

26 Introduction

27 It has been widely debated which policies, if any, could have mitigated the health impacts of the
28 initial stages of the COVID-19 pandemic in late 2019 and early 2020 as community transmission
29 became established and widespread. Early studies compared non-pharmaceutical interventions
30 (NPIs) such as mobility restrictions^{1,2}, school closures^{3,4}, voluntary home quarantine⁵ and testing
31 policies⁶, and optimized NPI parameters like testing frequency⁷, quarantine length⁸, testing
32 modality⁹, test pooling¹⁰ and intervention timing and ordering¹¹. While such NPIs undoubtedly
33 slowed the early spread of COVID-19¹² and previous outbreaks^{13,14}, there has been little
34 investigation of whether a separate strategy focused on earlier detection of COVID-19 would
35 have enabled more successful mitigation. In theory, earlier detection enables a response when the
36 outbreak is smaller: thus, resource-intensive mitigation strategies like test-trace-isolate become
37 less costly, and the earlier interventions are applied, the larger the number of infections and
38 deaths that can be delayed until healthcare capacity is increased¹⁵. However, the relevant
39 question is not whether early detection helps, but quantitatively how much of a difference it
40 would make. This question is especially urgent given current international and national policy
41 proposals to invest billions of dollars in such systems^{16,17}.

42 Researchers and policymakers have proposed immediate investments in systems to continuously
43 monitor for novel pathogens in (i) patients with infectious symptoms in hospitals and clinics^{18,19},
44 (ii) community wastewater treatment plants^{20,21}, and (iii) airplane sewage or bridge air on
45 international flights²²⁻²⁴, as well as other sites²⁵⁻²⁹. These three sites have attracted interest
46 because they have been frequent testing sites in COVID-19: hospitals since the pandemic's
47 beginning³⁰, and wastewater (including wastewater at treatment plants²⁰, within the sewershed³¹,
48 and locally near individual buildings³²) and air travel more recently^{33,34} because hospital cases
49 can lag community cases³⁵. COVID-19 also spurred methodological innovation and
50 characterization of sampling from these sites³⁶, particularly wastewater³⁷⁻³⁹. Detecting novel
51 pandemics at these sites has occasionally been piloted^{21,40} but has not been implemented at scale,
52 in part because it is unclear if these proposed systems sufficiently expedite detection of
53 outbreaks. The systems under consideration would use multiplex testing for conserved nucleic
54 acid sequences of known pathogen families, exploiting the fact that many past emerging diseases
55 belonged to such families, including SARS-CoV-2 (2019), Ebola (2013), MERS-CoV (2012),
56 and pandemic flu (2009). Proposed technologies include multiplex PCR⁴¹⁻⁴⁴, CRISPR-based
57 multiplex diagnostics⁴⁵, and metagenomic sequencing⁴⁶, possibly implemented with pooling¹⁰.

58 To determine whether early detection of novel pathogens at these sites could be effective in
59 changing the course of a pandemic, we first examined whether COVID-19 could have been
60 detected earlier in Wuhan if systems had been in place in advance to monitor hospitals,
61 wastewater or air travel. To do this, we developed, empirically validated, and mathematically
62 characterized a simulation-based model that predicts the number of cases at the time of detection
63 given a detection system and a set of outbreak epidemiological parameters. We then used this
64 model and COVID-19 epidemiological parameters⁴⁷ to estimate how early COVID-19 would
65 initially have been detected in Wuhan by the three early detection systems, and compare this to
66 the actual date of COVID-19 detection. Finally, we use our model to estimate detection times of
67 infectious agents with different epidemiological properties, such as mpox and polio in recent
68 outbreaks^{48,49}, to inform pathogen-agnostic surveillance for future pandemics.

69 Results

70 Model to estimate earliness of detection

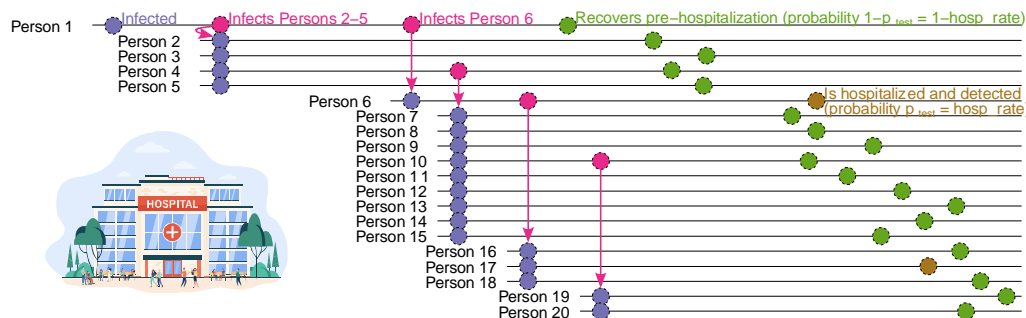
71 Previous research¹⁵ and our analysis (Supplementary text, figs. S1—5 and table S1) suggest that
72 earlier COVID-19 lockdowns could have delayed cases and deaths. Thus, it is critical to
73 understand which early detection systems, if any, could have effectively enabled earlier
74 response. To do this, we built a model that simulates outbreak spread and earliness of detection
75 for a given outbreak and detection system (Materials and methods, Supplementary materials).
76 This builds upon branching process models that have previously been used to model the spread
77 of COVID-19^{50,51} and other infectious diseases⁵². A traditional branching process model starts
78 from an index case and iteratively simulates each new generation of infections. Our model
79 follows this pattern, but with each new infection we also simulate whether the infected person is
80 detected by the detection system with some probability (Fig. 1A), and the simulation stops when
81 the number of detected individuals equals the detection threshold and the detection delay has
82 passed. Thus, each detection system is characterized by these three parameters of detection
83 probability, threshold, and delay (table S2). For example, in hospital monitoring, an infected
84 individual's detection probability is the probability they are sick enough to enter the hospital,
85 which is the hospitalization rate (assuming testing has a negligible false negative rate). In
86 systems that test individuals (hospital and air travel individual monitoring), the threshold is
87 measured in an absolute number of cases. In systems that test wastewater (wastewater
88 monitoring), the threshold is measured in terms of outbreak prevalence because wastewater
89 monitoring can only sample a small percentage of sewage flows, depending on the sampling
90 capacity⁵³; thus, a higher number of cases is required to trigger detection in a bigger community
91 (Materials and methods). We gathered literature estimates of detection system and outbreak
92 parameters (tables S2 and S3) and validated wastewater monitoring sensitivity in independent
93 data (fig. S6 and Materials and methods, Supplementary materials). We then empirically
94 validated the model by testing its ability to predict the detection times for the first COVID-19
95 outbreaks in 50 US states in 2020. We gathered the dates of the first COVID-19 case reported by
96 the public health department of each US state (table S4) as well as literature estimates of true
97 (tested and untested) statewide COVID-19 case counts in early 2020⁵⁴. Using our model, we
98 were able to predict the number of weeks until travel-based detection in each US state to within a
99 mean absolute error of 0.97 weeks (fig. S7 and S8). To check the robustness of our results, we
100 implemented a second, more complex model with varying reproduction number using a Monte
101 Carlo simulation-based package (EpiNow2⁵⁵). A list of model assumptions can be found in table
102 S5.

