

Manuscript version: Author's Accepted Manuscript

The version presented in WRAP is the author's accepted manuscript and may differ from the published version or Version of Record.

Persistent WRAP URL:

<http://wrap.warwick.ac.uk/151484>

How to cite:

Please refer to published version for the most recent bibliographic citation information. If a published version is known of, the repository item page linked to above, will contain details on accessing it.

Copyright and reuse:

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Publisher's statement:

Please refer to the repository item page, publisher's statement section, for further information.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk.

1 **Interventions targeting nonsymptomatic cases can be important to**
2 **prevent local outbreaks: SARS-CoV-2 as a case-study**

3

4 **Authors**

5 Francesca A. Lovell-Read^{1*}, Sebastian Funk², Uri Obolski^{3,4}, Christl A. Donnelly^{5,6}, Robin N.
6 Thompson^{1,2,7,8}

7

8 **Affiliations**

9 ¹Mathematical Institute, University of Oxford, Oxford, UK

10 ²Centre for Mathematical Modelling of Infectious Diseases, London School of Hygiene &
11 Tropical Medicine, London, UK

12 ³Porter School of the Environment and Earth Sciences, Tel Aviv University, Tel Aviv, Israel

13 ⁴School of Public Health, Tel Aviv University, Tel Aviv, Israel

14 ⁵Department of Statistics, University of Oxford, Oxford, UK

15 ⁶MRC Centre for Global Infectious Disease Analysis, Department of Infectious Disease
16 Epidemiology, Imperial College London, London, UK

17 ⁷Mathematics Institute, University of Warwick, Coventry, UK

18 ⁸The Zeeman Institute for Systems Biology and Infectious Disease Epidemiology Research,
19 University of Warwick, Coventry, UK

20

21 * **Corresponding author**

22 Address: Merton College, Merton Street, Oxford, OX1 4JD

23 E-mail: francesca.lovell-read@merton.ox.ac.uk. Tel: +44 (0)7511 676 592.

24

25 **Abstract**

26 During infectious disease epidemics, an important question is whether cases travelling to new
27 locations will trigger local outbreaks. The risk of this occurring depends on the transmissibility of
28 the pathogen, the susceptibility of the host population and, crucially, the effectiveness of
29 surveillance in detecting cases and preventing onward spread. For many pathogens, transmission
30 from presymptomatic and/or asymptomatic (together referred to as nonsymptomatic) infectious
31 hosts can occur, making effective surveillance challenging. Here, using SARS-CoV-2 as a case-
32 study, we show how the risk of local outbreaks can be assessed when nonsymptomatic
33 transmission can occur. We construct a branching process model that includes nonsymptomatic
34 transmission, and explore the effects of interventions targeting nonsymptomatic or symptomatic
35 hosts when surveillance resources are limited. We consider whether the greatest reductions in local
36 outbreak risks are achieved by increasing surveillance and control targeting nonsymptomatic or
37 symptomatic cases, or a combination of both. We find that seeking to increase surveillance of
38 symptomatic hosts alone is typically not the optimal strategy for reducing outbreak risks. Adopting
39 a strategy that combines an enhancement of surveillance of symptomatic cases with efforts to find
40 and isolate nonsymptomatic infected hosts leads to the largest reduction in the probability that
41 imported cases will initiate a local outbreak.

42

43 **Keywords**

44 mathematical modelling; infectious disease epidemiology; SARS-CoV-2; COVID-19;
45 presymptomatic infection; asymptomatic infection; disease surveillance

46

47 **1. Introduction**

48 Emerging epidemics represent a substantial challenge to human health worldwide [1-4]. When
49 cases are clustered in specific locations, two key questions are: i) Will exported cases lead to local
50 outbreaks in new locations? and ii) Which surveillance and control strategies in those new
51 locations will reduce the risk of local outbreaks?

52
53 Branching process models are used for a range of diseases to assess whether cases that are newly
54 arrived in a host population will generate a local outbreak driven by sustained local transmission
55 [5-11]. These models can also be used to predict the effectiveness of potential control
56 interventions. For example, early in the COVID-19 pandemic, Hellewell *et al.* [12] used
57 simulations of a branching process model to predict whether or not new outbreaks would fade out
58 under different contact tracing strategies. Thompson [13] estimated the probability of local
59 outbreaks analytically using a branching process model and found that effective isolation of
60 infectious hosts leads to a substantial reduction in the outbreak risk.

61
62 A factor that can hinder control interventions during any epidemic is the potential for individuals
63 to transmit a pathogen while not showing symptoms. For COVID-19, the incubation period has
64 been estimated to last approximately five or six days on average [14, 15], and presymptomatic
65 transmission can occur during that period [16-20]. Additionally, asymptomatic infected individuals
66 (those who never develop symptoms) are also thought to contribute to transmission [16, 21, 22].

67
68 Motivated by the need to assess the risk of outbreaks outside China early in the COVID-19
69 pandemic, we show how the risk that imported cases will lead to local outbreaks can be estimated
70 using a branching process model. Unlike standard approaches for estimating the probability of a
71 major epidemic analytically [23-26], nonsymptomatic individuals are included in the model

72 explicitly. Using a function that characterises the efficacy of interventions for different
73 surveillance efforts (denoted $f(\rho, \delta)$ in the model), we explore the effects of interventions that aim
74 to reduce this risk. Under the assumption that detected infected hosts are isolated effectively, we
75 consider whether it is most effective to dedicate resources to enhancing surveillance targeting
76 symptomatic individuals, to instead focus on increasing surveillance for nonsymptomatic
77 individuals, or to use a combination of these approaches.

78
79 We show that, when surveillance resources are limited, the maximum reduction in the outbreak
80 risk almost always corresponds to a mixed strategy involving enhanced surveillance of both
81 symptomatic and nonsymptomatic hosts. This remains the case even if the surveillance effort
82 required to find nonsymptomatic infected individuals is significantly larger than the effort required
83 to find symptomatic individuals. This highlights the benefits of not only seeking to find and isolate
84 symptomatic hosts, but also dedicating resources to detecting nonsymptomatic cases during
85 infectious disease epidemics.

86

87 **2. Methods**

88 **2.1 Model**

89 We consider a branching process model in which infectious individuals are classified as
90 asymptomatic (A), presymptomatic (I_1) or symptomatic (I_2). Hosts in any of these classes may
91 generate new infections. The parameter ξ represents the proportion of new infections that are
92 asymptomatic, so that a new infection either involves increasing A by one (with probability ξ) or
93 increasing I_1 by one (with probability $1 - \xi$).

94

95 Presymptomatic hosts may go on to develop symptoms (transition from I_1 to I_2) or be detected and
 96 isolated (so that I_1 decreases by one). Symptomatic individuals (I_2) can be isolated (so that I_2
 97 decreases by one) or be removed due to recovery or death (so that again I_2 decreases by one).
 98 Similarly, asymptomatic hosts may be detected and isolated, or recover (so that A decreases by
 99 one in either case).

100
 101 A schematic showing the different possible events in the model is shown in Fig 1A. The analogous
 102 compartmental differential equation model to the branching process model that we consider is
 103 given by

$$104 \quad \frac{dA}{dt} = \xi(\eta\beta A + \alpha\beta I_1 + \beta I_2) - \frac{\varepsilon\gamma}{1 - f(\rho_1, \delta)} A - \nu A,$$

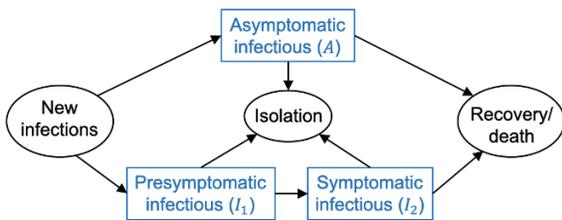
$$105 \quad \frac{dI_1}{dt} = (1 - \xi)(\eta\beta A + \alpha\beta I_1 + \beta I_2) - \frac{\varepsilon\gamma}{1 - f(\rho_1, \delta)} I_1 - \lambda I_1,$$

$$106 \quad \frac{dI_2}{dt} = \lambda I_1 - \frac{\gamma}{1 - f(\rho_2, \delta)} I_2 - \mu I_2.$$

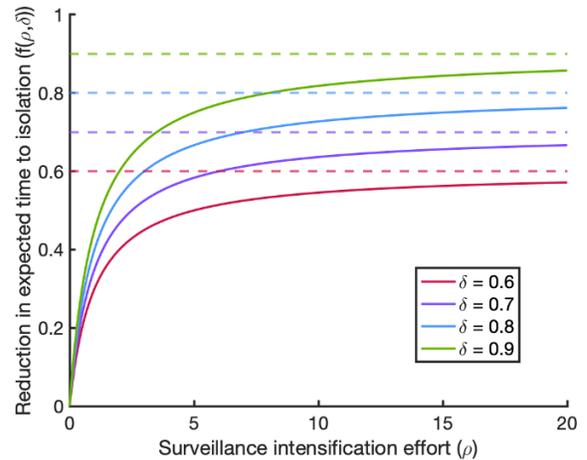
107 The parameters of the model, and the form of the function $f(\rho, \delta)$ that describes how the expected
 108 time to isolation is reduced for a given surveillance effort, are outlined below.

109

A.



B.



110

111 **Fig 1. The branching process model used in our analyses.** A. Schematic showing the different event
112 types in the branching process model. The parameters of the model are described in the text and in Table
113 1. B. The relationship between the surveillance intensification effort (ρ) and the proportional reduction in
114 the expected time to isolation ($f(\rho, \delta)$), shown for different values of the parameter δ (solid lines). The
115 parameter $\delta \in (0,1)$ represents the upper bound of $f(\rho, \delta)$ (dotted lines). This general functional
116 relationship between surveillance effort and isolation effectiveness is assumed to hold for surveillance of
117 both nonsymptomatic and symptomatic individuals, although nonsymptomatic hosts are more challenging
118 to detect than symptomatic hosts ($\varepsilon < 1$).

119
120 In our model, the parameter β and its scaled counterparts $\alpha\beta$ and $\eta\beta$ represent the rates at which
121 symptomatic, presymptomatic and asymptomatic hosts generate new infections, respectively.
122 Since we are modelling the beginning of a potential local outbreak, we assume that the size of the
123 susceptible population remains approximately constant and do not track the depletion of this
124 population. The parameter λ governs the rate at which presymptomatic individuals develop
125 symptoms, so that the expected duration of the presymptomatic period is $1/\lambda$ days in the absence
126 of interventions. Similarly, without interventions, the expected durations of the symptomatic and
127 asymptomatic infectious periods are $1/\mu$ days and $1/\nu$ days, respectively.

128
129 The baseline rate at which symptomatic individuals are detected and isolated is determined by the
130 parameter γ . Assuming that nonsymptomatic individuals are more difficult to detect than
131 symptomatic individuals, we take the analogous quantity for nonsymptomatic hosts to be $\varepsilon\gamma$,
132 where the scaling factor $\varepsilon < 1$ reflects the fact that interventions targeting nonsymptomatic hosts
133 are likely to be less effective for the same surveillance effort. We assume that the sensitivity of
134 surveillance is identical for presymptomatic and asymptomatic individuals, and therefore use the
135 same isolation rate for both of these groups.

136
137 The parameters ρ_1 and ρ_2 represent the surveillance intensification effort targeted at
138 nonsymptomatic and symptomatic hosts, respectively. The function $f(\rho, \delta) = \frac{\delta\rho}{1+\rho}$ governs the
139 proportional reduction in the expected time to isolation for a given surveillance effort, ρ (for a
140 similar approach in which the proportion of infectious cases prevented is assumed to be a function
141 of control effort, see Matthews *et al.* [27]). The functional form of $f(\rho, \delta)$ is chosen for three main
142 reasons. First, it generates a reduced expected time to isolation when the surveillance effort
143 increases. Second, since the proportional reduction in the expected time to isolation is bounded
144 above by the parameter $\delta \in (0,1)$, the isolation rate saturates and cannot increase indefinitely.
145 Third, the gradient $\frac{\partial f}{\partial \rho}$ decreases with the surveillance effort ρ , meaning that an increase in the
146 surveillance effort has a greater impact at low surveillance efforts compared to when this effort is
147 already large [27]. The function $f(\rho, \delta)$ is shown in Fig 1B for different values of the parameter δ .
148

149 **2.2 Reproduction number**

150 The basic reproduction number, R_0 , represents the expected number of secondary infections
151 generated by a single infected individual introduced at the start of their infection into a fully
152 susceptible population in the absence of intensified surveillance:

$$153 \quad R_0 = \frac{\xi\eta\beta}{\nu + \epsilon\gamma} + (1 - \xi) \left[\frac{\alpha\beta}{\lambda + \epsilon\gamma} + \frac{\lambda}{\lambda + \epsilon\gamma} \frac{\beta}{\gamma + \mu} \right].$$

154 This expression is the sum of the expected number of transmissions from a host who begins in the
155 asymptomatic class and from a host who begins in the presymptomatic infectious class, weighted
156 by the respective probabilities ξ and $1 - \xi$ that determine the chance that the host experiences a
157 fully asymptomatic course of infection. The expected number of transmissions from a host who
158 begins in the presymptomatic infectious class comprises transmissions occurring during the

159 incubation period and transmissions occurring during the symptomatic period, accounting for the
 160 possibility that the host is isolated prior to developing symptoms.

161
 162 The proportion of infections arising from presymptomatic hosts in the absence of intensified
 163 surveillance is then given by

$$164 \quad K_p = \frac{\frac{(1-\xi)\alpha}{\lambda + \epsilon\gamma}}{\frac{\xi\eta}{\nu + \epsilon\gamma} + (1-\xi) \left[\frac{\alpha}{\lambda + \epsilon\gamma} + \frac{\lambda}{\lambda + \epsilon\gamma} \frac{1}{\gamma + \mu} \right]}, \quad (1)$$

165 and the equivalent quantity for asymptomatic hosts is given by

$$166 \quad K_a = \frac{\frac{\xi\eta}{\nu + \epsilon\gamma}}{\frac{\xi\eta}{\nu + \epsilon\gamma} + (1-\xi) \left[\frac{\alpha}{\lambda + \epsilon\gamma} + \frac{\lambda}{\lambda + \epsilon\gamma} \frac{1}{\gamma + \mu} \right]}. \quad (2)$$

167

168 **2.3 Baseline values of model parameters**

169 Since this research was motivated by the need to estimate outbreak risks outside China in the
 170 initial stages of the COVID-19 pandemic, we used a baseline set of parameter values in our
 171 analyses that was informed by studies conducted during this pandemic (Table 1). Where possible,
 172 these parameter values were obtained from existing literature. However, we also performed
 173 sensitivity analyses to determine how our results varied when the parameter values were changed
 174 (see Supplementary Text S3 and Supplementary Figs S3-12). In Table 1, and throughout, rounded
 175 values are given to three significant figures.

