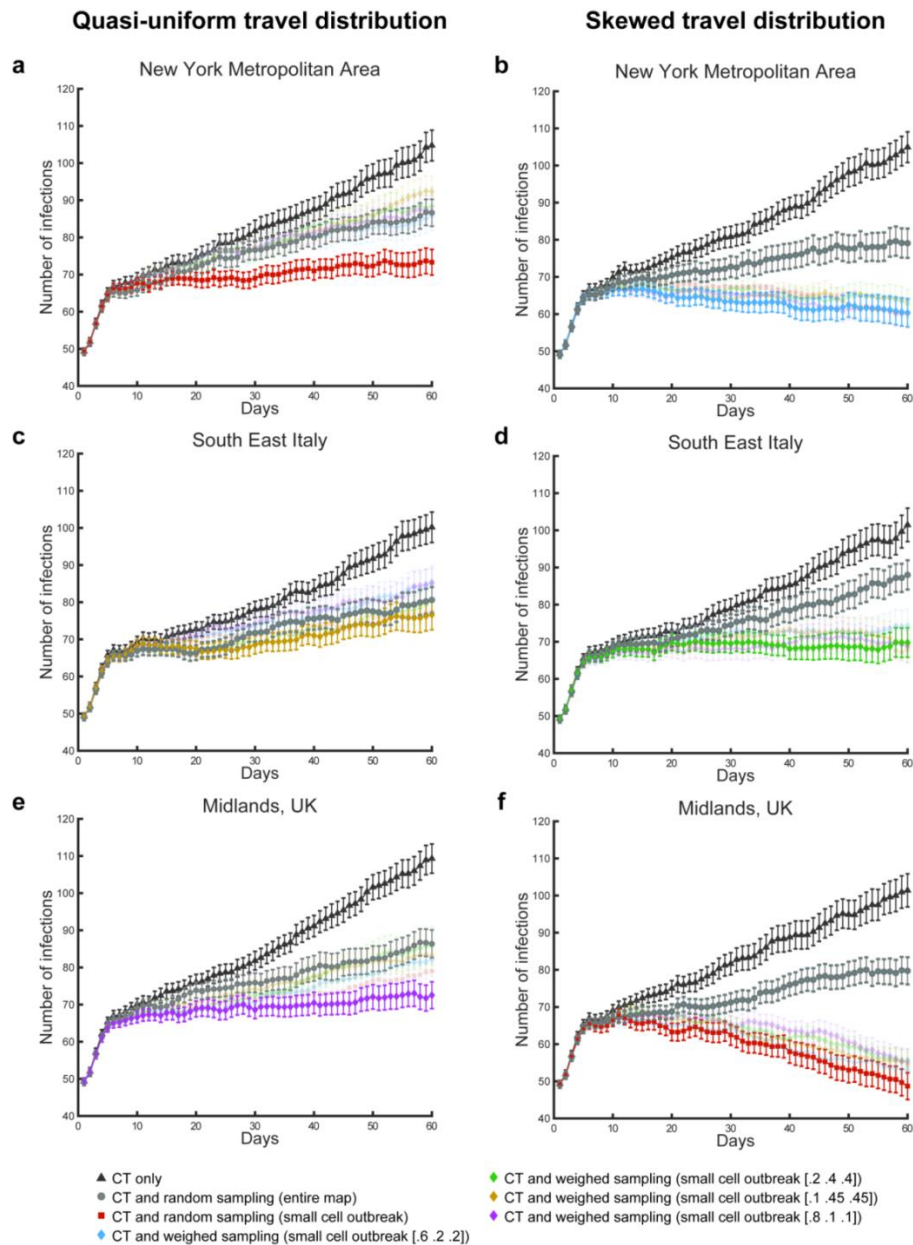
508

509 **Figure 1. Disease prevalence, agents tested and capacity.** Error bands (mean and standard
 510 deviation) represent the prevalence of COVID-19 in the population over 60 days of simulated time
 511 (a, c, e), and the associated number of daily tested agents in relation with the respective testing
 512 capacities (solid and dotted lines respectively in b, d, f). The 3x5 design was used to simulate three
 513 conditions of simulated disease incidence, e.g. due to different mitigation strategies in place, which
 514 regulated the growth in the number of infections (a-b: 15%, c-d: 25% and e-f: 35%), and five
 515 conditions of contact tracing and testing efficacy (100%, 80%, 60%, 40% and 20%).