103 Early detection's impact on COVID-19 detection in Wuhan

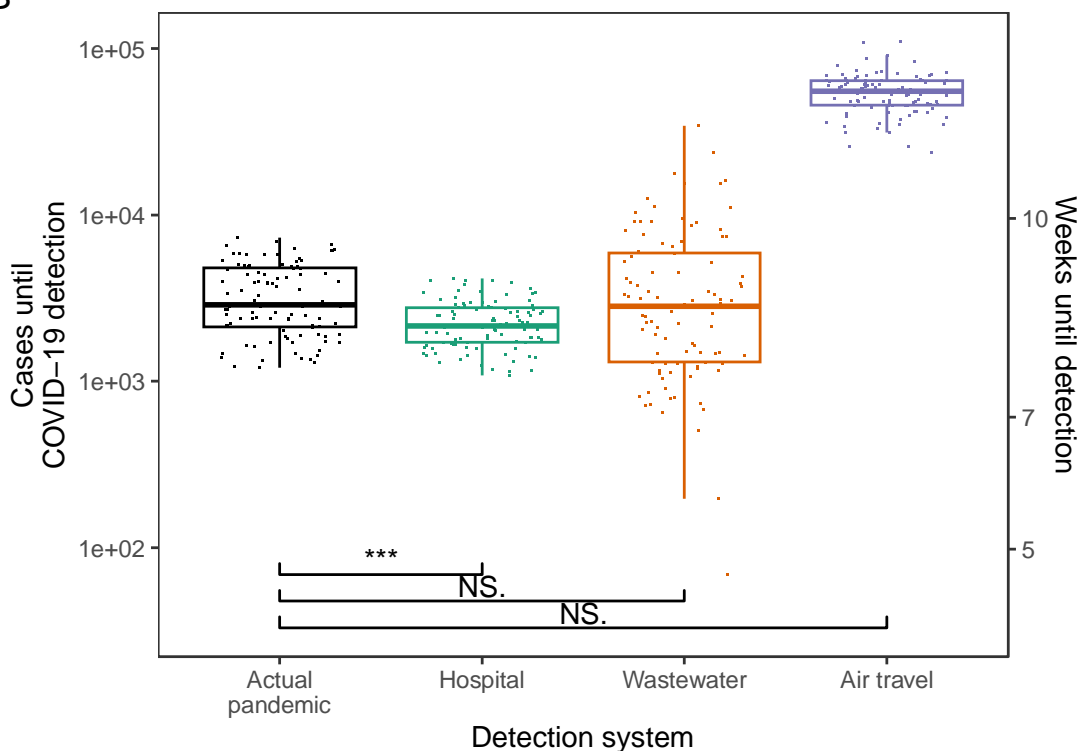
104 Next we use our model to examine the detection systems' ability to detect the first major
105 COVID-19 outbreak in Wuhan (Fig. 1B and table S2). To estimate cases at detection in the
106 actual pandemic, we used literature estimates of total (tested and untested) COVID-19 case
107 counts in Wuhan in late 2019 and early 2020⁵⁶. Our model shows that, on average, hospital
108 monitoring could have detected COVID-19 after 2,292 cases (standard error: 76 cases). In
109 reality, the pandemic was identified after 3,413 cases on average (standard error: 161 cases).
110 Thus, hospital monitoring would have caught the outbreak 1,121 cases earlier (approximately
111 0.43 weeks earlier), a statistically significant difference with $p = 1.9e-09$ and $t = -6.3$ ($df = 141$)
112 in a one-sided Welch two-sample t-test. Wastewater monitoring would have lagged detection in
113 the actual pandemic; it caught the outbreak after 4,575 cases (standard error: 523 cases), or 1,162
114 cases later, on average ($p = 0.018$; $t = 2.1$; $df = 118$). We tested this wastewater prediction

115 empirically by calculating the cases until COVID-19 wastewater detection in Massachusetts in
116 early 2020, using literature-estimated Massachusetts COVID-19 cases⁵⁴ and Massachusetts
117 wastewater SARS-CoV-2 PCR data⁵⁷; our model prediction was consistent with this analysis
118 (fig. S9). Because we model wastewater monitoring to detect later in larger communities
119 (Materials and methods, Supplementary materials), the Wuhan result is in part due to Wuhan's
120 650,000-person catchments. Wastewater monitoring would lead status quo detection of COVID-
121 19 in catchments smaller than 480,000 people, well above the global mean catchment size of
122 25,000 people⁵⁸. Air travel monitoring did not provide any acceleration of detection because of
123 the low probability of simultaneously traveling and being sick.

A



B



124

125 **Fig. 1. Comparison of COVID-19 detection times in the actual pandemic versus with**
 126 **proposed early detection systems. (A)** Schematic of first 20 infections in a simulated run of the
 127 detection model. In this run, Person 1 seeds an outbreak in a community covered by a hospital
 128 detection system. Each person infects a number of individuals determined by a draw from a

129 negative binomial distribution. Each person is then detected by the detection system with
130 probability p_{test} (gold) or goes undetected (olive); in the hospital system, p_{test} equals the
131 hospitalization rate. **(B)** Estimated cases until COVID-19 detection in the actual pandemic versus
132 model-simulated cases until detection for proposed detection systems (center line, median; box
133 limits, upper and lower quartiles; whiskers, point closest to 1.5x interquartile range). Estimates
134 for the actual pandemic are drawn from⁵⁶. Points for proposed detection systems are simulated
135 case counts from the model (actual pandemic (black), hospital (teal), wastewater (orange) and air
136 travel (purple)) assuming a Wuhan-sized catchment (650,000 people). Three, two, and one
137 asterisk(s) signify that the cases upon detection for the detection system are statistically
138 significantly lower than those in the actual pandemic at the 0.001, 0.01, and 0.05 levels,
139 respectively, in one-sided t-tests. NS. signifies not statistically significantly lower at $p=0.05$.
140 Equivalent weeks until detection are shown on the right y-axis.

141 Early detection's impact for other diseases: mathematical analysis and simulation

142 To make our model easily usable for pathogenic outbreaks beyond COVID-19, we derived a
143 compact formula that approximates the model's simulations. We observed that, without
144 accounting for the delay of g generations between the threshold case's infection and detection,
145 the number of cases until detection, C , is a random variable that follows a negative binomial
146 distribution by definition: each infected case is a Bernoulli trial, "success" in that trial occurs
147 when that case enters the detection system (with a probability we name p_{test}), and we count the
148 number of cases until the number of successes equals the detection threshold d . After accounting
149 for g and the basic reproduction number R_0 , we derived a formula approximating the mean of C
150 when the outbreak starts in a community covered by the detection system (see Supplementary
151 Text for full derivation):

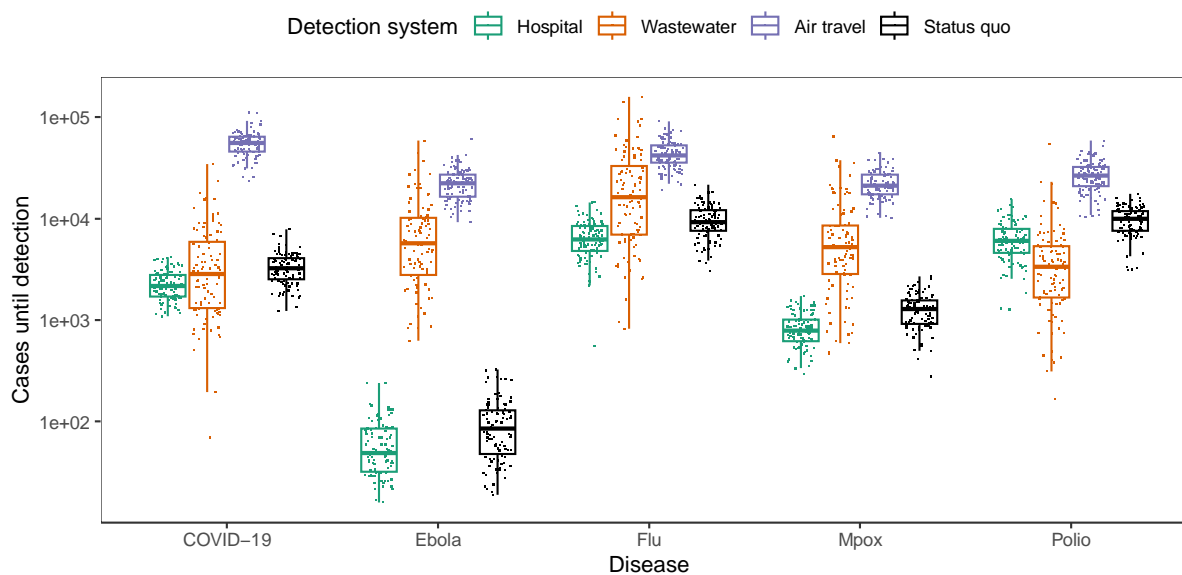
$$E[C] \approx \frac{d \times R_0^g}{p_{test}} \quad (1)$$

152 We confirmed our formula approximates the simulation model closely by comparing the
153 detection times predicted by both for all the detection systems for multiple diseases (fig. S10).
154 Thus, the formula allows us to interpret the model and the quantitative relationships between
155 detection times and various variables: the formula shows that the number of cases until detection
156 increases linearly with the detection threshold, increases polynomially with R_0 and exponentially
157 with the delay g as R_0^g , and decreases as the fraction of cases being tested increases. This
158 formula also makes the model easily usable for detection systems beyond the ones studied here.