176

177 **Table 1. Parameters of the model and the values used in the baseline version of our analysis.**

Parameter	Meaning	Baseline value	Justification
R_0	Expected number of secondary infections caused by a single	$R_0 = 3$	Within estimated range for SARS-CoV-2 [28-31]

	infected individual (when $\rho_1 = \rho_2 = 0$)		
ξ	Proportion of infections that are asymptomatic	$\xi = 0.2$	[32-34]
β	Rate at which symptomatic individuals generate new infections	$\beta = 0.336$ days ⁻¹	Chosen so that $R_0 = 3$
α	Relative infectiousness of presymptomatic individuals compared to symptomatic individuals	$\alpha = 2.78$	Chosen so that 48.9% of transmissions arise from presymptomatic hosts (i.e. $K_p = 0.489$) [16]
η	Relative infectiousness of asymptomatic individuals compared to symptomatic individuals	$\eta = 0.519$	Chosen so that 10.6% of transmissions arise from asymptomatic hosts (i.e. $K_a = 0.106$) [16]
γ	Isolation rate of symptomatic individuals without intensified surveillance	$\gamma = 0.0924$ days ⁻¹	Chosen so that $\frac{1}{\gamma+\mu} = 4.6$ days [35]
ε	Relative isolation rate of nonsymptomatic individuals without intensified surveillance (compared to symptomatic individuals)	$\varepsilon = 0.1$	Assumed; chosen within the range $\varepsilon \in (0,1)$ (for different values, see Fig S7)
λ	Rate at which presymptomatic individuals develop symptoms	$\lambda = 0.5$ days ⁻¹	[20]
μ	Recovery rate of symptomatic individuals	$\mu = 1/8$ days ⁻¹	[36-38]
ν	Recovery rate of asymptomatic individuals	$\nu = 0.1$ days ⁻¹	Chosen so that, in the absence of interventions, the expected duration of infection is identical for all infected hosts $\left(\frac{1}{\nu} = \frac{1}{\lambda} + \frac{1}{\mu}\right)$
δ	Upper bound on the fractional reduction in the time to isolation	$\delta = 0.8$	Assumed; chosen within the natural range $\delta \in (0,1)$ (for different values, see Fig S11)
ρ_1	Surveillance intensification effort targeted at nonsymptomatic hosts	ρ_1 allowed to vary in the range [0,20]	N/A – range of values explored
ρ_2	Surveillance intensification effort targeted at symptomatic hosts	ρ_2 allowed to vary in the range [0,20]	N/A – range of values explored

178
179 The value of the parameter governing the baseline rate at which symptomatic individuals are
180 isolated, γ , was chosen to match empirical observations which indicate that individuals who seek

181 medical care prior to recovery or death do so around four to six days after symptom onset [35].
182 Specifically, we assumed that the period of time to the first medical visit could be used a proxy for
183 the time to isolation, and chose γ so that the expected time period to isolation conditional on
184 isolation occurring during the symptomatic period was given by $\frac{1}{\gamma+\mu} = 4.6$ days [35]. This is
185 different to the time period that we refer to as the expected time to isolation for symptomatic hosts,
186 which is $\frac{1}{\gamma}$ days (see Methods).

187

188 **2.4 Probability of a local outbreak**

189 For stochastic simulations of compartmental epidemiological models starting from a small number
190 of hosts infected initially, there are generally two qualitatively different types of behaviour. The
191 pathogen may fade out rapidly, or case numbers may begin to increase exponentially (only starting
192 to fade out once the number of susceptible individuals has been sufficiently depleted, unless public
193 health measures are introduced to reduce transmission). Consequently, running many simulations
194 of those types of model with R_0 larger than but not close to one, the epidemic size is distributed
195 bimodally, with the total number of individuals ever infected falling into one of two distinct ranges
196 (for a simple example, see Supplementary Fig S1A; see also [39-41]). In that scenario, a natural
197 definition for the probability of a local outbreak is therefore the proportion of outbreak simulations
198 for which the total number of infected individuals falls into the higher of these two ranges.

199

200 Here, since we are considering the initial phase of potential local outbreaks, we instead considered
201 a branching process model in which depletion of susceptibles was not accounted for. If simulations
202 of branching process models are run, then in each simulation the pathogen either fades out with
203 few infections or case numbers generally increase indefinitely. The probability of a local outbreak

204 starting from a small number of infected hosts then corresponds to the proportion of simulations in
 205 which the pathogen does not fade out quickly and case numbers increase indefinitely instead. This
 206 again provides a natural definition of a local outbreak, since simulations can be partitioned into
 207 two distinct sets (for an example in which simulations of a simple branching process model are
 208 used to calculate the probability of a local outbreak, see Supplementary Fig S1B).

209
 210 As an alternative to repeated simulation, we instead use our branching process model (Fig 1A) to
 211 perform analytic calculations of the probability that a single imported infectious host initiates a
 212 local outbreak. To do this, we denote the probability of a local outbreak not occurring, starting
 213 from i presymptomatic hosts, j symptomatic hosts, and k asymptomatic hosts, by $q_{i,j,k}$. Starting
 214 from one presymptomatic host (so that $i = 1$ and $j = k = 0$), there are four possibilities for the
 215 next event. That host could:

216 i) generate a new asymptomatic infection (with probability $\frac{\xi\alpha\beta}{\alpha\beta+\lambda+\frac{\varepsilon\gamma}{1-f(\rho_1,\delta)}}$);

217 ii) generate a new presymptomatic infection (with probability $\frac{(1-\xi)\alpha\beta}{\alpha\beta+\lambda+\frac{\varepsilon\gamma}{1-f(\rho_1,\delta)}}$);

218 iii) develop symptoms (with probability $\frac{\lambda}{\alpha\beta+\lambda+\frac{\varepsilon\gamma}{1-f(\rho_1,\delta)}}$), or;

219 iv) be isolated (with probability $\frac{\frac{\varepsilon\gamma}{1-f(\rho_1,\delta)}}{\alpha\beta+\lambda+\frac{\varepsilon\gamma}{1-f(\rho_1,\delta)}}$).

220 These probabilities are obtained by considering the rates at which different possible events occur
 221 in the branching process model. Presymptomatic hosts generate new infections at rate $\alpha\beta$, and
 222 these new infections occur in asymptomatic and presymptomatic hosts with probabilities ξ and
 223 $1 - \xi$, respectively. Therefore, starting from a single presymptomatic host, new asymptomatic
 224 infections occur at rate $\xi\alpha\beta$, whilst new presymptomatic infections occur at rate $(1 - \xi)\alpha\beta$.

225 Additionally, presymptomatic hosts develop symptoms at rate λ , and are isolated at rate $\frac{\epsilon\gamma}{1-f(\rho_1, \delta)}$.

226 The overall rate at which events occur is the sum of these individual event rates:

227
$$\text{Total event rate} = \alpha\beta + \lambda + \frac{\epsilon\gamma}{1-f(\rho_1, \delta)}.$$

228 For each of the four possible next events (i-iv, above), the probability that event occurs next is the
 229 individual rate at which that event occurs divided by the total event rate, leading to the expressions
 230 given.

231
 232 We use these probabilities to condition on the event that occurs next in the branching process,
 233 following the introduction of a single presymptomatic infectious individual into the population. If
 234 that event is the generation of a new asymptomatic infection, which occurs with probability

235 $\frac{\xi\alpha\beta}{\alpha\beta+\lambda+\frac{\epsilon\gamma}{1-f(\rho_1, \delta)}}$, the probability that a local outbreak subsequently does not occur is $q_{1,0,1}$. Applying

236 analogous reasoning to the other possible events, we obtain

237
$$q_{1,0,0} = \frac{\xi\alpha\beta}{\alpha\beta+\lambda+\frac{\epsilon\gamma}{1-f(\rho_1, \delta)}} q_{1,0,1} + \frac{(1-\xi)\alpha\beta}{\alpha\beta+\lambda+\frac{\epsilon\gamma}{1-f(\rho_1, \delta)}} q_{2,0,0} + \frac{\lambda}{\alpha\beta+\lambda+\frac{\epsilon\gamma}{1-f(\rho_1, \delta)}} q_{0,1,0} + \frac{\frac{\epsilon\gamma}{1-f(\rho_1, \delta)}}{\alpha\beta+\lambda+\frac{\epsilon\gamma}{1-f(\rho_1, \delta)}} q_{0,0,0}.$$

238
 239 If there are no infectious hosts present in the population (i.e. $i = j = k = 0$), then a local outbreak
 240 will not occur and so $q_{0,0,0} = 1$. Assuming that transmission chains arising from two infectious
 241 individuals are independent gives $q_{1,0,1} = q_{1,0,0} q_{0,0,1}$ and $q_{2,0,0} = q_{1,0,0}^2$. Hence,

242
$$q_{1,0,0} = a\xi q_{1,0,0} q_{0,0,1} + a(1-\xi)q_{1,0,0}^2 + bq_{0,1,0} + (1-a-b), \quad (3)$$

243 where $a = \frac{\alpha\beta}{\alpha\beta+\lambda+\frac{\epsilon\gamma}{1-f(\rho_1, \delta)}}$, $b = \frac{\lambda}{\alpha\beta+\lambda+\frac{\epsilon\gamma}{1-f(\rho_1, \delta)}}$.

244

245 Similarly, considering the probability of a local outbreak failing to occur starting from a single
 246 symptomatic host gives

$$247 \quad q_{0,1,0} = \frac{\xi\beta}{\beta + \frac{\gamma}{1-f(\rho_2,\delta)} + \mu} q_{0,1,1} + \frac{(1-\xi)\beta}{\beta + \frac{\gamma}{1-f(\rho_2,\delta)} + \mu} q_{1,1,0} + \frac{\frac{\gamma}{1-f(\rho_2,\delta)} + \mu}{\beta + \frac{\gamma}{1-f(\rho_2,\delta)} + \mu} q_{0,0,0}.$$

248 As before, noting that $q_{0,0,0} = 1$ and assuming that different infection lineages are independent
 249 leads to

$$250 \quad q_{0,1,0} = c\xi q_{0,1,0} q_{0,0,1} + c(1-\xi)q_{1,0,0}q_{0,1,0} + (1-c), \quad (4)$$

$$251 \quad \text{where } c = \frac{\beta}{\beta + \frac{\gamma}{1-f(\rho_2,\delta)} + \mu}.$$

252 Finally, considering the probability of a local outbreak failing to occur starting from a single
 253 asymptomatic host gives

$$254 \quad q_{0,0,1} = d\xi q_{0,0,1}^2 + d(1-\xi)q_{1,0,0}q_{0,0,1} + (1-d), \quad (5)$$

$$255 \quad \text{where } d = \frac{\eta\beta}{\eta\beta + \nu + \frac{\epsilon\gamma}{1-f(\rho_1,\delta)}}.$$

256

257 Equations (3), (4) and (5) may be combined to give a single quartic equation for $q_{0,0,1}$, yielding
 258 four sets of solutions for $q_{1,0,0}$, $q_{0,1,0}$ and $q_{0,0,1}$ (see Supplementary Text S1). It is straightforward
 259 to verify that $q_{1,0,0} = q_{0,1,0} = q_{0,0,1} = 1$ is always a solution, and further solutions can be found
 260 numerically. The appropriate solution to take is the minimal non-negative real solution $q_{1,0,0} =$
 261 $q_{1,0,0}^*$, $q_{0,1,0} = q_{0,1,0}^*$, $q_{0,0,1} = q_{0,0,1}^*$ (see Supplementary Text S1). Then, the probability of a local
 262 outbreak occurring beginning from a single presymptomatic host is given by

$$263 \quad p_{1,0,0} = 1 - q_{1,0,0}^*,$$

264 with equivalent expressions holding for $p_{0,1,0}$ and $p_{0,0,1}$ (the probability of a local outbreak
 265 occurring beginning from a single symptomatic host or a single asymptomatic host, respectively).