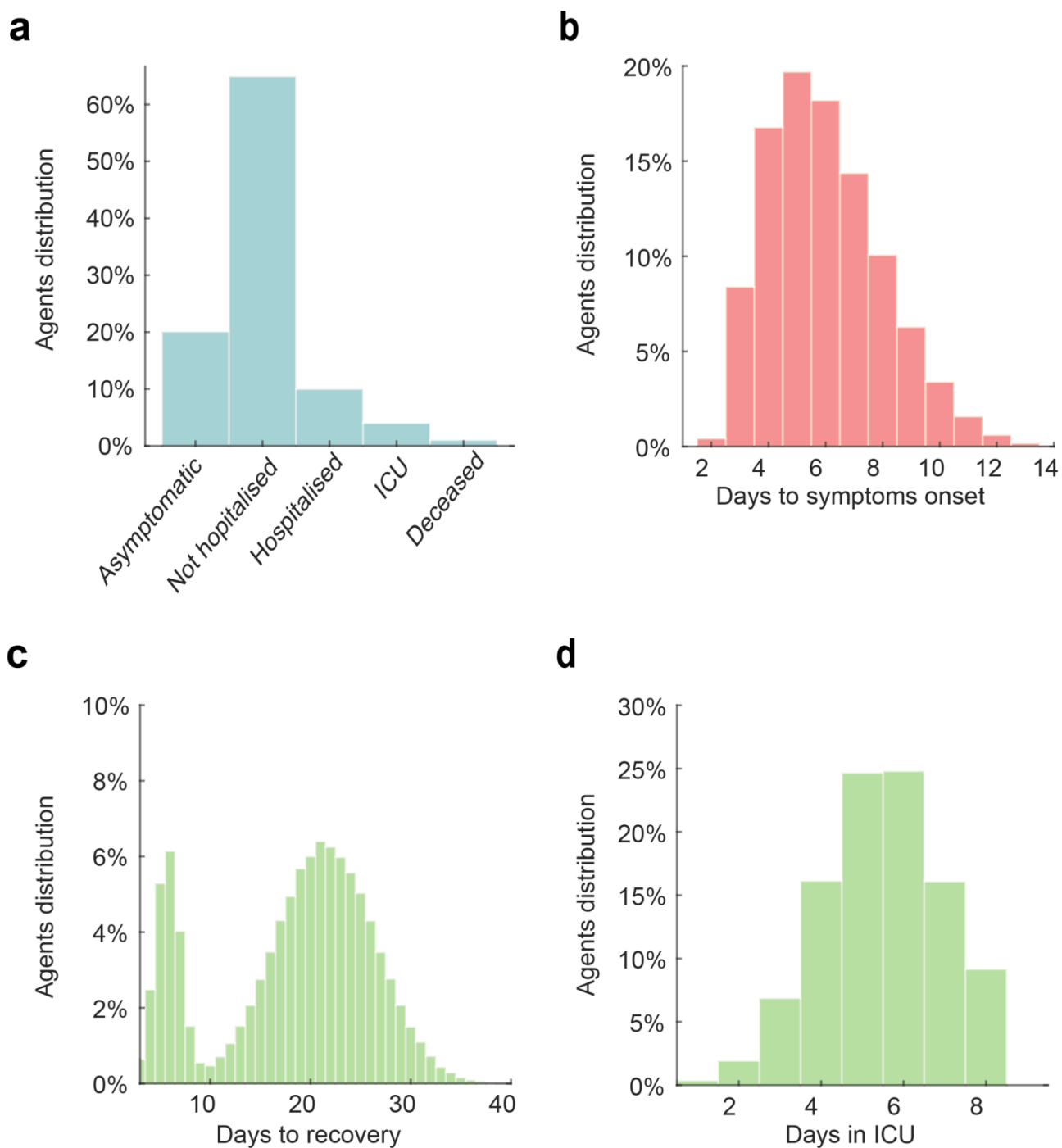
516

517 **Figure 2. Simulated evolution of the virus transmission over three regions.** Illustration of three
 518 different simulations for the scenario 1 (random seed 1) for the maps of New York metropolitan
 519 area (a), southeast Italy (b) and the Midlands in UK (c). All simulations display the (failed)
 520 containment of the disease transmission relying only on contact tracing, testing and isolation, under
 521 the conditions of medium incidence (25% daily increase in the number of infections), 20% contact
 522 tracing and testing efficacy and a distribution of travel cohorts of 40%, 30% and 30% for the short,
 523 medium and long travel range (respectively illustrated as black, blue and yellow squares in panel a).