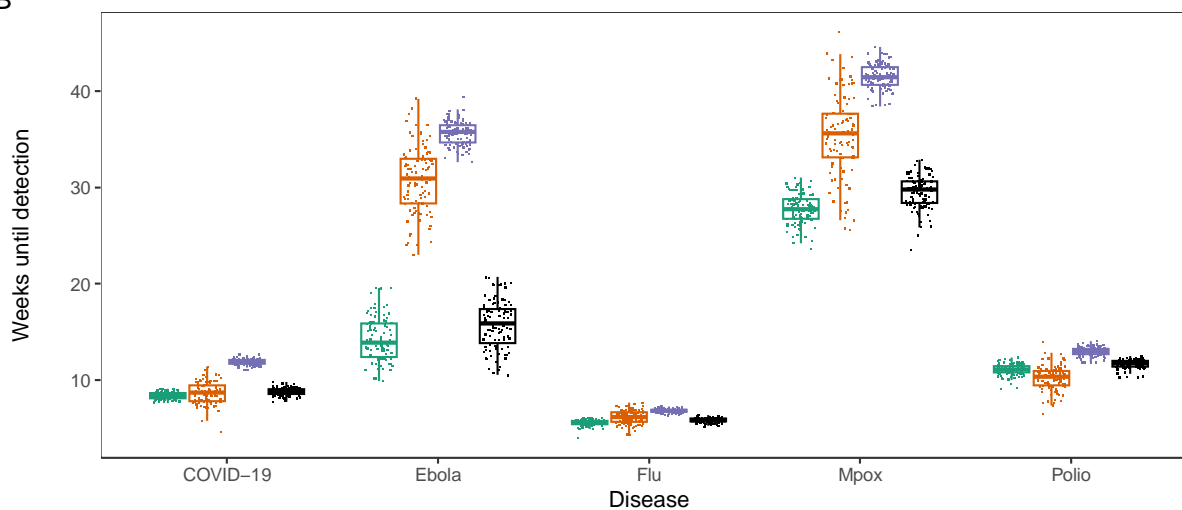
159 We applied our model to several outbreaks of recent interest—including COVID-19, mpox
160 (2022), polio (2013-2014), Ebola (2013-2016) and flu (2009 pandemic)—and found that the
161 detection systems vary in their success depending on the epidemiological parameters of the agent
162 (Figs. 2, S11 and S12, and table S3). For example, in our model hospital monitoring tends to
163 outperform wastewater monitoring when the hospitalization rate is high, as in the case of Ebola,
164 but tends to underperform for diseases like polio, in which the hospitalization rate is low and
165 when there is high asymptomatic spread in the delay from detection to hospitalization. This is
166 consistent with Equation (1), as well as previous observations that Ebola was first detected in
167 hospitals⁵⁹ and that wastewater monitoring has been more effective than hospital monitoring for
168 detecting polio⁶⁰. We also modeled the status quo detection times for these outbreaks: the
169 number of cases until these outbreaks were detected in the status quo, without the proposed

170 detection systems in place. We found that early detection systems can catch outbreaks when they
171 are up to 52% smaller (wastewater for polio) or 110 weeks earlier (hospital for HIV/AIDS) (figs.
172 S13, S14, S15 and S16). Similar results hold for the more complex model: the relative median
173 detection times of the three systems remain the same 97% of the time across the five main
174 diseases (29/30 pairwise comparisons) (fig. S17).

A



B

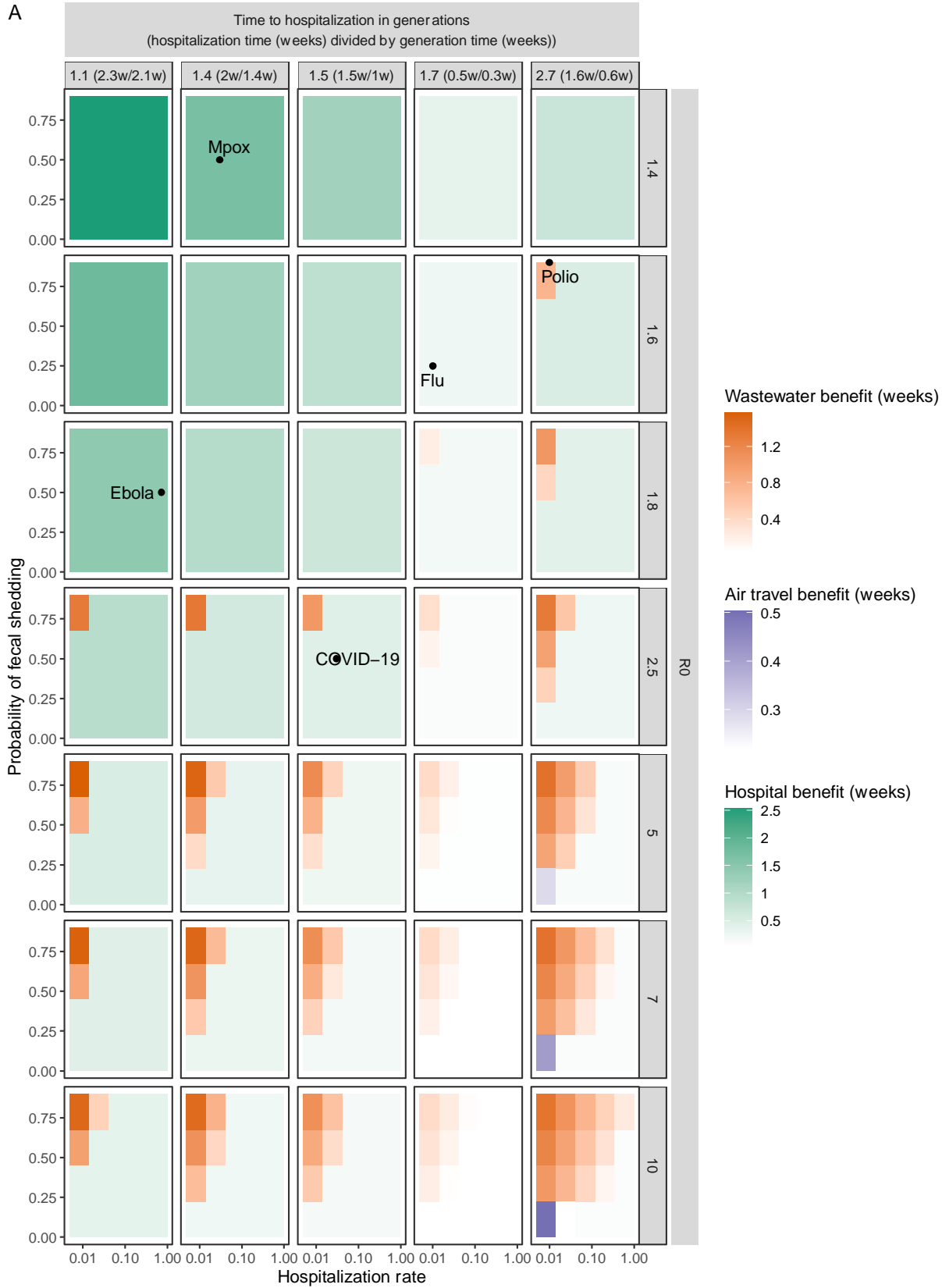


C

Diseases	COVID-19	Ebola (2013–2016)	Flu (2009 pandemic)	Mpox (2022)	Polio (2013–2014)
Hospitalization rate	0.03	0.72	0.01	0.03	0.01
R0	2.50	1.80	1.60	1.40	1.60
Serial interval (weeks)	1.00	2.10	0.30	1.40	0.60
Time to hospitalization (weeks)	1.50	2.30	0.50	2.00	1.60
Probability of fecal shedding	0.50	0.50	0.25	0.50	0.90
Dispersion	0.70	0.10	0.10	0.10	0.10

176 **Fig. 2. Comparison of detection systems for different infectious diseases.** (A) Earliness of
177 detection for detection systems in cases across infectious diseases (hospital (teal), wastewater
178 (orange), air travel (purple) and status quo (black)) in a 650,000-person catchment (center line,
179 median; box limits, upper and lower quartiles; whiskers, point closest to 1.5x interquartile
180 range). (B) Earliness of detection for detection systems in weeks across infectious diseases in a
181 650,000-person catchment. (C) Epidemiological parameters of the studied diseases.

182 Because future infectious diseases are likely to have different epidemiological parameters, we
183 generalized the previous analysis and calculated detection times for many possible diseases
184 spanning the epidemiological parameter space (Figs. 3 and S18). As expected, hospital
185 monitoring is the best system for diseases with higher hospitalization rates and lower times to
186 hospitalization. For diseases with higher R0s and times to hospitalization, wastewater monitoring
187 emerges as the best system more often, because hospital monitoring has a longer detection delay
188 (mainly the time from infection to hospitalization) than wastewater (mainly the time from
189 infection to fecal shedding), during which cases grow exponentially with R0. However, this
190 holds mainly for diseases with high probability of fecal shedding and low hospitalization rate.
191 Air travel monitoring, which did not perform well in the previous modeled diseases (Figs. 1 and
192 2), actually performed best for a few diseases for which fecal shedding is low (disadvantaging
193 wastewater monitoring) and the time to hospitalization and R0 are too large (disadvantaging
194 hospital monitoring).



196 **Fig. 3. Comparison of detection systems across the space of possible diseases of varying**
197 **epidemiological parameters. (A)** Average weeks gained over status quo detection by the
198 proposed detection systems across the epidemiological space of possible diseases. Within each
199 panel, each uniformly colored cell corresponds to a specific disease with the hospitalization rate
200 and probability of fecal shedding indicated on the x- and y-axes, as well as the R_0 and time to
201 hospitalization (generations) indicated by the panel row and column. The cell has a hue
202 corresponding to the detection system that detects the disease the earliest (hospital (teal),
203 wastewater (orange) and air travel (purple)) and an intensity corresponding to the number of
204 weeks gained by the earliest system over status quo detection. Times are calculated by the
205 derived mathematical approximation in a 650,000-person catchment.