266

267 Throughout, we consider the probability p of a local outbreak starting from a single
268 nonsymptomatic host entering the population, accounting for the possibility that the
269 nonsymptomatic host is either presymptomatic or asymptomatic:

$$270 \quad p = (1 - \xi)p_{1,0,0} + \xi p_{0,0,1}.$$

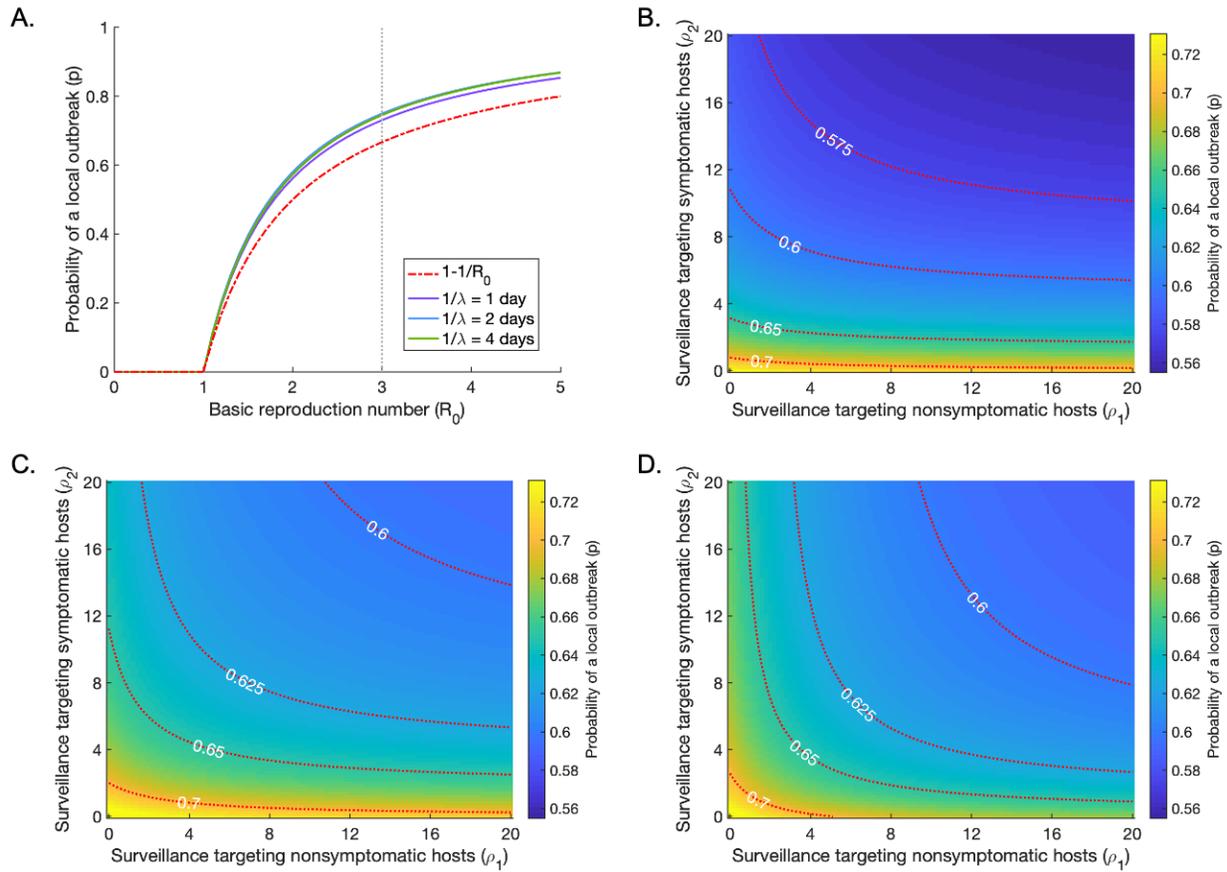
271

272 **3. Results**

273 **3.1 Probability of a local outbreak**

274 We considered the effect of R_0 and the duration of the presymptomatic and asymptomatic periods
275 on the probability of a local outbreak when a nonsymptomatic host enters a new host population
276 (Fig 2). We examined presymptomatic periods of length $1/\lambda = 1$ day, $1/\lambda = 2$ days and $1/\lambda = 4$
277 days; in each case, the duration of the asymptomatic period ($1/\nu$ days) was adjusted so that the
278 relative proportion of infections arising from asymptomatic hosts compared to presymptomatic
279 hosts remained fixed ($K_a/K_p = 0.218$, as in the baseline case). If instead nonsymptomatic
280 infections are not accounted for, the infectious period follows an exponential distribution and the
281 probability of a local outbreak is given by $p = 1 - 1/R_0$ (red dash-dotted line in Fig 2A).
282 Including nonsymptomatic infection in the model therefore led to an increased risk of a local
283 outbreak in the absence of surveillance intensification (Fig 2A).

284



285
 286 **Fig 2. The effect of the duration of the presymptomatic and asymptomatic periods on the probability**
 287 **of a local outbreak (p), starting from a single nonsymptomatic host. A.** The probability of a local
 288 outbreak as a function of the basic reproduction number R_0 , for presymptomatic periods of lengths $1/\lambda = 1$
 289 day (purple), $1/\lambda = 2$ days (blue) and $1/\lambda = 4$ days (green) in the absence of enhanced surveillance ($\rho_1 =$
 290 $\rho_2 = 0$). In each case, the duration of the asymptomatic period ($1/\nu$) is adjusted so that the relative
 291 proportion of infections arising from asymptomatic hosts compared to presymptomatic hosts remains
 292 constant ($K_a/K_p = 0.218$, as in the baseline case). The red dash-dotted line indicates the probability of a
 293 local outbreak in the absence of nonsymptomatic transmission. The vertical grey dotted line indicates $R_0 =$
 294 3, the baseline value used throughout. B. The probability of a local outbreak as a function of the surveillance
 295 intensification efforts ρ_1 and ρ_2 , for $1/\lambda = 1$ day. C. The analogous figure to B but with $1/\lambda = 2$ days. D.
 296 The analogous figure to B but with $1/\lambda = 4$ days. Red dotted lines indicate contours of constant local
 297 outbreak probability (i.e. lines on which the probability of a local outbreak takes the values shown). The

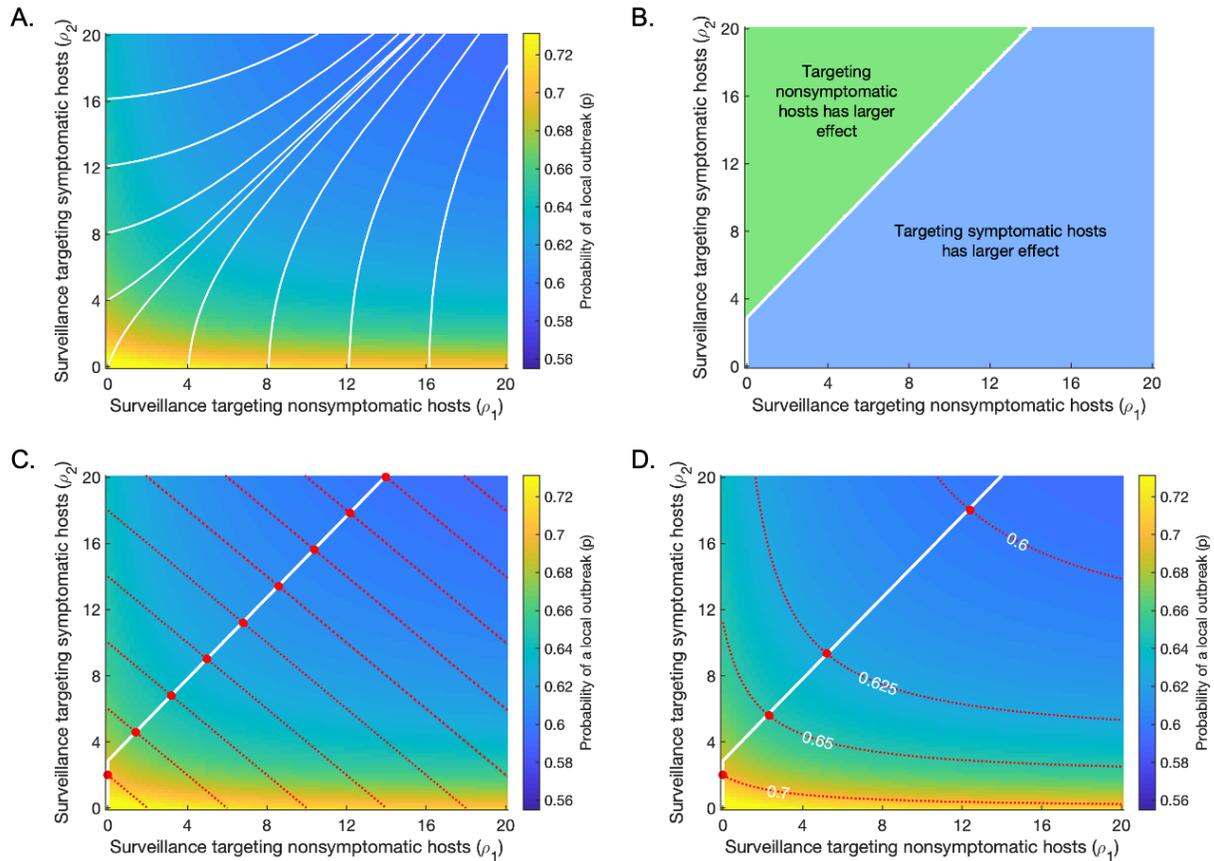
298 value of β is varied in each panel to fix $R_0 = 3$. All other parameter values are held fixed at the values in
299 Table 1 (except where stated).
300
301 We then considered the dependence of the probability of a local outbreak on the intensity of
302 surveillance targeting nonsymptomatic and symptomatic hosts (Fig 2B-D). The maximum value of
303 the surveillance intensification effort that we considered (given by ρ_1 or ρ_2 values of 20)
304 corresponded to a 76% reduction in the expected time to isolation (blue line in Fig 1B), i.e. a 76%
305 reduction in $\frac{1}{\varepsilon\gamma}$ or $\frac{1}{\gamma}$.
306
307 The length of the presymptomatic and asymptomatic periods significantly affected the dependence
308 of the probability of a local outbreak on the level of surveillance targeted at nonsymptomatic and
309 symptomatic hosts. In Fig 2B, in which the duration of the presymptomatic period was 1 day,
310 increasing surveillance targeted at nonsymptomatic hosts (ρ_1) had a limited effect on the
311 probability of a local outbreak, while increasing surveillance targeted at symptomatic hosts (ρ_2)
312 had a more significant effect. For example, increasing the surveillance effort targeted at
313 nonsymptomatic hosts to $\rho_1 = 5$ (a 67% reduction in the time to isolation) only reduced the
314 probability of a local outbreak from 0.730 to 0.716, whereas the equivalent effort targeted at
315 symptomatic hosts ($\rho_2 = 5$) reduced the probability to 0.630. As shown in Figs 3C and D,
316 however, when the presymptomatic and asymptomatic periods were longer, the benefit of directing
317 surveillance resources towards detecting nonsymptomatic individuals increased. This was because
318 longer presymptomatic and asymptomatic periods increased the proportion of infections generated
319 by nonsymptomatic individuals ($K_p + K_a$, see equations (1) and (2)); a presymptomatic period of 1
320 day, 2 days and 4 days corresponded to values of $K_p + K_a$ equal to 0.424, 0.595 and 0.746,
321 respectively.

322

323 **3.2 Optimising surveillance enhancement**

324 We next considered in more detail the impact of surveillance targeted at nonsymptomatic hosts
325 (ρ_1) relative to the impact of surveillance targeted at symptomatic hosts (ρ_2). For our baseline
326 parameter values, we considered the probability of a local outbreak starting from a single imported
327 nonsymptomatic individual for a range of values of ρ_1 and ρ_2 . We calculated the steepest descent
328 contours (white lines in Fig 3A) numerically using a gradient maximisation approach, in which at
329 each point the contour direction was determined by minimising the local outbreak probability over
330 a fixed search radius (see Supplementary Text S2 and Supplementary Fig S2). These contours
331 indicate how ρ_1 and ρ_2 should be altered to maximise the reduction in the probability of a local
332 outbreak. In this case, enhancing surveillance targeting both symptomatic and nonsymptomatic
333 hosts is always optimal (the steepest descent contours are neither horizontal nor vertical).

334



335
 336 **Fig 3. Optimal surveillance strategies to reduce the probability of a local outbreak (p) starting from a**
 337 **single nonsymptomatic host.** A. The local outbreak probability for different values of ρ_1 and ρ_2 , with the
 338 steepest descent contours overlaid (white lines). For the maximum reduction in the probability of a local
 339 outbreak at each point, surveillance must be enhanced for both nonsymptomatic and symptomatic
 340 individuals, with different levels of prioritisation depending on the current values of ρ_1 and ρ_2 . B. Values of
 341 ρ_1 and ρ_2 for which increasing surveillance for nonsymptomatic hosts (i.e. increasing ρ_1) is more effective
 342 at reducing the local outbreak probability than increasing surveillance for symptomatic hosts (i.e. increasing
 343 ρ_2) (green region) and vice versa (blue region). The white line represents the steepest descent contour
 344 starting from $\rho_1 = \rho_2 = 0$, under the constraint that surveillance can only be enhanced for either
 345 symptomatic or nonsymptomatic hosts at any time. The diagonal section of the steepest descent contour is
 346 made up of small horizontal and vertical sections. C. Strategies for minimising the local outbreak probability
 347 for a given fixed total surveillance effort ($\rho_1 + \rho_2 = C$). Red dotted lines indicate contours on which $\rho_1 +$
 348 ρ_2 is constant (i.e. lines on which $\rho_1 + \rho_2$ takes the values shown); red circles indicate the points along these

349 contours at which the local outbreak probability is minimised. The white line indicates the optimal
350 surveillance enhancement strategy if the maximum possible surveillance level (i.e. the maximum value of
351 $\rho_1 + \rho_2 = C$) is increased. D. Strategies for minimising the surveillance effort required to achieve a pre-
352 specific risk level (an “acceptable” local outbreak probability). Red dotted lines indicate contours of
353 constant local outbreak probability (i.e. lines on which the probability of a local outbreak takes the values
354 shown); red circles indicate the points along these contours at which the total surveillance effort $\rho_1 + \rho_2$ is
355 minimised. The white line indicates the optimal strategy to follow if the pre-specified risk level is increased
356 or reduced.

357

358 We then considered a scenario in which, at any time, it is only possible to direct resources towards
359 enhancing surveillance of either nonsymptomatic individuals or symptomatic individuals (e.g.
360 antigen testing of nonsymptomatic contacts of known infectious individuals, or screening for
361 symptomatic individuals at public events). In Fig 3B, the blue region represents values of ρ_1 and
362 ρ_2 for which enhancing surveillance targeting symptomatic hosts (i.e. increasing ρ_2) leads to a
363 larger reduction in the local outbreak probability than enhancing surveillance targeting
364 nonsymptomatic hosts (i.e. increasing ρ_1). In contrast, in the green region, enhancing surveillance
365 of nonsymptomatic individuals is more effective than enhancing surveillance of symptomatic
366 individuals. The white line represents the steepest descent contour starting from $\rho_1 = \rho_2 = 0$,
367 under the constraint that surveillance can only be enhanced for symptomatic or nonsymptomatic
368 hosts at any time.

369

370 Practical deployment of surveillance is often subject to logistical constraints, and policy-makers
371 may wish to design surveillance strategies to achieve a specific objective – for example, to
372 maximise the effectiveness of limited resources or to minimise the cost of achieving a desired
373 outcome. We therefore also considered the following two examples of such objectives.