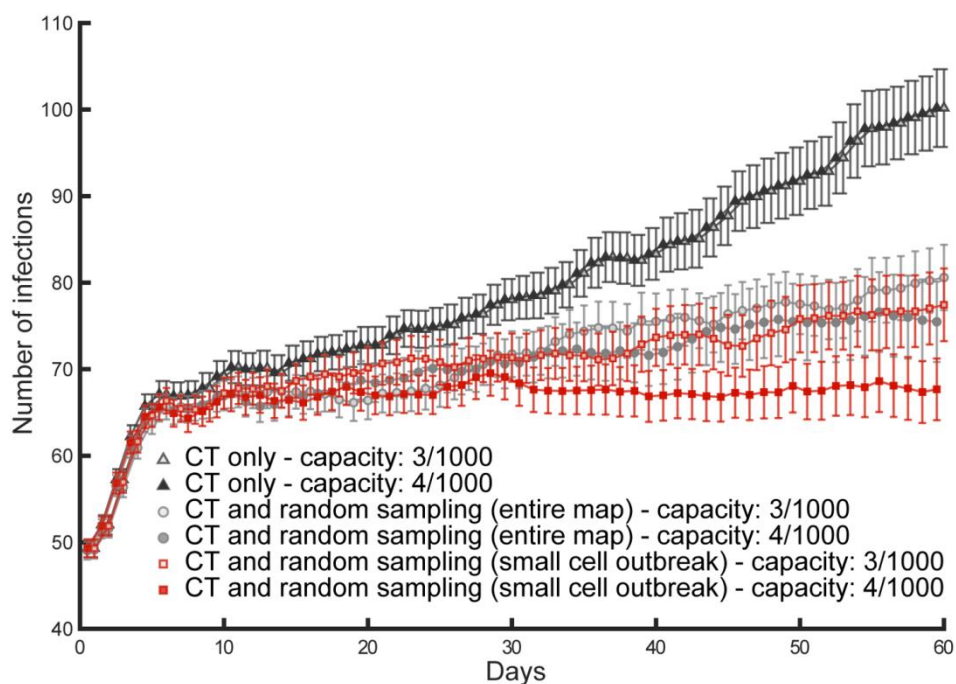
524

525 **Figure 3. Effects of test and isolation policies on the virus effective reproduction.** Mean number
 526 of infections per day, associated with different conditions and testing policies. Legends and error
 527 bars (standard errors) are depicted for the policies of contact tracing and testing (CT), alone (black
 528 triangles), contact tracing and testing jointly with random sampling across the entire map (grey
 529 circles) and the combination of contact tracing and testing jointly with the best performing sampling
 530 policies. Note that these optimal policies change depending on the simulated conditions of
 531 geographical distribution and travel behaviour of the population. Under all conditions, the optimal
 532 sampling policy to aid contact tracing focuses on small cells (equivalent to a small sector for the
 533 travel behaviour) centred on the coordinates of the most severe outbreak recorded in the previous
 534 day of simulated time. For two conditions, the optimal sampling is random within this cell (a, f).
 535 For the remaining conditions, the sampling is weighted: 60%-20%-20% (b), 10%-45%-45% (c),
 536 20%-40%-40% (d), 80%-10%-10% (e), for short, medium and long distance travel cohort.



537

538 **Figure 4. Simulation settings.** The histograms represent the distribution of symptoms (a), days
539 required for the symptoms onset (b), days required for recovery after symptoms onset (c) and days
540 required in intensive care units (d). Note that the bimodal distribution of the days to recovery is due
541 to the presence of asymptomatic subjects who are characterised by a shorter recovery time (marking
542 the end of viral shedding). The days spent in the intensive care units are considered as part of the
543 time required to recovery, when recovery is possible.



545

546 **Supplementary figure 1. Increased testing capacity and testing policies.** These simulations
547 illustrate the effects of an increase of testing capacity from 3 to 4 tests per thousand agents, jointly
548 with 20% contact tracing and testing (CT) efficacy. The contact tracing and testing process, when
549 considered alone (filled triangles for high capacity and empty triangles for high capacity), does not
550 exhaust the initial testing capacity due to low efficacy, so that an increase in capacity is ineffective
551 as it simply increases the number of unused tests per day. Instead, improved containment of the
552 disease transmission is found both for contact tracing and testing jointly with random sampling over
553 the entire population (filled circles for high capacity and empty squares for low capacity), as well as
554 for contact tracing and testing jointly with random sampling over a small sector centred on the most
555 recent outbreak (filled squares for high capacity and empty squares for low capacity). The latter
556 mixed policy succeeds in keeping the number of daily infections constant ($R_{\text{effective}} \approx 1$), once the
557 capacity is increased.