206 Discussion

207 Our results show that the benefits of early detection systems vary from marginal (0.4 weeks
208 earlier for COVID-19) to significant (110 weeks earlier for HIV/AIDS) (Figs. 1B, 2, and S16).
209 Our detection time model (Fig. 1A) can be used for many diseases and detection systems,
210 including other systems beyond this study^{25,26}, by varying the fraction of the infected population
211 being tested in each system. Some further points are worth emphasizing. First, early detection
212 only aids mitigation if it leads to a coordinated early response. Many factors beyond detection
213 affect the pace of response, including the economic and political feasibility of lockdowns, the
214 availability of medicines and personal protective equipment, and whether there are pre-
215 determined policies to be implemented upon detection. Second, when deciding to invest in these
216 systems, one must consider factors such as cost-effectiveness and whether the system provides
217 evidence of disease severity. Although wastewater monitoring gives earlier detection than
218 hospital monitoring in multiple diseases (Fig. 3A), it does not discriminate between mild and
219 severe disease (although sequencing could detect lineages known to cause severe illness). In
220 contrast, hospital monitoring provides evidence that the detected pathogen produces symptoms
221 that require hospital treatment. Third, our model is meant to be used now, in advance of future
222 pandemics, and not in the early months of a novel pandemic, because early detection systems
223 must be set up in advance of the next pandemic to be effective. Because we do not know the
224 epidemiological parameters of the next pandemic, our study assesses how these systems would
225 perform for a wide, representative variety of diseases with different epidemiological parameters,
226 in order to quantify these systems' benefits in general.

227 These results can inform ongoing international and national policy debates about which policies
228 are needed to mitigate future pandemics. In the wake of COVID-19, the World Health
229 Organization Intergovernmental Negotiating Body is actively negotiating a new treaty on
230 international pandemic preparedness which updates the International Health Regulations (2005).
231 Drafts of this treaty highlight "early warning and alert systems" as key measures¹⁶. Similarly, the
232 presidential administration of the United States has proposed investing \$5.3 billion over 7 to 10
233 years in early warning and real-time monitoring systems, including in hospitals and
234 wastewater¹⁷. In this study, we have assessed detection systems' detection times and have
235 developed a model to assess current and future detection system proposals. Along with additional
236 cost-effectiveness analysis and technical pilots²¹, these results can help inform which detection
237 systems are most effective and thus worth funding in pandemic preparedness efforts.

238 Methods

239 Description of model predicting cases at detection.

240 Our branching process-based model predicts the cumulative number of cases at the time of
241 detection for a given detection system and outbreak. It follows the approach of branching process
242 simulation models used previously to model the spread of COVID-19^{50,51} and other infectious
243 diseases⁵², but with the main added step of simulating each infected person's chance of being
244 detected by the detection system. The values for parameters for detection systems can be found
245 in table S2. The values for epidemiological parameters for outbreaks (R_0 , serial interval,
246 dispersion, hospitalization rate, and time to hospitalization) can be found in Fig. 2 and table S3.
247 As in previous models⁵², we assume the offspring distribution (the number of secondary cases
248 infected by each primary case) is negative binomial with mean R_0 .

249 We generally follow past detection system proposals^{19,21} to determine the implementation details
250 of each system in our model. Our model assumes the following. In hospital monitoring, hospitals
251 would test for high-priority pathogen families (e.g. coronaviruses) in patients presenting with
252 severe infectious symptoms in hospital emergency departments¹⁹. Similarly, in wastewater
253 monitoring, governments would test for pathogens in city wastewater treatment plants daily, and
254 monitor for high and increasing levels of high-priority pathogen families²¹. In air travel
255 monitoring, we model testing of individual symptomatic passengers (differs from proposals to
256 monitor airplane sewage²² or bridge air) on incoming international flights for the same
257 pathogens. The parameters of these systems are shown in table S2.

258 Our model also accounts for different delays involved in different detection systems. For
259 example, if the 500th case of a COVID-19-like outbreak triggers the detection threshold in both
260 the hospital and wastewater monitoring systems, because of the significant delay from infection
261 to hospitalization compared to the delay from infection to fecal shedding, the wastewater system
262 would catch the outbreak earlier.

263 In systems that test individuals (hospital and air travel individual monitoring), the threshold is
264 measured in an absolute *number* of cases. In systems that test wastewater (community and air
265 travel wastewater monitoring), the threshold or sensitivity is measured in prevalence (cases as a
266 *percentage* of the population)^{53,61,62}. To predict the number of cases and time to detection, we
267 need to convert this percentage back to a number of cases, so the wastewater detection time
268 depends on the catchment population size.

269 To estimate wastewater sensitivity measured in prevalence, we used data from^{53, 53} conducted
270 PCR testing for SARS-CoV-2 1687 longitudinal wastewater samples from 353 sampling
271 locations in 40 US states in early 2020, and synced these with publicly reported local daily new
272 COVID-19 case counts. This enables us to estimate a distribution of the wastewater sensitivity:
273 the lowest case count required to trigger positive detection in wastewater. Of the 353 sampling
274 locations, 47 had both SARS-CoV-2-positive and negative samples such that local case counts
275 on days of positive samples were all higher than those on days of negative samples. We thus
276 knew each sampling location's sensitivity is between the maximum of case counts on negative
277 sample days and the minimum of case counts on positive sample days. We took the midpoint of
278 this maximum and minimum as the location's sensitivity; this gave us 47 local sensitivities. We
279 fitted this to a log-normal distribution with a median of 2.5 daily new cases per 100,000 people.
280 As expected, this distribution is similarly shaped but slightly left-shifted of the distribution in

281 Figure 2b of⁵³ (median 3.7 per 100,000), because the latter distribution is an upper bound of the
282 former.

283 To use this distribution in our model, in each simulation run, we first randomly drew a
284 wastewater sensitivity from this distribution, and then we needed to convert this *reported*
285 *incidence i* to the *true (reported and unreported) number* of cases shedding fecally into public
286 wastewater systems up to the time of wastewater detection. We converted as follows. Let day T
287 be the day on which the incidence *i* is reported. First we assumed the wastewater SARS-CoV-2
288 level on day T is proportional to the number of COVID-19 cases who are fecally shedding on
289 day T, which we estimate as the number of fecal shedders infected 2 days before, given the
290 dominant peak in fecal shedding on day 2 of infection⁶¹. We infer the number of fecal shedders
291 infected on day T-2 from the incidence as follows. To account for underreporting, we first
292 estimate a true daily incidence of $5.7 \times i$ with symptom onset on day T, based on estimates of the
293 ratio of true (dated by symptom onset) to reported (dated by reporting date) COVID-19 cases in
294 the United States in early 2020⁵⁴. (This study's abstract reports true cases are 5-50x reported
295 ones, but this refers to the early March 1-April 4, 2020, period. We calculated the factor of 5.7
296 from the study's data when we use the fuller March 1-May 16 period, which overlaps better with
297 the February-June 2020 period in⁵³ and reflects less underreporting as the pandemic developed
298 and testing capacity increased. We calculated this underreporting factor as an average of state-
299 level underreporting factors, weighted by frequency of each state among the wastewater samples
300 in⁵³.) Finally, we multiply by (a) the fraction of cases who shed fecally (0.5⁶³) and (b) the
301 fraction of people connected to central sewage (0.8 in the US⁶⁴, which is the area from which
302 the⁵³ threshold is derived). This gives us the one-time prevalence of cases *p* who contribute to the
303 wastewater SARS-CoV-2 level on day T. For a given catchment with population *c*, this one-time
304 number of cases is *cp*, and we estimate the cumulative number of fecal shedders up to this time
305 as $\sum_{t=T}^0 \frac{cp}{R_0^{t/7}} \approx \int_{t=0}^T R_0^{t/7}$, where $T = \log_{R_0^{t/7}}(cp)$ is the number of days for the daily exponential
306 outbreak incidence curve to grow from 1 to *cp* cases.

307 To check this estimate, we identified studies that compared wastewater and hospital COVID-19
308 trends^{20,53}.²⁰ found that trends in wastewater SARS-CoV-2 values led trends in hospital
309 admissions by 1-4 days in New Haven (catchment size $2e+05$). We estimate that wastewater
310 detection would lead hospital detection of COVID-19 in New Haven by -0.8 to 3 weeks (90%
311 CI). This is consistent with the 1-4d lead estimate from²⁰. Similarly,⁵³ found that trends in
312 wastewater led those in clinical data by 4 days in Massachusetts (catchment size 2,300,000).
313 Their clinical data are dated by date of reporting rather than sample gathering; assuming that
314 hospital admissions are 5 days ahead of tests by date of reporting²⁰, then wastewater is 5d-4d=1
315 day behind hospital admissions. We estimate that wastewater detection would lead hospital
316 detection of COVID-19 in Massachusetts by -4 to -0.09 weeks (90% CI). This is consistent with
317 the 1-day lag estimate from⁵³.

318 Validation of model in US states.

319 We gathered two sources of data for each state: dates of COVID-19 detection and COVID-19
320 case counts in early 2020. For the former, we searched media reports and US state public health
321 press releases to determine the dates of the first COVID-19 case reported in each US state.
322 Sources for each state's detection date are listed in table S4. We were able to identify such dates
323 for all 50 states.