374

375 **Objective 1: Minimise the probability of a local outbreak for a fixed total surveillance effort.**

376 First, we considered the question: given a fixed maximum surveillance effort ($\rho_1 + \rho_2 = C$), how
377 should surveillance be targeted at nonsymptomatic and symptomatic hosts? This involves setting
378 the values of ρ_1 and ρ_2 to minimise the local outbreak probability. The optimal strategies in this
379 case are shown in Fig 3C. The red dotted lines represent contours along which the total
380 surveillance effort $\rho_1 + \rho_2$ is held constant (i.e. different values of C). On each contour, the red
381 circle indicates the point at which the local outbreak probability is minimised.

382

383 If surveillance resources are increased (i.e. C increases), a further question is how surveillance
384 should then be increased. In Fig 3C, the white line represents the contour of steepest descent,
385 under the constraint that the total change in surveillance effort ($\rho_1 + \rho_2$) is held constant at each
386 step (rather than a constant search radius, as in Fig 3A – for more details, see Supplementary Text
387 S2 and Supplementary Fig S2). This contour coincides exactly with that shown in Fig 3B.

388

389 These results indicate that, if surveillance resources are such that C is greater than 2.8
390 (corresponding to a 59% reduction in time to isolation of symptomatic hosts), the optimal
391 surveillance strategy involves both enhanced surveillance of symptomatic individuals and
392 nonsymptomatic individuals (the red dots correspond to strictly positive values of both ρ_1 and ρ_2 ,
393 unless C is less than 2.8).

394

395 **Objective 2: Minimise the total surveillance effort to achieve a pre-specified reduction in the**

396 **probability of a local outbreak.** Second, we considered the question: given a pre-specified
397 acceptable risk level (i.e. probability of a local outbreak), how should the surveillance level

398 targeted at nonsymptomatic and symptomatic hosts be chosen? This involves choosing ρ_1 and ρ_2
399 to minimise $\rho_1 + \rho_2$ along a given contour corresponding to a fixed local outbreak probability (red
400 dotted lines in Fig 3D). On each contour, the red circle indicates the point along that contour at
401 which the total surveillance effort $\rho_1 + \rho_2$ is minimised. These optimal points also lie exactly
402 along the line on which enhancing surveillance targeted at symptomatic hosts is equally effective
403 compared to enhancing surveillance targeted at nonsymptomatic hosts.

404
405 As long as the target local outbreak probability is less than 0.69, optimal surveillance involves
406 enhanced surveillance of nonsymptomatic individuals as well as symptomatic individuals. For
407 example, in order to reduce the local outbreak probability to 0.6, the optimal approach is to deploy
408 resources such that $\rho_1 = 12.4$ (a 74% reduction in time to isolation of nonsymptomatic
409 individuals) and $\rho_2 = 18.0$ (a 76% reduction in time to isolation of symptomatic individuals).

410
411 Plots analogous to Fig 3D in which the parameters were varied from their baseline values are
412 shown in Supplementary Figs S3-12. In each case that we considered, our main finding was
413 unchanged. There always exists a threshold local outbreak probability such that, if the target local
414 outbreak probability is below this threshold, the optimal strategy for further reduction in the local
415 outbreak probability involves enhancing surveillance targeting both nonsymptomatic and
416 symptomatic individuals.

417

418 **4. Discussion**

419 A key component of infectious disease epidemic management is inferring the risk of outbreaks in
420 different locations [5-8, 11, 41, 42]. Surveillance and control strategies can be introduced to reduce
421 the risk that imported cases will lead to local outbreaks [12, 13, 43-46]. However, for a range of

422 pathogens, public health measures are hindered by nonsymptomatic infectious hosts who can
423 transmit the pathogen yet are challenging to detect [16, 42, 44, 47-49].

424
425 Here, we showed how the probability of a local outbreak can be estimated using a branching
426 process model that accounts for transmission from nonsymptomatic infected individuals (Fig 1).
427 The model can be used to assess the local outbreak probability for different surveillance strategies
428 that target nonsymptomatic or symptomatic hosts (Fig 2). Previous studies have shown that
429 detection of nonsymptomatic infections can be a key component of epidemic forecasting [42] and
430 containment [44], and have demonstrated the benefits of identifying and isolating infectious
431 nonsymptomatic hosts to reduce transmission [16, 17]. We focused instead on investigating how
432 surveillance should be targeted at nonsymptomatic or symptomatic hosts in order to reduce the
433 probability that cases imported to new locations will trigger a local outbreak (Fig 3A,B). We also
434 showed how the optimal surveillance level targeting these two groups can be assessed when
435 surveillance resources are limited and policy-makers have specific objectives (Fig 3C,D). In each
436 case, our main conclusion was that surveillance for nonsymptomatic infected hosts ($\rho_1 > 0$) can
437 be an important component of reducing the local outbreak risk during epidemics. This result has
438 broad implications, and our analysis could be extended to assess the potential for containing
439 outbreaks at their source using a range of specific interventions targeting symptomatic and
440 nonsymptomatic hosts.

441
442 Our goal here was to use the simplest possible model to explore the effects of surveillance of
443 nonsymptomatic and symptomatic individuals on the risk of local outbreaks. However, this model
444 is not without its limitations. One area of uncertainty is the precise values of the parameters
445 governing pathogen transmission and control. In this article, we chose a baseline set of parameter

446 values that is consistent with the findings of studies conducted during the COVID-19 pandemic,
447 although constructing a detailed transmission model for this pandemic was not our main focus. For
448 example, we set the relative rates at which presymptomatic and asymptomatic individuals generate
449 new infections compared to symptomatic individuals so that 48.9% of transmissions arise from
450 presymptomatic infectors, and 10.6% arise from asymptomatic infectors [16]. While this is in line
451 with reported estimates [50, 51], there is substantial variation between studies. Similarly, the
452 proportion of individuals who experience a fully asymptomatic course of infection (denoted by ξ
453 in our model) is subject to a considerable degree of uncertainty. Here, we chose $\xi = 0.2$ as the
454 baseline value [32-34] but estimates in the literature range from 0.04 to over 0.8 [33, 52-54]. We
455 therefore also conducted sensitivity analyses in which we explored a range of different values of
456 model parameters (Supplementary Text S3 and Supplementary Figs S3-12). In each case that we
457 considered, our main conclusion was unchanged: surveillance of nonsymptomatic individuals can
458 contribute to reducing the risk of local outbreaks. This result is expected to hold for epidemics of
459 any pathogen for which nonsymptomatic individuals contribute significantly to transmission.

460
461 For our modelling approach to be used to make precise quantitative predictions during epidemics,
462 it would be necessary to update the model to include the range of different specific surveillance
463 and control interventions that are in place. For example, detection of nonsymptomatic infected
464 individuals is facilitated by contact tracing and antigen testing, which are carried out routinely
465 during epidemics and can be included in models explicitly [12, 44, 55, 56]. Reductions in contacts
466 due to social distancing strategies and school or workplace closures could also be accounted for
467 [57, 58], although such interventions are often introduced after a local outbreak has begun rather
468 than in the initial phase of a potential local outbreak as considered here. We modelled the level of
469 surveillance targeted at nonsymptomatic and symptomatic hosts in a simple way, using a function

470 describing the relationship between surveillance effort and effectiveness (Fig 1B). We assumed
471 that this general functional relationship could be applied to interventions targeting both
472 symptomatic and nonsymptomatic hosts, accounting for logistical differences in the ease of
473 targeting either group by scaling the effectiveness of surveillance for nonsymptomatic hosts using
474 the parameter ε (results are shown for different values of ε in Supplementary Fig S8). In principle,
475 it would be possible to include entirely different functional forms describing the relationship
476 between surveillance effort and effectiveness for strategies targeting symptomatic and
477 nonsymptomatic individuals, and these could be tailored to the effects of particular interventions.
478 If different public health measures are included in the model explicitly, then it would be possible
479 to increase the accuracy of assessments of the relative public health benefits of specific
480 interventions that only target symptomatic individuals (e.g. screening for passengers with
481 heightened temperatures at airports [59, 60]) compared to interventions that also target
482 nonsymptomatic hosts (e.g. travel bans or quarantine of all inbound passengers [61, 62]). Of
483 course, this would require data from which the relative effectiveness of different measures could
484 be inferred.

485
486 The underlying transmission model could also be extended to include additional realism in several
487 ways. Transmission dynamics are influenced by marked heterogeneities in the patterns of contacts
488 between individuals in different age groups [63, 64], and, for COVID-19, susceptibility to
489 infection, the likelihood of developing symptoms, and the average severity of those symptoms
490 increase with age [65, 66]. Age-dependent variation in the proportion of asymptomatic cases in
491 particular implies that the optimal balance of surveillance between symptomatic and
492 nonsymptomatic hosts may differ between age groups. An age-structured version of the model
493 presented here is a focus of our ongoing research. Similarly, for a range of infectious diseases, the

494 distribution characterising the number of secondary infections generated by each infected host (the
495 offspring distribution) exhibits a high degree of overdispersion [67-70]. For a fixed value of R_0 , a
496 higher degree of overdispersion increases the likelihood that initial cases will fade out without
497 leading to a local outbreak [71, 72], and suggests that greater reductions in local outbreak risks
498 could theoretically be achieved for the same surveillance effort, if potential superspreaders or
499 superspreading events can be identified and targeted.

500
501 Despite the necessary simplifications, we have shown how the risk of local outbreaks can be
502 estimated during epidemics using a branching process model that includes nonsymptomatic
503 infectious hosts explicitly. Determining the extent to which nonsymptomatic individuals contribute
504 to transmission is essential early in emerging epidemics of a novel pathogen. As we have shown, if
505 transmissions occur from nonsymptomatic infectors, dedicating surveillance resources towards
506 finding nonsymptomatic cases can be an important component of public health measures that aim
507 to prevent local outbreaks.

508

509 **Acknowledgements**

510 Thank you to Andrew Wood and Charles Tolkien-Gillett for proofreading.

511

512 **Funding**

513 FALR acknowledges funding from the Biotechnology and Biological Sciences Research Council
514 (UKRI-BBSRC), grant number BB/M011224/1. SF and RNT acknowledge funding from the
515 Wellcome Trust, grant number 210758/Z/18/Z. RNT also acknowledges funding from Christ
516 Church (University of Oxford) via a Junior Research Fellowship. CAD acknowledges funding
517 from the MRC Centre for Global Infectious Disease Analysis (reference MR/R015600/1), jointly

518 funded by the UK Medical Research Council (MRC) and the UK Foreign, Commonwealth &
519 Development Office (FCDO) under the MRC/FCDO Concordat agreement, and is also part of the
520 EDCTP2 programme supported by the European Union.

521

522 **References**

523 [1] Bloom, D.E. & Cadarette, D. 2019 Infectious Disease Threats in the Twenty-First Century:
524 Strengthening the Global Response. *Front. Immunol.* **10**. (doi:10.3389/fimmu.2019.00549).

525 [2] Morens, D.M. & Fauci, A.S. 2013 Emerging Infectious Diseases: Threats to Human Health
526 and Global Stability. *Public Library of Science: Pathogens* **9**, e1003467.

527 (doi:10.1371/journal.ppat.1003467).

528 [3] Sigfrid, L., Maskell, K., Bannister, P.G., Ismail, S.A., Collinson, S., Regmi, S., Blackmore, C.,
529 Harriss, E., Longuere, K.-S., Gobat, N., et al. 2020 Addressing challenges for clinical research
530 responses to emerging epidemics and pandemics: a scoping review. *BMC Med.* **18**, 190-190.

531 (doi:10.1186/s12916-020-01624-8).

532 [4] Thompson, R.N. & Brooks-Pollock, E. 2019 Detection, forecasting and control of infectious
533 disease epidemics: modelling outbreaks in humans, animals and plants. *Philos. Trans. R. Soc.*

534 *Lond., B, Biol. Sci.* **374**, 20190038-20190038. (doi:10.1098/rstb.2019.0038).

535 [5] Althaus, C.L., Low, N., Musa, E.O., Shuaib, F. & Gsteiger, S. 2015 Ebola virus disease
536 outbreak in Nigeria: Transmission dynamics and rapid control. *Epidemics* **11**, 80-84.

537 (doi:10.1016/j.epidem.2015.03.001).

538 [6] Boldog, P., Tekeli, T., Vizi, Z., Denes, A., Bartha, F.A. & Rost, G. 2020 Risk Assessment of
539 Novel Coronavirus COVID-19 Outbreaks Outside China. *J. Clin. Med.* **9**, 571.

540 (doi:doi.org/10.3390/jcm9020571).

- 541 [7] Daon, Y., Thompson, R.N. & Obolski, U. 2020 Estimating COVID-19 outbreak risk through
542 air travel. *J. Travel Med.* **27**. (doi:10.1093/jtm/taaa093).
- 543 [8] Lashari, A.A. & Trapman, P. 2018 Branching process approach for epidemics in dynamic
544 partnership network. *J. Math. Biol.* **76**, 265-294. (doi:10.1007/s00285-017-1147-0).
- 545 [9] Nishiura, H., Cook, A.R. & Cowling, B.J. 2011 Assortativity and the Probability of Epidemic
546 Extinction: A Case Study of Pandemic Influenza A (H1N1-2009). *Interdiscip. Perspect. Infect.*
547 *Dis.* **2011**, 194507-194507. (doi:10.1155/2011/194507).
- 548 [10] Thompson, R.N., Jalava, K. & Obolski, U. 2019 Sustained transmission of Ebola in new
549 locations: more likely than previously thought. *The Lancet Infectious Diseases* **19**, 1058-1059.
550 (doi:10.1016/S1473-3099(19)30483-9).
- 551 [11] Thompson, R.N., Thompson, C.P., Pelerman, O., Gupta, S. & Obolski, U. 2019 Increased
552 frequency of travel in the presence of cross-immunity may act to decrease the chance of a global
553 pandemic. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* **374**, 20180274. (doi:10.1098/rstb.2018.0274).
- 554 [12] Hellewell, J., Abbott, S., Gimma, A., Bosse, N.I., Jarvis, C.I., Russell, T.W., Munday, J.D.,
555 Kucharski, A.J., Edmunds, W.J., Sun, F., et al. 2020 Feasibility of controlling COVID-19
556 outbreaks by isolation of cases and contacts. *The Lancet Global Health* **8**, e488-e496.
557 (doi:10.1016/S2214-109X(20)30074-7).
- 558 [13] Thompson, R.N. 2020 Novel Coronavirus Outbreak in Wuhan, China, 2020: Intense
559 Surveillance Is Vital for Preventing Sustained Transmission in New Locations. *Journal of Clinical*
560 *Medicine* **9**, 498. (doi:10.3390/jcm9020498).
- 561 [14] Lauer, S.A., Grantz, K.H., Bi, Q., Jones, F.K., Zheng, Q., Meredith, H.R., Azman, A.S.,
562 Reich, N.G. & Lessler, J. 2020 The Incubation Period of Coronavirus Disease 2019 (COVID-19)
563 From Publicly Reported Confirmed Cases: Estimation and Application. *Ann. Intern. Med.* **172**,
564 577-582. (doi:10.7326/M20-0504).

565 [15] McAloon, C., Collins, A., Hunt, K., Barber, A., Byrne, A.W., Butler, F., Casey, M., Griffin,
566 J., Lane, E., McEvoy, D., et al. 2020 Incubation period of COVID-19: a rapid systematic review
567 and meta-analysis of observational research. *BMJ Open* **10**, e039652. (doi:10.1136/bmjopen-2020-
568 039652).

569 [16] Ferretti, L., Wymant, C., Kendall, M., Zhao, L., Nurtay, A., Abeler-Dorner, L., Parker, M.,
570 Bonsall, D. & Fraser, C. 2020 Quantifying SARS-CoV-2 transmission suggests epidemic control
571 with digital contact tracing. *Science* **368**, eabb6936. (doi:10.1126/science.abb6936).

572 [17] Moghadas, S.M., Fitzpatrick, M.C., Sah, P., Pandey, A., Shoukat, A., Singer, B.H. & Galvani,
573 A. 2020 The implications of silent transmission for the control of COVID-19 outbreaks. *Proc.*
574 *Natl. Acad. Sci. U. S. A.* **117**, 17513-17515. (doi:10.1073/pnas.2008373117).

575 [18] Tindale, L.C., Stockdale, J.E., Coombe, M., Garlock, E.S., Lau, W.Y.V., Saraswat, M.,
576 Zhang, L., Chen, D., Wallinga, J. & Colijn, C. 2020 Evidence for transmission of COVID-19 prior
577 to symptom onset. *eLife* **9**, e57149. (doi:10.7554/eLife.57149).

578 [19] Tong, Z.-D., Tang, A., Li, K.-F., Li, P., Wang, H.-L., Yi, J.-P., Zhang, Y.-L. & Yan, J.-B.
579 2020 Potential Presymptomatic Transmission of SARS-CoV-2, Zhejiang Province, China, 2020.
580 *Emerging Infectious Diseases* **26**, 1052-1054. (doi:10.3201/eid2605.200198).

581 [20] Wei, W.E., Li, Z., Chiew, C.J., Yong, S.E., Toh, M.P. & Lee, V.J. 2020 Presymptomatic
582 Transmission of SARS-CoV-2 — Singapore, January 23–March 16, 2020. *Morbidity and*
583 *Mortality Weekly Report* **69**, 411-415. (doi:10.15585/mmwr.mm6914e1external icon).

584 [21] Bai, Y., Yao, L., Wei, T., Tian, F., Jin, D.-Y., Chen, L. & Wang, M. 2020 Presumed
585 Asymptomatic Carrier Transmission of COVID-19. *J. Am. Med. Assoc.* **323**, 1406-1407.
586 (doi:10.1001/jama.2020.2565).

- 587 [22] Furukawa, N.W., Brooks, J.T. & Sobel, J. 2020 Evidence Supporting Transmission of Severe
588 Acute Respiratory Syndrome Coronavirus 2 While Presymptomatic or Asymptomatic. *Emerging*
589 *Infectious Diseases* **26**, e201595. (doi:10.3201/eid2607.201595).
- 590 [23] Anderson, D. & Watson, R. 1980 On the spread of a disease with gamma distributed latent
591 and infectious periods. *Biometrika* **67**, 191-198. (doi:10.1093/biomet/67.1.191).
- 592 [24] Bartlett, M.S. 1964 The Relevance of Stochastic Models for Large-Scale Epidemiological
593 Phenomena. *J. Roy. Stat. Soc. Ser. C. (Appl. Stat.)* **13**, 2-8. (doi:10.2307/2985217).
- 594 [25] Craft, M.E., Beyer, H.L. & Haydon, D.T. 2013 Estimating the Probability of a Major
595 Outbreak from the Timing of Early Cases: An Indeterminate Problem? *PLoS One* **8**, e57878-
596 e57878. (doi:10.1371/journal.pone.0057878).
- 597 [26] Lloyd, A.L., Zhang, J. & Root, A.M. 2007 Stochasticity and heterogeneity in host-vector
598 models. *J. R. Soc. Interface* **4**, 851-863. (doi:10.1098/rsif.2007.1064).
- 599 [27] Matthews, L., Haydon, D.T., Shaw, D.J., Chase-Topping, M.E., Keeling, M.J. & Woolhouse,
600 M.E.J. 2003 Neighbourhood control policies and the spread of infectious diseases. *Proceedings of*
601 *the Royal Society B: Biological Sciences* **270**, 1659-1666. (doi:10.1098/rspb.2003.2429).
- 602 [28] Lai, C.-C., Shih, T.-P., Ko, W.-C., Tang, H.-J. & Hsueh, P.-R. 2020 Severe acute respiratory
603 syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): The
604 epidemic and the challenges. *International Journal of Antimicrobial Agents* **55**, 105924.
605 (doi:10.1016/j.ijantimicag.2020.105924).
- 606 [29] Liu, Y., Gayle, A.A., Wilder-Smith, A. & Rocklöv, J. 2020 The reproductive number of
607 COVID-19 is higher compared to SARS coronavirus. *Journal of Travel Medicine* **27**, taaa021.
608 (doi:10.1093/jtm/taaa021).

- 609 [30] Zhai, P., Ding, Y., Wu, X., Long, J., Zhong, Y. & Li, Y. 2020 The epidemiology, diagnosis
610 and treatment of COVID-19. *International Journal of Antimicrobial Agents* **55**, 105955-105955.
611 (doi:10.1016/j.ijantimicag.2020.105955).
- 612 [31] Zhao, S., Lin, Q., Ran, J., Musa, S.S., Yang, G., Wang, W., Lou, Y., Gao, D., Yang, L., He,
613 D., et al. 2020 Preliminary estimation of the basic reproduction number of novel coronavirus
614 (2019-nCoV) in China, from 2019 to 2020: A data-driven analysis in the early phase of the
615 outbreak. *International Journal of Infectious Diseases* **92**, 214-217.
616 (doi:10.1016/j.ijid.2020.01.050).
- 617 [32] Buitrago-Garcia, D., Egli-Gany, D., Counotte, M.J., Hossmann, S., Imeri, H., Ipekci, A.M.,
618 Salanti, G. & Low, N. 2020 Occurrence and transmission potential of asymptomatic and
619 presymptomatic SARS-CoV-2 infections: A living systematic review and meta-analysis. *PLoS*
620 *Med.* **17**, e1003346. (doi:10.1371/journal.pmed.1003346).
- 621 [33] Byambasuren, O., Cardona, M., Bell, K., Clark, J., McLaws, M.-L. & Glasziou, P. 2020
622 Estimating the extend of asymptomatic COVID-19 and its potential for community transmission:
623 Systematic review and meta-analysis. *JAMMI* **5**, 223-234. (doi:10.3138/jammi-2020-0030).
- 624 [34] He, J., Guo, Y., Mao, R. & Zhang, J. 2020 Proportion of asymptomatic coronavirus disease
625 2019: A systematic review and meta-analysis. *J. Med. Virol.* **93**, 820-830.
626 (doi:10.1002/jmv.26326).
- 627 [35] Li, Q., Guan, X., Wu, P., Wang, X., Zhou, L., Tong, Y., Ren, R., Leung, K.S.M., Lau,
628 E.H.Y., Wong, J.Y., et al. 2020 Early Transmission Dynamics in Wuhan, China, of Novel
629 Coronavirus–Infected Pneumonia. *N. Engl. J. Med.* **382**, 1199-1207.
630 (doi:10.1056/NEJMoa2001316).
- 631 [36] Arons, M.M., Hatfield, K.M., Reddy, S.C., Kimball, A., James, A., Jacobs, J.R., Taylor, J.,
632 Spicer, K., Bardossy, A.C., Oakley, L.P., et al. 2020 Presymptomatic SARS-CoV-2 Infections and

633 Transmission in a Skilled Nursing Facility. *The New England Journal of Medicine* **382**, 2081-
634 2090. (doi:10.1056/NEJMoa2008457).

635 [37] Bullard, J., Dust, K., Funk, D., Strong, J.E., Alexander, D., Garnett, L., Boodman, C., Bello,
636 A., Hedley, A., Schiffman, Z., et al. 2020 Predicting Infectious Severe Acute Respiratory
637 Syndrome Coronavirus 2 From Diagnostic Samples. *Clinical Infectious Diseases*, ciaa638.
638 (doi:10.1093/cid/ciaa638).

639 [38] Wolfel, R., Corman, V.M., Guggemos, W., Seilmaier, M., Zange, S., Muller, M.A.,
640 Niemeyer, D., Jones, T.C., Vollmar, P., Rothe, C., et al. 2020 Virological assessment of
641 hospitalized patients with COVID-2019. *Nature* **581**, 465-469. (doi:10.1038/s41586-020-2196-x).

642 [39] Nasell, I. 1995 The Threshold Concept in Stochastic Epidemic and Endemic Models. In
643 *Epidemic Models: Their Structure and Relation to Data* (ed. D. Mollison), pp. 71-83. Cambridge,
644 UK, Cambridge University Press.

645 [40] Britton, T. 2010 Stochastic epidemic models: A survey. *Math. Biosci.* **225**, 24-35.
646 (doi:10.1016/j.mbs.2010.01.006).

647 [41] Thompson, R.N., Gilligan, C.A. & Cunniffe, N.J. 2020 Will an outbreak exceed available
648 resources for control? Estimating the risk from invading pathogens using practical definitions of a
649 severe epidemic. *J. R. Soc. Interface* **17**. (doi:10.1098/rsif.2020.0690).

650 [42] Thompson, R.N., Gilligan, C.A. & Cunniffe, N.J. 2016 Detecting Presymptomatic Infection
651 Is Necessary to Forecast Major Epidemics in the Earliest Stages of Infectious Disease Outbreaks.
652 *PLoS Comput. Biol.* **12**, e1004836. (doi:10.1371/journal.pcbi.1004836).

653 [43] Charleston, B., Bankowski, B.M., Gubbins, S., Chase-Topping, M.E., Schley, D., Howey, R.,
654 Barnett, P.V., Gibson, D., Juleff, N.D. & Woolhouse, M.E.J. 2011 Relationship Between Clinical
655 Signs and Transmission of an Infectious Disease and the Implications for Control. *Science* **332**,
656 726-729. (doi:10.1126/science.1199884).

657 [44] Fraser, C., Riley, S., Anderson, R.M. & Ferguson, N.M. 2004 Factors that make an infectious
658 disease outbreak controllable. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 6146-6151.
659 (doi:10.1073/pnas.0307506101).

660 [45] Martinez, L. 2000 Global infectious disease surveillance. *Int. J. Infect. Dis.* **4**, 222-228.
661 (doi:10.1016/s1201-9712(00)90114-0).

662 [46] Milinovich, G.J., Williams, G.M., Clements, A.C.A. & Hu, W. 2014 Internet-based
663 surveillance systems for monitoring emerging infectious diseases. *The Lancet Infectious Diseases*
664 **14**, 160-168. (doi:10.1016/S1473-3099(13)70244-5).

665 [47] Gu, Y., Komiya, N., Kamiya, H., Yasui, Y., Taniguchi, K. & Okabe, N. 2011 Pandemic
666 (H1N1) 2009 Transmission during Presymptomatic Phase, Japan. *Emerging Infectious Diseases*
667 **17**, 1737-1739. (doi:10.3201/eid1709.101411).

668 [48] Ip, D.K.M., Lau, L.L.H., Leung, N.H.L., Fang, V.J., Chan, K.-H., Chu, D.K.W., Leung,
669 G.M., Peiris, J.S.M., Uyeki, T.M. & Cowling, B.J. 2016 Viral Shedding and Transmission
670 Potential of Asymptomatic and Paucisymptomatic Influenza Virus Infections in the Community.
671 *Clin. Infect. Dis.* **64**, 736-742. (doi:10.1093/cid/ciw841).

672 [49] ten Bosch, Q.A., Clapham, H.E., Lambrechts, L., Duong, V., Buchy, P., Althouse, B.M.,
673 Lloyd, A.L., Waller, L.A., Morrison, A.C., Kitron, U., et al. 2018 Contributions from the silent
674 majority dominate dengue virus transmission. *PLoS Pathog.* **14**, e1006965-e1006965.
675 (doi:10.1371/journal.ppat.1006965).

676 [50] He, X., Lau, E.H.Y., Wu, P., Deng, X., Wang, J., Hao, X., Lau, Y.C., Wong, J.Y., Guan, Y.,
677 Tan, X., et al. 2020 Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nat.*
678 *Med.* **26**, 672-675. (doi:10.1038/s41591-020-0869-5).

679 [51] Ren, X., Li, Y., Yang, X., Li, Z., Cui, J., Zhu, A., Zhao, H., Yu, J., Nie, T., Ren, M., et al.
680 2020 Evidence for pre-symptomatic transmission of coronavirus disease 2019 (COVID-19) in
681 China. *Influenza Other Respi. Viruses*, 1-9. (doi:10.1111/irv.12787).

682 [52] Yanes-Lane, M., Winters, N., Fregonese, F., Bastos, M., Perlman-Arrow, S., Campbell, J.R.
683 & Menzies, D. 2020 Proportion of asymptomatic infection among COVID-19 positive persons and
684 their transmission potential: A systematic review and meta-analysis. *PLoS One* **15**.
685 (doi:10.1371/journal.pone.0241536).

686 [53] Meyerowitz, E.A., Richterman, A., Bogoch, I.I., Low, N. & Cevik, M. 2020 Towards an
687 accurate and systematic characterisation of persistently asymptomatic infection with SARS-CoV-
688 2. *The Lancet Infectious Diseases*. (doi:10.1016/S1473-3099(20)30837-9).