324 For the latter, we used literature estimates of true (tested and untested) COVID-19 case counts,
325 which incorporate COVID-19 mortality data to deal with variation in testing capacity among
326 states⁵⁴. We received a time series of weekly symptomatic COVID-19 case estimates for March
327 1-May 10, 2020 and divided by a symptomatic rate of 0.55 to get an estimate of total
328 (symptomatic and asymptomatic) cases⁴⁷. We specifically used estimates from the adjusted
329 mMAP (mortality maximum a posteriori) method because⁵⁴ had mMAP estimates for all 50
330 states, whereas other methods from the same study were missing estimates for various states. We
331 fit an exponential curve of case counts in each state to extrapolate cases back to January 2020. In
332 the data we received, all states had case data for all weeks in March 1-May 10, 2020.

333 We used our model to predict the weeks until detection in each US state (y-axis in Fig. S7).
334 Because most US states detected their first case by travel (table S4), we modeled a travel-based
335 detection system similarly to how we modeled the aforementioned detection systems. We
336 simulated a growing stream of imported travel cases (R_0^i cases for the i -th generation and global
337 $R_0 = 2.5$), and as for the other detection systems, we simulated infection and detection steps for
338 each generation, except that we only allowed travel-associated cases to be detected. We assumed
339 that the state COVID-19 outbreaks had the same values for all epidemiological parameters
340 except for R_0 , which we allowed to vary by state to account for state-specific conditions. We
341 obtained state-specific R_0 values from⁶⁵. The values for shared parameters were obtained from
342 literature (table S3). We used a detection delay of 12 days (5-day incubation period⁴⁷ plus 7-day
343 test and reporting turnaround in early 2020 in the US⁶⁶) because many first cases were detected
344 following symptoms. The only parameter we were unable to precisely estimate from literature
345 was the probability of a travel case being detected. We noted that this rate was at most the
346 COVID-19 symptomatic rate (0.55^{47}) and at least the hospitalization rate (0.03^{47}): in the highest-
347 detecting scenario, every symptomatic case would volunteer to be tested; in the lowest-detecting
348 scenario, only hospitalized travel cases would get flagged for testing. So we chose a rate of 0.1,
349 near the two rates' geometric mean. The predicted detection time for each state (the y-value
350 reported in Fig. S7) was the mean of 100 simulations.

351 We compared these predictions to ground truth estimates in each state (x-axis in Fig. S7). These
352 ground truth estimates were calculated by summing the aforementioned weekly case counts from
353 the first week of January 2020 until the date of detection in that state (Fig. S8).

354 Comparison of COVID-19 detection times in the actual pandemic versus with proposed early 355 detection systems.

356 We used our model to examine whether the early detection systems could have detected COVID-
357 19 earlier than in the actual pandemic. To do this, we used two data sources: (1) literature
358 estimates of total (tested and untested) COVID-19 case counts in late 2019 and early 2020⁵⁶ and
359 (2) simulation output from our model. We then used (1) to calculate the cumulative number of
360 cases when COVID-19 was actually detected, and compared this to results from (2).

361 For (1), we chose to use estimates from⁵⁶, which quantifies both the time of SARS-CoV-2
362 introduction into humans and the time series of cases following said introduction. These
363 estimates are based on phylodynamic rooting methods applied to SARS-CoV-2 sequence data,
364 combined with epidemic simulations and accounting for epidemiological data on the first known
365 cases of COVID-19. These estimates improve upon previous attempts to time SARS-CoV-2's

366 introduction into humans, which are solely based on phylodynamic rooting methods to quantify
367 the time to the most recent common ancestor of SARS-CoV-2 sequences⁶⁷.

368 As instructed by⁵⁶, we utilized
369 ‘BEAST.primary.IH.Dec10_16.linB.Dec15_25.linA.cumulativeInfections.timedGEMF_combine
370 d.stats.pickle’ from GitHub⁶⁸ to obtain the distribution of daily case counts. Conditioning on the
371 fact that there were at least six COVID-19-related hospitalizations by 2019-12-29⁶⁹, we
372 narrowed the distribution to those epidemic simulations with the top 25 percent of
373 hospitalizations and case counts. We simulated 100 draws from this distribution, and then took
374 the number of cases on 2019-12-29 in each simulation to get 100 values for the distribution of
375 cumulative cases at detection in the actual pandemic (‘Actual pandemic’ boxplot in Fig. 1B). We
376 chose 2019-12-29 as the date that COVID-19 was detected in the actual pandemic, because this
377 was the date of the first report of an outbreak of pneumonia cases to health authorities in
378 Wuhan⁷⁰.

379 For (2), we ran our model for COVID-19 (see table S3 for the epidemiological parameters used)
380 and all three detection systems (100 simulations for each system). For each detection system, this
381 gave us the estimated number of cases until detection of COVID-19 if that system had been in
382 place at the start of the pandemic. We assumed the system was present in the community in
383 which COVID-19 originated. We compared each system to the actual pandemic, and determined
384 that detection could have occurred earlier with the system if there was a statistically significant
385 difference in cases until detection between the actual pandemic and the simulated world with the
386 system (Fig. 1B). Statistical significance was assessed by a 1-sided t-test in which the alternative
387 hypothesis was that the detection system performed better.

388 We could empirically test our model predictions for the cases until wastewater detection by using
389 literature-estimated total COVID-19 cases in Massachusetts⁵⁴ and Massachusetts wastewater
390 SARS-CoV-2 data⁵⁷ in early 2020. We aimed to use these to estimate the cases until COVID-19
391 wastewater detection in Massachusetts in early 2020, but because Massachusetts wastewater
392 sampling for COVID-19 started only after the Massachusetts outbreak was underway,
393 wastewater samples were positive for SARS-CoV-2 on the first day of testing, so this first day of
394 testing was later than when wastewater detection could have caught SARS-CoV-2 if wastewater
395 detection had been in place in advance. Thus, we could only calculate an upper bound on the true
396 cases until detection. We utilized the wastewater time series from the Massachusetts Water
397 Resources Authority (MWRA) website and synced it with the COVID-19 case count time series
398 (Fig. S9). We multiplied the Massachusetts statewide cases by 0.33 (equal to
399 2,300,000/6,900,000) because the MWRA data covers 2,300,000 people, out of 6,900,000 people
400 in Massachusetts in 2020. We then summed these case counts up to the date of apparent
401 wastewater detection to get an upper bound for cases at detection, and checked whether our
402 model prediction was consistent with this bound.

403 Model-simulated cases required to trigger COVID-19 detection versus mathematical
404 approximations.

405 We compared the model simulations of cases until detection with our derived mathematical
406 formula, Equation (1) (Fig. S10). The points in Fig. S10 are the same as in Fig. 2A. The dashed
407 lines are generated by plugging values into Equation (1) for each detection system: we plugged

408 in the detection threshold, detection probability, outbreak R_0 , and detection delay (measured in
409 number of generations, i.e. serial intervals) for d , p_{test} , R_0 , and g , respectively.

410 Comparison of detection systems for different infectious diseases.

411 We applied our model to several outbreaks of recent interest: COVID-19, mpox (2022), polio
412 (2013-2014), Ebola (2013-2016) and flu (2009 pandemic) (Fig. 2A). Because of the lack of data
413 on the number of cases at the time of detection in previous outbreaks (except for the COVID-19
414 data used in Fig. 1B), we used our model to estimate status quo detection times for the outbreaks.
415 Because many recent outbreaks have been detected in healthcare settings^{59,69,71,72}, we assumed
416 status quo detection was similar to hospital monitoring, except with a lower detection probability
417 per case (p_{test}) to reflect that symptomatic cases are less likely to be tested for a panel of
418 diseases without the proposed systematic, proactive testing scheme. The per-case detection
419 probability for status quo was set to 0.67 times that of hospital monitoring to match our modeled
420 status quo detection times for COVID-19 with those estimated independently by⁵⁶ (Fig. 1B).

421 **Data availability:** Data are available in the supplement and at [https://github.com/abliu/early-](https://github.com/abliu/early-detection/releases)
422 [detection/releases](https://github.com/abliu/early-detection/releases).

423 **Code availability:** Code is available at <https://github.com/abliu/early-detection/releases>.

424 References

- 425 1. Kraemer, M. U. G. *et al.* [The effect of human mobility and control measures on the](#)
426 [COVID-19 epidemic in China](#). *Science* **368**, 493–497 (2020).
- 427 2. Meyerowitz-Katz, G. *et al.* [Is the cure really worse than the disease? The health impacts](#)
428 [of lockdowns during COVID-19](#). *BMJ Global Health* **6**, e006653 (2021).
- 429 3. Yamey, G. & Walensky, R. P. [Covid-19: Re-opening universities is high risk](#). *BMJ* **370**,
430 m3365 (2020).
- 431 4. Brauner, J. M. *et al.* [Inferring the effectiveness of government interventions against](#)
432 [COVID-19](#). *Science* **371**, eabd9338 (2021).
- 433 5. Ferguson, N. M. *et al.* [Report 9 - Impact of non-pharmaceutical interventions \(NPIs\) to](#)
434 [reduce COVID-19 mortality and healthcare demand](#).
435 [https://www.imperial.ac.uk/medicine/departments/school-public-health/infectious-disease-](https://www.imperial.ac.uk/medicine/departments/school-public-health/infectious-disease-epidemiology/mrc-global-infectious-disease-analysis/covid-19/report-9-impact-of-npis-on-covid-19/)
436 [epidemiology/mrc-global-infectious-disease-analysis/covid-19/report-9-impact-of-npis-on-covid-](https://www.imperial.ac.uk/medicine/departments/school-public-health/infectious-disease-epidemiology/mrc-global-infectious-disease-analysis/covid-19/report-9-impact-of-npis-on-covid-19/)
437 [19/](https://www.imperial.ac.uk/medicine/departments/school-public-health/infectious-disease-epidemiology/mrc-global-infectious-disease-analysis/covid-19/report-9-impact-of-npis-on-covid-19/) (2020).
- 438 6. Levine-Tiefenbrun, M. *et al.* [SARS-CoV-2 RT-qPCR Test Detection Rates Are](#)
439 [Associated with Patient Age, Sex, and Time since Diagnosis](#). *The Journal of Molecular*
440 *Diagnostics* **24**, 112–119 (2022).
- 441 7. Larremore, D. B. *et al.* [Test sensitivity is secondary to frequency and turnaround time for](#)
442 [COVID-19 screening](#). *Science Advances* eabd5393 (2020) doi:[10.1126/sciadv.abd5393](https://doi.org/10.1126/sciadv.abd5393).
- 443 8. Liu, A. B. *et al.* [Association of COVID-19 Quarantine Duration and Postquarantine](#)
444 [Transmission Risk in 4 University Cohorts](#). *JAMA Network Open* **5**, e220088 (2022).

- 445 9. Wyllie, A. L. *et al.* [Saliva or Nasopharyngeal Swab Specimens for Detection of SARS-](#)
446 [CoV-2](#). *New England Journal of Medicine* **383**, 1283–1286 (2020).
- 447 10. Yelin, I. *et al.* [Evaluation of COVID-19 RT-qPCR Test in Multi sample Pools](#). *Clinical*
448 *Infectious Diseases* **71**, 2073–2078 (2020).
- 449 11. Karin, O. *et al.* [Cyclic exit strategies to suppress COVID-19 and allow economic activity](#).
450 (2020) doi:[10.1101/2020.04.04.20053579](https://doi.org/10.1101/2020.04.04.20053579).
- 451 12. Flaxman, S. *et al.* [Estimating the effects of non-pharmaceutical interventions on COVID-](#)
452 [19 in Europe](#). *Nature* **584**, 257–261 (2020).
- 453 13. Hatchett, R. J., Mecher, C. E. & Lipsitch, M. [Public health interventions and epidemic](#)
454 [intensity during the 1918 influenza pandemic](#). *Proceedings of the National Academy of Sciences*
455 **104**, 7582–7587 (2007).
- 456 14. Peak, C. M., Childs, L. M., Grad, Y. H. & Buckee, C. O. [Comparing nonpharmaceutical](#)
457 [interventions for containing emerging epidemics](#). *Proceedings of the National Academy of*
458 *Sciences* **114**, 4023–4028 (2017).
- 459 15. Pei, S., Kandula, S. & Shaman, J. [Differential effects of intervention timing on COVID-](#)
460 [19 spread in the United States](#). *Science Advances* **6**, eabd6370 (2020).
- 461 16. Bureau of the Intergovernmental Negotiating Body, World Health Organization.
462 [Conceptual zero draft for the consideration of the Intergovernmental Negotiating Body at its](#)
463 [third meeting](#). 32 (2022).
- 464 17. Lander, E. & Sullivan, J. [American Pandemic Preparedness: Transforming Our](#)
465 [Capabilities](#). 27 (2021).
- 466 18. Botti-Lodovico, Y. *et al.* [The Origins and Future of Sentinel: An Early-Warning System](#)
467 [for Pandemic Preemption and Response](#). *Viruses* **13**, 1605 (2021).
- 468 19. Ecker, D. J. [How to Snuff Out the Next Pandemic](#). *Scientific American Blog Network*
469 (2020).
- 470 20. Peccia, J. *et al.* [Measurement of SARS-CoV-2 RNA in wastewater tracks community](#)
471 [infection dynamics](#). *Nature Biotechnology* **38**, 1164–1167 (2020).
- 472 21. The Nucleic Acid Observatory Consortium. [A Global Nucleic Acid Observatory for](#)
473 [Biodefense and Planetary Health](#). *arXiv:2108.02678 [q-bio]* (2021).
- 474 22. Hjelmsø, M. H. *et al.* [Metagenomic analysis of viruses in toilet waste from long distance](#)
475 [flights—A new procedure for global infectious disease surveillance](#). *PLOS ONE* **14**, e0210368
476 (2019).
- 477 23. Muntean, J., Howard, K. & Atwood, P. [CDC has tested wastewater from aircraft amid](#)
478 [concerns over Covid-19 surge in China](#). *CNN* (2023).

- 479 24. Nordahl Petersen, T. *et al.* [Meta-genomic analysis of toilet waste from long distance](#)
480 [flights; a step towards global surveillance of infectious diseases and antimicrobial resistance.](#)
481 *Scientific Reports* **5**, 11444 (2015).
- 482 25. Mina, M. J. *et al.* [A Global Immunological Observatory to meet a time of pandemics.](#)
483 *eLife* **9**, e58989 (2020).
- 484 26. Chu, H. Y. *et al.* [Early Detection of Covid-19 through a Citywide Pandemic Surveillance](#)
485 [Platform.](#) *New England Journal of Medicine* **383**, 185–187 (2020).
- 486 27. Brownstein, J. S., Freifeld, C. C. & Madoff, L. C. [Digital Disease Detection —](#)
487 [Harnessing the Web for Public Health Surveillance.](#) *The New England journal of medicine* **360**,
488 2153–2157 (2009).
- 489 28. Dugas, A. F. *et al.* [Influenza Forecasting with Google Flu Trends.](#) *PLOS ONE* **8**, e56176
490 (2013).
- 491 29. Bajema, N., Beaver, W. & Parthemore, C. [Toward a Global Pathogen Early Warning](#)
492 [System: Building on the Landscape of Biosurveillance Today.](#)
493 <https://councilonstrategicrisks.org/wp-content/uploads/2021/07/Toward-A-Global-Pathogen->
494 [Early-Warning-System_2021_07_20-1.pdf](#) (2021).
- 495 30. Lee, V. J., Chiew, C. J. & Khong, W. X. [Interrupting transmission of COVID-19:](#)
496 [Lessons from containment efforts in Singapore.](#) *Journal of Travel Medicine* **27**, (2020).
- 497 31. Keshaviah, A. *et al.* [Wastewater monitoring can anchor global disease surveillance](#)
498 [systems.](#) *The Lancet Global Health* **11**, e976–e981 (2023).
- 499 32. Petros, B. A. *et al.* [Multimodal surveillance of SARS-CoV-2 at a university enables](#)
500 [development of a robust outbreak response framework.](#) *Med* **0**, (2022).
- 501 33. Boehm, A. B. *et al.* [Regional Replacement of SARS-CoV-2 Variant Omicron BA.1 with](#)
502 [BA.2 as Observed through Wastewater Surveillance.](#) *Environmental Science & Technology*
503 *Letters* **9**, 575–580 (2022).
- 504 34. Kim, J. E., Lee, J. H., Lee, H., Moon, S. J. & Nam, E. W. [COVID-19 screening center](#)
505 [models in South Korea.](#) *Journal of Public Health Policy* **42**, 15–26 (2021).
- 506 35. Meakin, S. *et al.* [Comparative assessment of methods for short-term forecasts of COVID-](#)
507 [19 hospital admissions in England at the local level.](#) *BMC Medicine* **20**, 86 (2022).
- 508 36. Arizti-Sanz, J. *et al.* [Streamlined inactivation, amplification, and Cas13-based detection](#)
509 [of SARS-CoV-2.](#) *Nature Communications* **11**, 5921 (2020).
- 510 37. Gregory, D. A., Wieberg, C. G., Wenzel, J., Lin, C.-H. & Johnson, M. C. [Monitoring](#)
511 [SARS-CoV-2 Populations in Wastewater by Amplicon Sequencing and Using the Novel](#)
512 [Program SAM Refiner.](#) *Viruses* **13**, 1647 (2021).