689 [54] Oran, D.P. & Topol, E.J. 2020 Prevalence of Asymptomatic SARS-CoV-2 Infection: A
690 Narrative Review. *Ann. Intern. Med.* **173**, 362-367. (doi:10.7326/M20-3012).

691 [55] Klinkenberg, D., Fraser, C. & Heesterbeek, H. 2006 The Effectiveness of Contact Tracing in
692 Emerging Epidemics. *PLoS One* **1**, e12-e12. (doi:10.1371/journal.pone.0000012).

693 [56] Webb, G., Browne, C., Huo, X., Seydi, O., Seydi, M. & Magal, P. 2015 A Model of the 2014
694 Ebola Epidemic in West Africa with Contact Tracing. *PLoS Curr.* **7**.
695 (doi:10.1371/currents.outbreaks.846b2a31ef37018b7d1126a9c8adf22a).

696 [57] Davies, N.G., Kucharski, A.J., Eggo, R.M., Gimma, A., Edmunds, W.J., Jombart, T.,
697 O'Reilly, K., Endo, A., Hellewell, J., Nightingale, E.S., et al. 2020 Effects of non-pharmaceutical
698 interventions on COVID-19 cases, deaths, and demand for hospital services in the UK: a
699 modelling study. *The Lancet Public Health* **5**, E375-E385. (doi:10.1016/S2468-2667(20)30133-
700 X).

701 [58] Thompson, R.N. 2020 Epidemiological models are important tools for guiding COVID-19
702 interventions. *BMC Med.* **18**. (doi:10.1186/s12916-020-01628-4).

703 [59] Nishiura, H. & Kamiya, K. 2011 Fever screening during the influenza (H1N1-2009)
704 pandemic at Narita International Airport, Japan. *BMC Infect. Dis.* **11**, 111-111. (doi:10.1186/1471-
705 2334-11-111).

706 [60] Quilty, B.J., Clifford, S., group, C.n.w., Flasche, S. & Eggo, R.M. 2020 Effectiveness of
707 airport screening at detecting travellers infected with novel coronavirus (2019-nCoV).
708 *Eurosurveillance* **25**, 2000080-2000080. (doi:10.2807/1560-7917.ES.2020.25.5.2000080).

709 [61] Chinazzi, M., Davis, J.T., Ajelli, M., Gioannini, C., Litvinova, M., Merler, S., y Piontti, A.P.,
710 Mu, K., Rossi, L., Sun, K., et al. 2020 The effect of travel restrictions on the spread of the 2019
711 novel coronavirus (COVID-19) outbreak. *Science* **368**, 395-400. (doi:10.1126/science.aba9757).

712 [62] Dickens, B.L., Koo, J.R., Lim, J.T., Sun, H., Clapham, H.E., Wilder-Smith, A. & Cook, A.R.
713 2020 Strategies at points of entry to reduce importation risk of COVID-19 cases and reopen travel.
714 *J. Travel Med.*, taaa141-taaa141. (doi:10.1093/jtm/taaa141).

715 [63] Prem, K., Cook, A.R. & Jit, M. 2017 Projecting social contact matrices in 152 countries using
716 contact surveys and demographic data. *PLOS Computational Biology* **13**, e1005697.
717 (doi:10.1371/journal.pcbi.1005697).

718 [64] Zhang, J., Litvinova, M., Liang, Y., Wang, Y., Wang, W., Zhao, S., Wu, Q., Merler, S.,
719 Viboud, C., Vespignani, A., et al. 2020 Changes in contact patterns shape the dynamics of the
720 COVID-19 outbreak in China. *Science* **368**, 1481-1486. (doi:10.1126/science.abb8001).

721 [65] Kang, S.-J. & Jung, S.I. 2020 Age-Related Morbidity and Mortality among Patients with
722 COVID-19. *J. Infect. Chemother.* **52**, 154-164. (doi:10.3947/ic.2020.52.2.154).

723 [66] Davies, N.G., Klepac, P., Liu, Y., Prem, K., Jit, M., Pearson, C.A.B., Quilty, B.J., Kucharski,
724 A.J., Gibbs, H., Clifford, S., et al. 2020 Age-dependent effects in the transmission and control of
725 COVID-19 epidemics. *Nature Medicine* **26**, 1205-1211. (doi:10.1038/s41591-020-0962-9).

726 [67] Endo, A., Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working
727 Group, Abbott, S., Kucharski, A. & Funk, S. 2020 Estimating the overdispersion in COVID-19
728 transmission using outbreak sizes outside China [version 3; peer review: 2 approved]. *Wellcome*
729 *Open Research* **5**. (doi:10.12688/wellcomeopenres.15842.3).

730 [68] Lau, M.S.Y., Grenfell, B., Thomas, M., Bryan, M., Nelson, K. & Lopman, B. 2020
731 Characterizing superspreading events and age-specific infectiousness of SARS-CoV-2
732 transmission in Georgia, USA. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 22430-22435.
733 (doi:10.1073/pnas.2011802117).

734 [69] Woolhouse, M.E.J., Dye, C., Etard, J.-F., Smith, T., Charlwood, J.D., Garnett, G.P., Hagan,
735 P., Hii, J.L.K., Ndhlovu, P.D., Quinell, R.J., et al. 1997 Heterogeneities in the transmission of
736 infectious agents: Implications for the design of control programs. *Proceedings of the National*
737 *Academy of Sciences* **94**, 338-342. (doi:10.1073/pnas.94.1.338).

738 [70] Thompson, R.N., Hollingsworth, T.D., Isham, V., Arribas-Bel, D., Ashby, B., Britton, T.,
739 Challenor, P., Chappell, L.H.K., Clapham, H., Cunniffe, N.J., et al. 2020 Key questions for
740 modelling COVID-19 exit strategies. *Proceedings of the Royal Society B: Biological Sciences* **287**,
741 20201405. (doi:10.1098/rspb.2020.1405).

742 [71] Lloyd-Smith, J.O., Schreiber, S.J., Kopp, P.E. & Getz, W.M. 2005 Superspreading and the
743 effect of individual variation on disease emergence. *Nature* **438**, 355-359.
744 (doi:10.1038/nature04153).

745 [72] Lipsitch, M., Cohen, T., Cooper, B., Robins, J.M., Ma, S., James, L., Gopalakrishna, G.,
746 Chew, S.K., Tan, C.C., Samore, M.H., et al. 2003 Transmission Dynamics and Control of Severe
747 Acute Respiratory Syndrome. *Science* **300**, 1966-1970. (doi:10.1126/science.1086616).

748 [73] Norris, J.R. 1997 *Markov Chains (Cambridge Series in Statistical and Probabilistic*
749 *Mathematics)*. Cambridge, Cambridge University Press.

750

751 **Supplementary Text**

752

753 **Text S1. Probability of a local outbreak**

754 In the Methods section of the main text, we outlined an approach for deriving the probability of a local outbreak
755 starting from a single infectious host in either the presymptomatic, symptomatic or asymptomatic classes. Here,
756 we provide more details about that derivation. The probability of a local outbreak not occurring, starting from i
757 presymptomatic hosts, j symptomatic hosts and k asymptomatic hosts, is denoted by $q_{i,j,k}$. If we consider the
758 temporal evolution of (i, j, k) to be a Markov process on the state space $M = \mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0}$, then $q_{i,j,k}$ is the
759 hitting probability of the state $(0,0,0)$ starting from the state (i, j, k) . The vector of hitting probabilities $q^* =$
760 $\{q_{i,j,k}^* \mid (i, j, k) \in M\}$ is therefore the minimal non-negative (real) solution to the following system of equations:

761 $q_{0,0,0} = 1,$

762 $q_{i,j,k} = \sum_{(l,m,n) \in M} p_{(i,j,k),(l,m,n)} q_{l,m,n}$ for $(i, j, k) \neq (0,0,0),$

763 where $p_{(i,j,k),(l,m,n)}$ is the transition probability from state (i, j, k) to state (l, m, n) [73].

764 Here, minimality means that if $\hat{q} = \{\hat{q}_{i,j,k} \mid (i, j, k) \in M\}$ is another non-negative real solution, then $q_{i,j,k}^* \leq$
765 $\hat{q}_{i,j,k}$ for all $(i, j, k) \in M$.

766

767 From this, equations (3), (4) and (5) in the main text are obtained. These equations may be reduced to the
768 following quartic equation for $q_{0,0,1}$:

769
$$\omega_4 q_{0,0,1}^4 + \omega_3 q_{0,0,1}^3 + \omega_2 q_{0,0,1}^2 + \omega_1 q_{0,0,1} + \omega_0 = 0,$$

770 where

771 $\omega_4 = d(a - d)(d - c)\xi;$

772 $\omega_3 = cd(a - d)(1 - d)\xi - bd^3(1 - c)(1 - \xi) + (d - c)[d - a - ad\xi + d^2(a - 1 + b(1 - \xi) + \xi)];$

773 $\omega_2 = c(1 - d)[d - a - ad\xi + d^2(a - 1 + b(1 - \xi) + \xi)] + (d - c)[d(d - 2a - 1) + 2a];$

774 $\omega_1 = c(1 - d)[d(d - 2a - 1) + 2a] - a(1 - d)^2(d - c);$

775 $\omega_0 = -ac(1 - d)^3.$

776 The parameters a, b, c, d and ξ are as defined in the main text.

777

778 This yields four solutions for $q_{0,0,1}$ and four corresponding solutions for each of $q_{1,0,0}$ and $q_{0,1,0}$. One solution is
779 always given by $q_{1,0,0} = q_{0,1,0} = q_{0,0,1} = 1$; the other solutions may be found numerically. As described above,
780 we take the minimal non-negative real solution $q_{1,0,0} = q_{1,0,0}^*, q_{0,1,0} = q_{0,1,0}^*, q_{0,0,1} = q_{0,0,1}^*$, and observe that the
781 probability of a local outbreak occurring starting from i presymptomatic hosts, j symptomatic hosts and k
782 asymptomatic hosts is simply $1 - q_{i,j,k}$, giving the result stated in the main text.

783

784 If required, this result can be confirmed for specific model parameter values via repeated stochastic simulation of
785 the branching process model, starting from i presymptomatic hosts, j symptomatic hosts and k asymptomatic
786 hosts. As described in the main text, in simulations of branching process models, initial cases typically either
787 fade out or go on to cause a local outbreak. There is a natural definition of a local outbreak as a simulation in
788 which a large number of infections occur (see Supplementary Fig S1B). The probability of a local outbreak then
789 corresponds to the proportion of simulations in which large numbers of infections occur.

790

791 We note here that although this is a natural way to define a local outbreak, alternative definitions exist that may
792 be more appropriate in particular contexts. This is discussed by Thompson *et al.* (reference [41] in the main
793 text), who consider three practically relevant definitions of an outbreak based on different criteria for measuring
794 severity.

795

796 **Text S2. Computation of steepest descent contours**

797 The steepest descent contours shown in Fig 3A of the main text were computed using a gradient maximisation
798 approach, in which at each point the contour direction was determined by minimising the local outbreak
799 probability over a fixed search radius (Fig S2 A,B). Starting from ρ_1, ρ_2 , at each step we considered increasing
800 ρ_1 by an amount $\Delta\rho_1$ and increasing ρ_2 by an amount $\Delta\rho_2$ subject to the constraint $(\Delta\rho_1)^2 + (\Delta\rho_2)^2 = r^2$,
801 where r is a small pre-specified constant. To achieve this, we scanned over the circular arc $\Delta\rho_1 =$
802 $r \cos \theta, \Delta\rho_2 = r \sin \theta$, for $\theta \in [0, \pi/2]$ (Fig S2 A). In practice, this range was discretised into 33 search

803 directions evenly spaced between 0 and $\pi/2$. We then selected the pair of $\Delta\rho_1, \Delta\rho_2$ values for which the local
804 outbreak probability evaluated at $\rho_1 + \Delta\rho_1, \rho_2 + \Delta\rho_2$ was minimised. The process was then repeated beginning
805 from $\rho_{1_{new}} = \rho_1 + \Delta\rho_1, \rho_{2_{new}} = \rho_2 + \Delta\rho_2$ (Fig S2 B). The white line in Fig 3B, which divides the region in
806 which increasing ρ_1 has a greater effect on the local outbreak probability from the region in which increasing ρ_2
807 has a greater effect on the local outbreak probability, was computed in an analogous way, with the additional
808 restriction that we only considered the search directions $\theta = 0$ and $\theta = \pi/2$ (i.e. intensifying only surveillance
809 of nonsymptomatic or symptomatic hosts; see Fig S2 C).

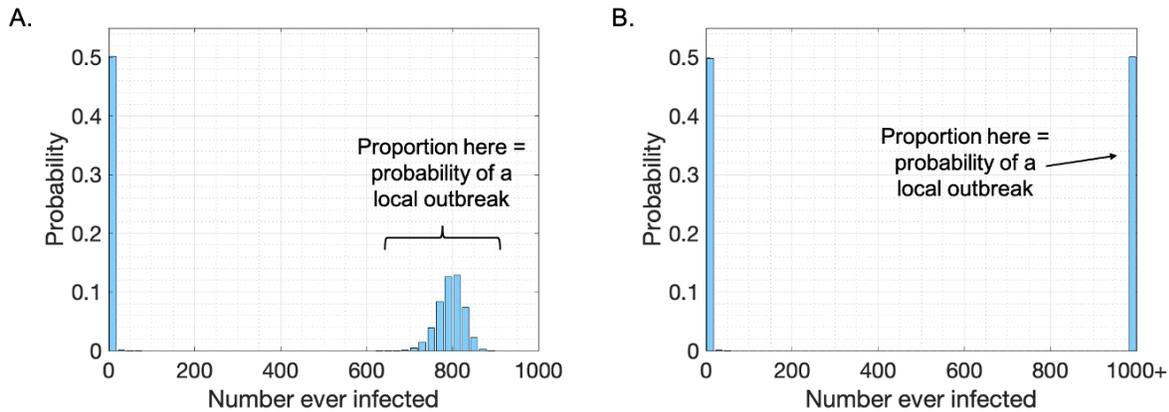
810
811 In Fig 3C, the white line represents the contour of steepest descent under the constraint that the total change in
812 surveillance effort ($\Delta\rho_1 + \Delta\rho_2 = S$) is held constant at each step, rather than fixing the search radius $(\Delta\rho_1)^2 +$
813 $(\Delta\rho_2)^2 = r^2$ as in Fig 3A. Therefore, instead of scanning over a circular arc, at each step we scan along the line
814 $\Delta\rho_1 = s, \Delta\rho_2 = S - s$, where c varies in the range $[0, S]$ (Fig S2 D). Otherwise, the process is completely
815 analogous to that described above.