- 513 38. Li, C. *et al.* Population Normalization in SARS-CoV-2 Wastewater-Based Epidemiology:
514 Implications from Statewide Wastewater Monitoring in Missouri. (2022)
515 doi:[10.1101/2022.09.08.22279459](https://doi.org/10.1101/2022.09.08.22279459).
- 516 39. Robinson, C. A. *et al.* Defining biological and biophysical properties of SARS-CoV-2
517 genetic material in wastewater. *Science of The Total Environment* **807**, 150786 (2022).
- 518 40. Bibby, K. & Peccia, J. Identification of Viral Pathogen Diversity in Sewage Sludge by
519 Metagenome Analysis. *Environmental Science & Technology* **47**, 1945–1951 (2013).
- 520 41. Creager, H. M. *et al.* Clinical evaluation of the BioFire® Respiratory Panel 2.1 and
521 detection of SARS-CoV-2. *Journal of Clinical Virology* **129**, 104538 (2020).
- 522 42. Edin, A., Eilers, H. & Allard, A. Evaluation of the Biofire Filmarray Pneumonia panel
523 plus for lower respiratory tract infections. *Infectious Diseases* **52**, 479–488 (2020).
- 524 43. Murphy, C. N. *et al.* Multicenter Evaluation of the BioFire FilmArray
525 Pneumonia/Pneumonia Plus Panel for Detection and Quantification of Agents of Lower
526 Respiratory Tract Infection. *Journal of Clinical Microbiology* **58**, e00128–20 (2020).
- 527 44. Quick, J. *et al.* Multiplex PCR method for MinION and Illumina sequencing of Zika and
528 other virus genomes directly from clinical samples. *Nature Protocols* **12**, 1261–1276 (2017).
- 529 45. Ackerman, C. M. *et al.* Massively multiplexed nucleic acid detection with Cas13. *Nature*
530 **582**, 277–282 (2020).
- 531 46. Chiu, C. Y. & Miller, S. A. Clinical metagenomics. *Nature Reviews Genetics* **20**, 341–
532 355 (2019).
- 533 47. Bar-On, Y. M., Sender, R., Flamholz, A. I., Phillips, R. & Milo, R. A quantitative
534 compendium of COVID-19 epidemiology. *arXiv:2006.01283 [q-bio]* (2020).
- 535 48. Du, Z. *et al.* Reproduction number of monkeypox in the early stage of the 2022 multi-
536 country outbreak. *Journal of Travel Medicine* taac099 (2022) doi:[10.1093/jtm/taac099](https://doi.org/10.1093/jtm/taac099).
- 537 49. U.S. Centers for Disease Control and Prevention (CDC). United States confirmed as
538 country with circulating vaccine-derived poliovirus. *CDC* (2022).
- 539 50. Bradshaw, W. J., Alley, E. C., Huggins, J. H., Lloyd, A. L. & Esvelt, K. M. Bidirectional
540 contact tracing could dramatically improve COVID-19 control. *Nature Communications* **12**, 232
541 (2021).
- 542 51. Hellewell, J. *et al.* Feasibility of controlling COVID-19 outbreaks by isolation of cases
543 and contacts. *The Lancet Global Health* **8**, e488–e496 (2020).
- 544 52. Lloyd-Smith, J. O., Schreiber, S. J., Kopp, P. E. & Getz, W. M. Superspreading and the
545 effect of individual variation on disease emergence. *Nature* **438**, 355–359 (2005).
- 546 53. Wu, F. *et al.* Wastewater surveillance of SARS-CoV-2 across 40 U.S. States from
547 February to June 2020. *Water Research* **202**, 117400 (2021).

- 548 54. Lu, F. S. *et al.* Estimating the cumulative incidence of COVID-19 in the United States
549 using influenza surveillance, virologic testing, and mortality data: Four complementary
550 approaches. *PLOS Computational Biology* **17**, e1008994 (2021).
- 551 55. Abbott, S. *et al.* Epiforecasts/EpiNow2: 1.3.4 release. (2023)
552 doi:[10.5281/zenodo.7611804](https://doi.org/10.5281/zenodo.7611804).
- 553 56. Pekar, J. E. *et al.* The molecular epidemiology of multiple zoonotic origins of SARS-
554 CoV-2. *Science* **377**, 960–966 (2022).
- 555 57. Massachusetts Water Resources Authority. [MWRA - Wastewater COVID-19 Tracking](#).
556 (2022).
- 557 58. Adhikari, S. & Halden, R. U. Opportunities and limits of wastewater-based epidemiology
558 for tracking global health and attainment of UN sustainable development goals. *Environment*
559 *International* **163**, 107217 (2022).
- 560 59. Sack, K., Fink, S., Belluck, P. & Nossiter, A. How Ebola Roared Back. *The New York*
561 *Times* (2014).
- 562 60. Brouwer, A. F. *et al.* Epidemiology of the silent polio outbreak in Rahat, Israel, based on
563 modeling of environmental surveillance data. *Proceedings of the National Academy of Sciences*
564 **115**, E10625–E10633 (2018).
- 565 61. Wu, F. *et al.* SARS-CoV-2 RNA concentrations in wastewater foreshadow dynamics and
566 clinical presentation of new COVID-19 cases. *Science of The Total Environment* **805**, 150121
567 (2022).
- 568 62. Soller, J. *et al.* Modeling infection from SARS-CoV-2 wastewater concentrations:
569 Promise, limitations, and future directions. *Journal of Water and Health* **20**, 1197–1211 (2022).
- 570 63. Jones, D. L. *et al.* Shedding of SARS-CoV-2 in feces and urine and its potential role in
571 person-to-person transmission and the environment-based spread of COVID-19. *The Science of*
572 *the Total Environment* **749**, 141364 (2020).
- 573 64. Biobot Analytics. *The Effect of Septic Systems on Wastewater-Based Epidemiology*.
574 [http://biobot.io/wp-](http://biobot.io/wp-content/uploads/2022/09/BIOBOT_WHITEPAPER_EFFECT_OF_SEPTIC_V01-1.pdf)
575 [content/uploads/2022/09/BIOBOT_WHITEPAPER_EFFECT_OF_SEPTIC_V01-1.pdf](http://biobot.io/wp-content/uploads/2022/09/BIOBOT_WHITEPAPER_EFFECT_OF_SEPTIC_V01-1.pdf) (2022).
- 576 65. Mallela, A. *et al.* Bayesian Inference of State-Level COVID-19 Basic Reproduction
577 Numbers across the United States. *Viruses* **14**, 157 (2022).
- 578 66. Goldberg, C. Mass. Public Health Lab Can Now Test For New Coronavirus, Speeding
579 Results. (2020).
- 580 67. Pekar, J., Worobey, M., Moshiri, N., Scheffler, K. & Wertheim, J. O. Timing the SARS-
581 CoV-2 index case in Hubei province. *Science* **372**, 412–417 (2021).
- 582 68. Pekar, J. E. *et al.* [Sars-cov-2-origins / multi-introduction](#). *GitHub* (2022).

- 583 69. Yuan, Yao, Yujie, Ma, Jialu, Zhou & Wenkun, Hou. [Xinhua Headlines: Chinese doctor](#)
584 [recalls first encounter with mysterious virus - Xinhua English.news.cn](#). *Xinhua* (2020).
- 585 70. The 2019-nCoV Outbreak Joint Field Epidemiology Investigation Team & Li, Q. [An](#)
586 [Outbreak of NCIP \(2019-nCoV\) Infection in China — Wuhan, Hubei Province, 2019–2020](#).
587 *China CDC Weekly* **2**, 79–80 (2020).
- 588 71. Rosenthal, E. [THE SARS EPIDEMIC: THE PATH; From China’s Provinces, a Crafty](#)
589 [Germ Breaks Out](#). *The New York Times* (2003).
- 590 72. Garrett, L. *The Coming Plague*. (1994).
- 591 73. Center for Systems Science and Engineering (CSSE) at Johns Hopkins University.
592 [COVID-19 Data Repository by the Center for Systems Science and Engineering \(CSSE\) at Johns](#)
593 [Hopkins University](#). (2022).
- 594 74. Mathieu, E. *et al.* [Total COVID-19 tests per 1,000 people](#). *Our World in Data* (2020).
- 595 75. Harris, T. E. *The theory of branching processes*.
596 <https://www.rand.org/content/dam/rand/pubs/reports/2009/R381.pdf> (1964).
- 597 76. Yoo, S. J. *et al.* [Frequent Detection of Pandemic \(H1N1\) 2009 Virus in Stools of](#)
598 [Hospitalized Patients](#). *Journal of Clinical Microbiology* **48**, 2314–2315 (2010).
- 599 77. Vetter, P. *et al.* [Ebola Virus Shedding and Transmission: Review of Current Evidence](#).
600 *The Journal of Infectious Diseases* **214**, S177–S184 (2016).
- 601 78. Independent Evaluation Department, Asia Development Bank. *People’s Republic of*
602 *China: Wuhan Wastewater and Stormwater Management Project*.
603 <https://www.adb.org/sites/default/files/evaluation-document/188852/files/pvr-447.pdf> (2016).
- 604 79. Faes, C. *et al.* [Time between Symptom Onset, Hospitalisation and Recovery or Death:](#)
605 [Statistical Analysis of Belgian COVID-19 Patients](#). *International Journal of Environmental*
606 *Research and Public Health* **17**, 7560 (2020).
- 607 80. He, D. *et al.* [Low dispersion in the infectiousness of COVID-19 cases implies difficulty](#)
608 [in control](#). *BMC Public Health* **20**, 1558 (2020).
- 609 81. Wang, S. *et al.* [Serial intervals and incubation periods of the monkeypox virus clades](#).
610 *Journal of Travel Medicine* taac105 (2022) doi:[10.1093/jtm/taac105](https://doi.org/10.1093/jtm/taac105).
- 611 82. DeWitt, M. E. *et al.* [Global monkeypox case hospitalisation rates: A rapid systematic](#)
612 [review and meta-analysis](#). *eClinicalMedicine* **54**, 101710 (2022).
- 613 83. Reda, A., Hemmeda, L., Brakat, A. M., Sah, R. & El-Qushayri, A. E. [The clinical](#)
614 [manifestations and severity of the 2022 monkeypox outbreak among 4080 patients](#). *Travel*
615 *Medicine and Infectious Disease* **50**, 102456 (2022).
- 616 84. Ryan, E. T., Hill, D. R., Solomon, T., Aronson, N. & Endy, T. P. *Hunter’s Tropical*
617 *Medicine and Emerging Infectious Diseases E-Book*. (Elsevier Health Sciences, 2019).