816
817 **Text S3. Robustness of results to parameter values used**
818 We conducted supplementary analyses to investigate how our results are affected by varying the parameters
819 from their baseline values given in Table 1. We performed sensitivity analyses on the values of $R_0, \xi, K_p, K_a, \gamma +$
820 $\mu, \epsilon, \lambda, \nu$ and δ . For each of these, we present plots analogous to Fig 3D for six different values of the relevant
821 parameter (Figs S3-S12). In each case we considered, our qualitative message was unchanged: whenever the
822 maximum acceptable risk level was below a particular threshold value, the optimal strategy involved
823 surveillance targeting both nonsymptomatic and symptomatic individuals.

824
825 **Text S4. Details on computer code**
826 All computer code was written in the MATLAB programming environment (version R2019a).

827
828
829

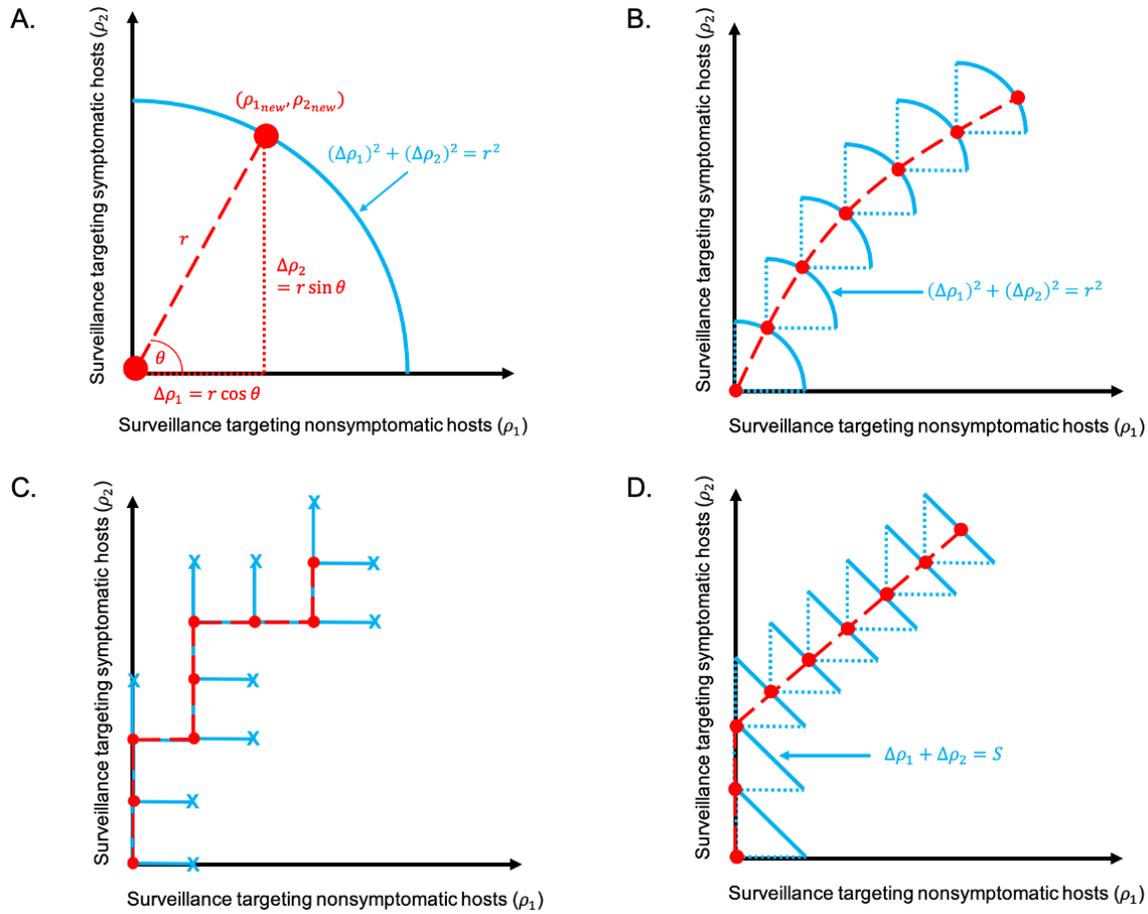
830 **Supplementary Figures**



831

832 **Fig S1. Illustrating the definition of the term ‘local outbreak’ with a simple example.** A. The total number
 833 of individuals ever infected (final epidemic size) in each of 100,000 simulations of a stochastic SIR (Susceptible-
 834 Infected-Removed) model with basic reproduction number $R_0 = 2$, beginning from a single infectious host each
 835 time. Within each simulation, each event is either an infection event (with probability $\frac{R_0 S}{R_0 S + N}$) or a removal event
 836 (with probability $\frac{N}{R_0 S + N}$), and simulations are run until the pathogen fades out (I hits zero). This provides a
 837 natural partitioning between simulations that fade out quickly and those that go on to become local outbreaks. In
 838 50% of simulations, fewer than 20 infections occurred in total (left hand peak); initial cases did not lead to
 839 sustained transmission in the population. In the remaining 50% of simulations (local outbreaks), between 620
 840 and 900 individuals were infected in total each time. The probability of a local outbreak is then defined as the
 841 proportion of simulations for which the final epidemic size lies within this natural upper range, here equal to 0.5.
 842 B. The analogous figure to panel A, but for the branching process version of the SIR model in which depletion of
 843 susceptibles is not accounted for (i.e. $S = N$ throughout each simulation). In this model, each event is either an
 844 infection event (with probability $\frac{R_0}{R_0 + 1}$) or a removal event (with probability $\frac{1}{R_0 + 1}$), and simulations are run until
 845 either the pathogen fades out (I hits zero) or 1000 infections have occurred. The rightmost bar corresponds to
 846 simulations in which 1000 infections occurred. There is again a natural partitioning between simulations that
 847 fade out quickly and those that go on to become local outbreaks, with the probability of a local outbreak
 848 matching the equivalent value for the stochastic SIR model (panel A).

849

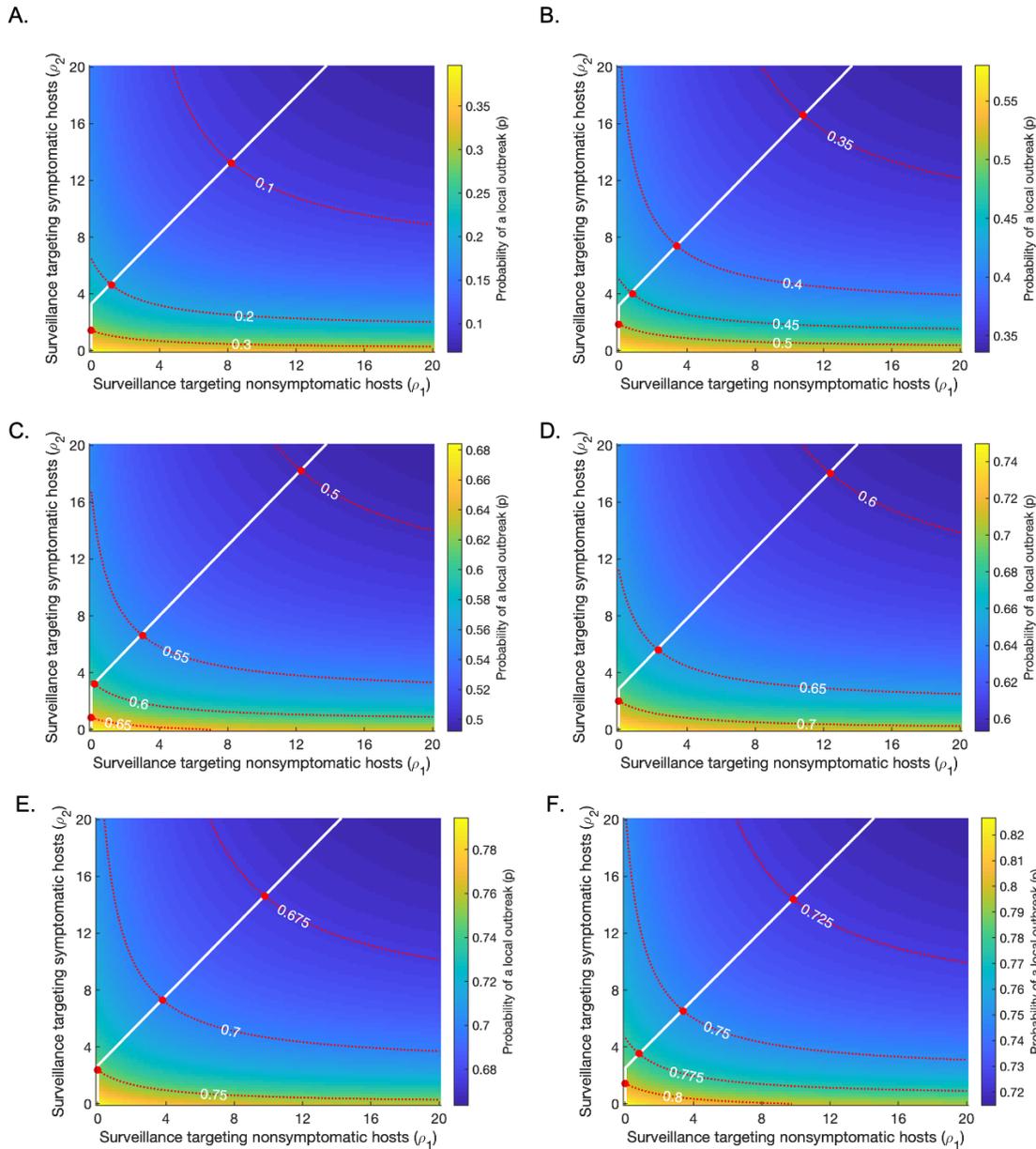


850

851 **Fig S2. Computation of the steepest descent contours shown in the main text.** A. To compute the steepest
 852 descent contours shown in Fig 3A of the main text, we increment ρ_1 and ρ_2 by scanning over a constant search
 853 radius $(\Delta\rho_1)^2 + (\Delta\rho_2)^2 = r^2$ (blue arc), and moving to the point $(\rho_{1_{new}}, \rho_{2_{new}})$ along that arc at which the local
 854 outbreak probability is minimised. B. The process shown in A is repeated to generate the complete contour (red
 855 dashed line). C. The analogous figure to B, in which the search direction is limited to directly to the right ($\theta =$
 856 0) or directly upwards ($\theta = \pi/2$). This procedure is used to generate the contour in Fig 3B in the main text. D.
 857 The analogous figure to B, in which the total change in surveillance effort ($\Delta\rho_1 + \Delta\rho_2 = S$) is held constant at
 858 each step, rather than the search radius. This procedure is used to generate the contour in Figs 3C and D in the
 859 main text.