- 618 85. Mailhe, M. *et al.* Clinical characteristics of ambulatory and hospitalized patients with
619 monkeypox virus infection: An observational cohort study. *Clinical Microbiology and Infection*
620 (2022) doi:[10.1016/j.cmi.2022.08.012](https://doi.org/10.1016/j.cmi.2022.08.012).
- 621 86. Tiwari, A. *et al.* [Monkeypox outbreak: Wastewater and environmental surveillance](#)
622 [perspective](#). *Science of The Total Environment* **856**, 159166 (2023).
- 623 87. Wolfe, M. K. *et al.* Detection of monkeypox viral DNA in a routine wastewater
624 monitoring program. (2022) doi:[10.1101/2022.07.25.22278043](https://doi.org/10.1101/2022.07.25.22278043).
- 625 88. U.S. Centers for Disease Control and Prevention (CDC) National Center for Health
626 Statistics. [Immunization](#). (2022).
- 627 89. World Health Organization. [Circulating vaccine-derived poliovirus type 3 – Israel](#).
628 (2022).
- 629 90. Estivariz, Concepcion F., Link-Gelles, Ruth & Shimabukuro, Tom. [Pinkbook:](#)
630 [Poliomyelitis](#) CDC. (2022).
- 631 91. Nelson, K. E. & Williams, C. M. Determinants of Epidemic Growth. In Chapter 6:
632 Infectious Disease Dynamics. in *Infectious Disease Epidemiology: Theory and Practice* 135
633 (Jones & Bartlett Learning, 2013).
- 634 92. Chung, P. W., Huang, Y. C., Chang, L. Y., Lin, T. Y. & Ning, H. C. [Duration of](#)
635 [enterovirus shedding in stool](#). *Journal of microbiology, immunology, and infection* **34**, 167–170
636 (2001).
- 637 93. Onorato, I. M. *et al.* [Mucosal Immunity Induced by Enhanced-Potency Inactivated and](#)
638 [Oral Polio Vaccines](#). *The Journal of Infectious Diseases* **163**, 1–6 (1991).
- 639 94. Wong, Z. S. Y., Bui, C. M., Chughtai, A. A. & Macintyre, C. R. [A systematic review of](#)
640 [early modelling studies of Ebola virus disease in West Africa](#). *Epidemiology & Infection* **145**,
641 1069–1094 (2017).
- 642 95. Drake, J. M. *et al.* [Ebola Cases and Health System Demand in Liberia](#). *PLOS Biology* **13**,
643 e1002056 (2015).
- 644 96. Van Kerkhove, M. D., Bento, A. I., Mills, H. L., Ferguson, N. M. & Donnelly, C. A. [A](#)
645 [review of epidemiological parameters from Ebola outbreaks to inform early public health](#)
646 [decision-making](#). *Scientific Data* **2**, 150019 (2015).
- 647 97. Qureshi, A. I. *et al.* [High Survival Rates and Associated Factors Among Ebola Virus](#)
648 [Disease Patients Hospitalized at Donka National Hospital, Conakry, Guinea](#). *Journal of Vascular*
649 *and Interventional Neurology* **8**, S4–S11 (2015).
- 650 98. Fraser, C. *et al.* [Pandemic Potential of a Strain of Influenza A \(H1N1\): Early Findings](#).
651 *Science (New York, N.Y.)* **324**, 1557–1561 (2009).

- 652 99. Heijnen, L. & Medema, G. Surveillance of Influenza A and the pandemic influenza A
653 (H1N1) 2009 in sewage and surface water in the Netherlands. *Journal of Water and Health* **9**,
654 434–442 (2011).
- 655 100. Presanis, A. M. *et al.* The Severity of Pandemic H1N1 Influenza in the United States,
656 from April to July 2009: A Bayesian Analysis. *PLOS Medicine* **6**, e1000207 (2009).
- 657 101. Furuya-Kanamori, L. *et al.* Heterogeneous and Dynamic Prevalence of Asymptomatic
658 Influenza Virus Infections. *Emerging Infectious Diseases* **22**, 1052–1056 (2016).
- 659 102. Lessler, J., Reich, N. G. & Cummings, D. A. T. Outbreak of 2009 Pandemic Influenza A
660 (H1N1) at a New York City School. *New England Journal of Medicine* **361**, 2628–2636 (2009).
- 661 103. Louie, J. K. *et al.* Factors Associated With Death or Hospitalization Due to Pandemic
662 2009 Influenza A(H1N1) Infection in California. *JAMA* **302**, 1896–1902 (2009).
- 663 104. Vynnycky, E. & White, R. *An Introduction to Infectious Disease Modelling*. (Oxford
664 University Press, 2010).
- 665 105. Morgan, D. *et al.* HIV-1 infection in rural Africa: Is there a difference in median time to
666 AIDS and survival compared with that in industrialized countries? *AIDS (London, England)* **16**,
667 597–603 (2002).
- 668 106. Todd, J. *et al.* Time from HIV seroconversion to death: A collaborative analysis of eight
669 studies in six low and middle-income countries before highly active antiretroviral therapy. *AIDS*
670 (London, England) **21**, S55 (2007).
- 671 107. CASCADE. Time from HIV-1 seroconversion to AIDS and death before widespread use
672 of highly-active antiretroviral therapy: A collaborative re-analysis. Collaborative Group on AIDS
673 Incubation and HIV Survival including the CASCADE EU Concerted Action. Concerted Action
674 on SeroConversion to AIDS and Death in Europe. *Lancet (London, England)* **355**, 1131–1137
675 (2000).
- 676 108. Crum-Cianflone, N. F. HIV and the Gastrointestinal Tract. *Infectious diseases in clinical*
677 *practice (Baltimore, Md.)* **18**, 283–285 (2010).
- 678 109. Yolken, R. H., Li, S., Perman, J. & Viscidi, R. Persistent Diarrhea and Fecal Shedding of
679 Retroviral Nucleic Acids in Children Infected with Human Immunodeficiency Virus. *The*
680 *Journal of Infectious Diseases* **164**, 61–66 (1991).
- 681 110. World Health Organization. *Tuberculosis (TB)*. (2022).
- 682 111. Behr, M. A., Edelstein, P. H. & Ramakrishnan, L. Revisiting the timetable of
683 tuberculosis. *The BMJ* **362**, k2738 (2018).
- 684 112. Ma, Y., Horsburgh, C. R., White, L. F. & Jenkins, H. E. Quantifying TB transmission: A
685 systematic review of reproduction number and serial interval estimates for tuberculosis.
686 *Epidemiology and Infection* **146**, 1478–1494 (2018).

687 **Acknowledgments:** We gratefully acknowledge Mauricio Santillana, Nicholas B. Link, Fred S.
688 Lu, and Andre T. Nguyen for sharing their estimates on COVID-19 incidence in the US states.
689 We also acknowledge Jonathan Pekar, Joel Wertheim, and Michael Worobey for sharing their
690 estimates on COVID-19 incidence in late 2019 and early 2020. We finally acknowledge Michael
691 McLaren, Becky Ward, and Quincey Justman for feedback on the manuscript.

692 **Funding:**

693 Lynch Foundation Fellows Program in Systems Biology at Harvard Medical School (ABL)

694 National Library of Medicine grant T15LM007092 (DL)

695 National Institutes of Health grant R01GM120122 (APJ)

696 CDC contract 200-2016-91779 (WPH)

697 National Institutes of Health grant R01GM120122 (MS)

698 This project has been funded (in part) by contract 200-2016-91779 with the Centers for Disease
699 Control and Prevention.

700 Disclaimer: The findings, conclusions, and views expressed are those of the author(s) and do not
701 necessarily represent the official position of the Centers for Disease Control and Prevention
702 (CDC).

703 **Author contributions:**

704 Conceptualization: ABL, MS

705 Methodology: ABL, APJ, MS

706 Software: ABL

707 Validation: ABL

708 Formal analysis: ABL, DL

709 Investigation: ABL, DL

710 Resources: MS

711 Data Curation: ABL

712 Writing - Original Draft: ABL, MS

713 Writing - Review & Editing: ABL, DL, APJ, WPH, MS

714 Visualization: ABL, MS

715 Supervision: ABL, WPH, MS

716 Project administration: ABL, MS

717 Funding acquisition: MS

718 **Competing interests:** WPH is a member of the scientific advisory board and has stock options
719 in BioBot Analytics. MS is a cofounder of Rhinostics and consults for the diagnostic consulting
720 company Vectis Solutions LLC. The other authors declare that they have no competing interests.