860

861



862

863

864

865

866

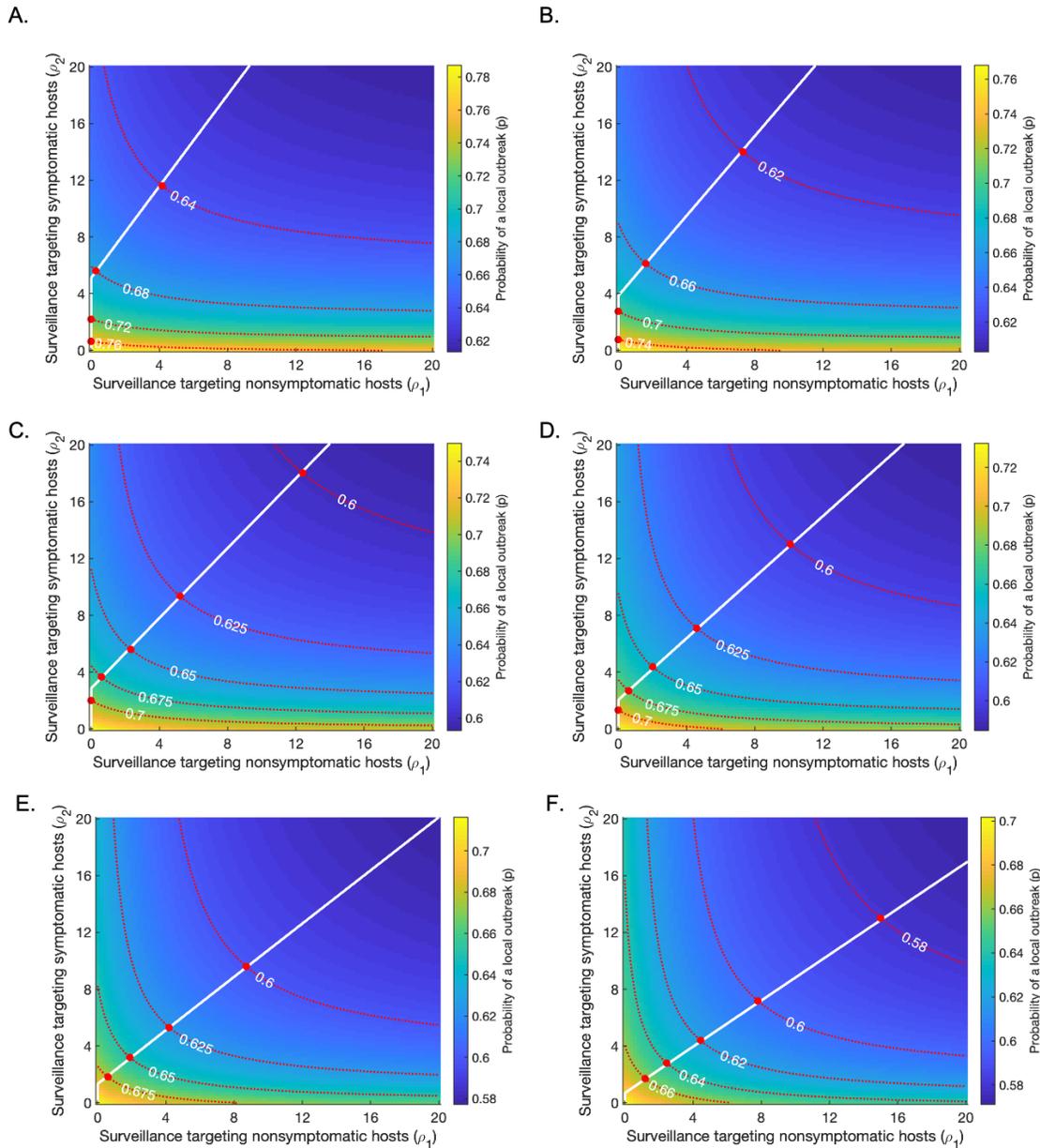
867

868

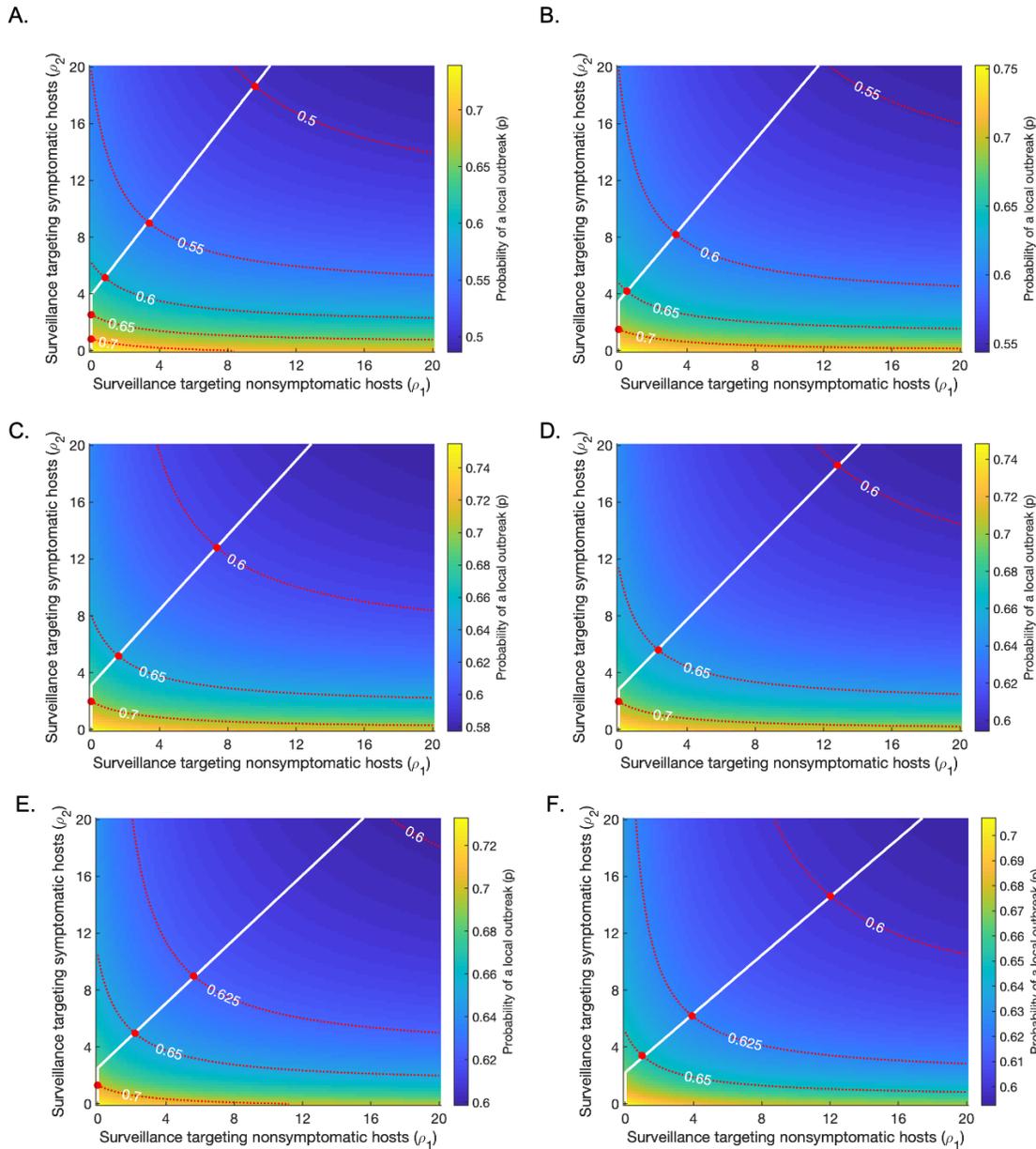
869

870

Fig S3. Varying the basic reproduction number R_0 from its baseline value ($R_0 = 3$). Plots are analogous to Fig 3D in the main text, showing strategies for minimising the surveillance effort required to achieve a pre-specified risk level (an “acceptable” local outbreak probability). Red dotted lines represent contours along which the probability of a local outbreak is constant, as labelled; red circles indicate the points along these contours at which the total surveillance effort $\rho_1 + \rho_2$ is minimised. The white line indicates the optimal strategy to follow if the pre-specified risk level is reduced. Apart from R_0 and β (which is changed in each panel to set the value of R_0), all parameters are held fixed at their baseline values given in Table 1. A. $R_0 = 1.5$. B. $R_0 = 2$. C. $R_0 = 2.5$. D. $R_0 = 3$ (baseline). E. $R_0 = 3.5$. F. $R_0 = 4$.

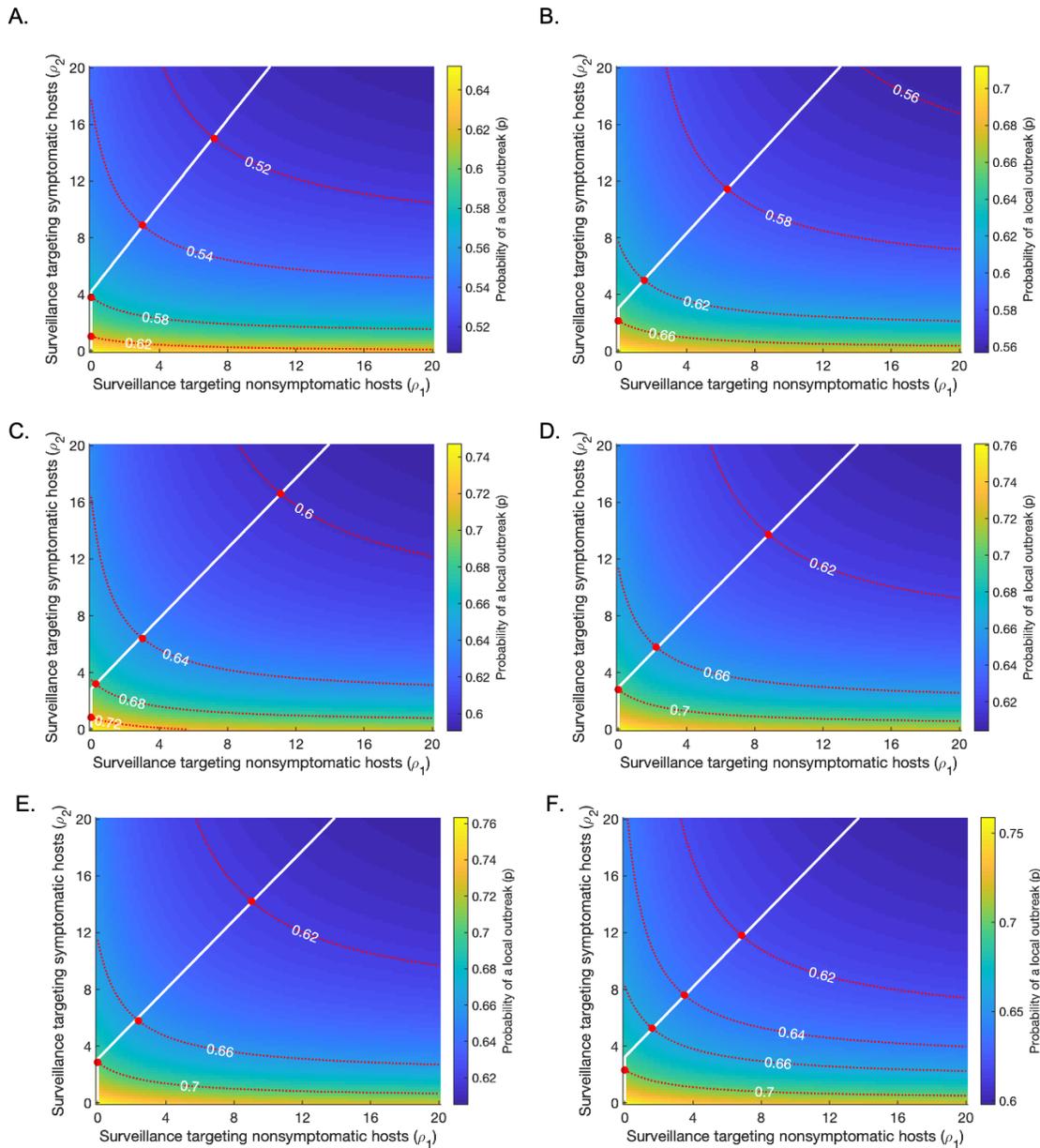


871
 872 **Fig S4. Varying the proportion of infections from asymptomatic infectors, ξ , from its baseline value ($\xi =$**
 873 **0.2).** Plots are analogous to Fig 3D in the main text, showing strategies for minimising the surveillance effort
 874 required to achieve a pre-specified risk level (an “acceptable” local outbreak probability). Red dotted lines
 875 represent contours along which the probability of a local outbreak is constant, as labelled; red circles indicate the
 876 points along these contours at which the total surveillance effort $\rho_1 + \rho_2$ is minimised. The white line indicates
 877 the optimal strategy to follow if the pre-specified risk level is reduced. Apart from ξ and β (which is changed in
 878 each panel to set $R_0 = 3$), all parameters are held fixed at their baseline values given in Table 1. A. $\xi = 0$. B.
 879 $\xi = 0.1$. C. $\xi = 0.2$ (baseline). D. $\xi = 0.3$. E. $\xi = 0.4$. F. $\xi = 0.5$.



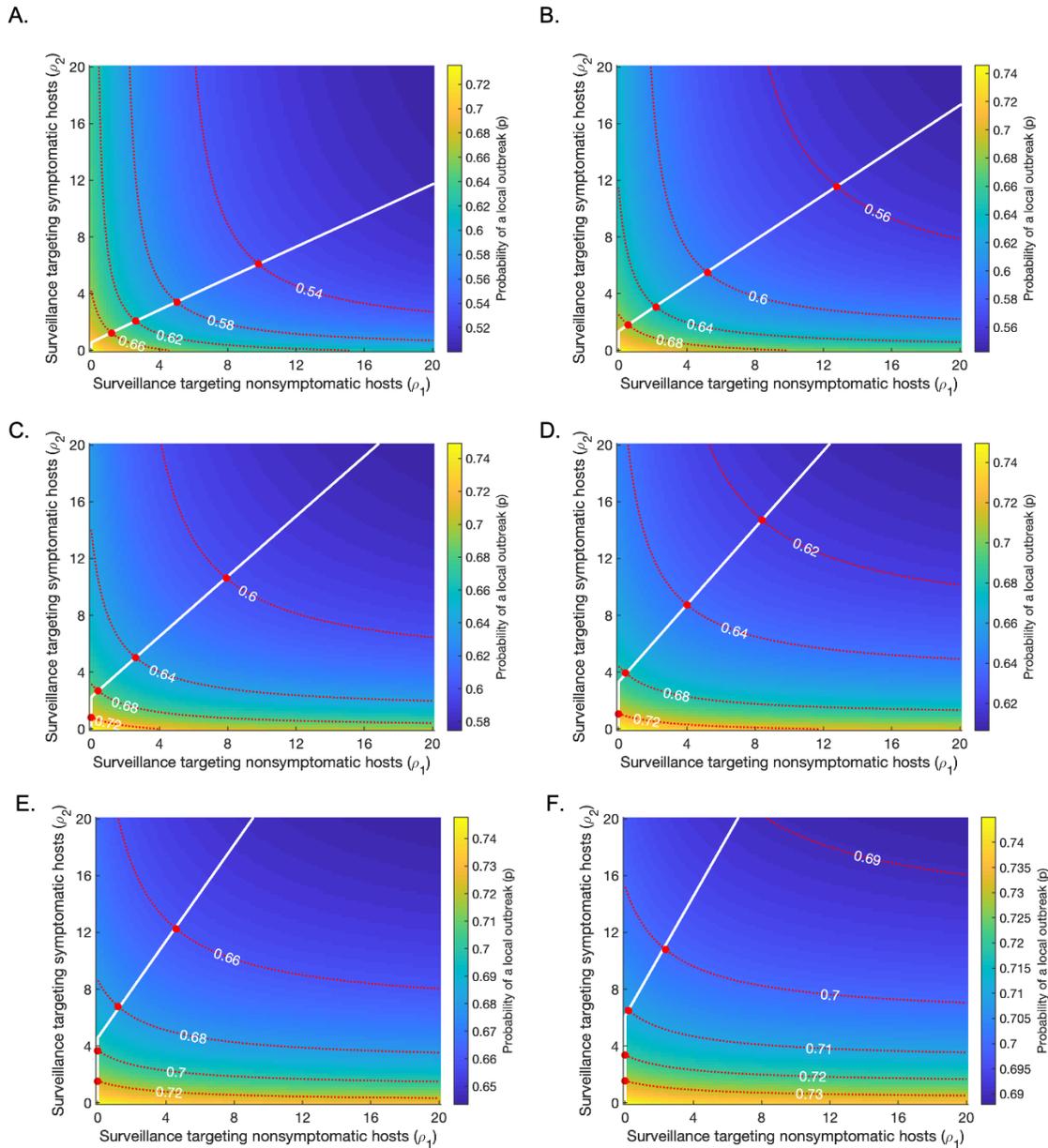
880
 881 **Fig S5. Varying the proportion of infections arising from presymptomatic hosts in the absence of**
 882 **intensified surveillance (K_p , given by expression (1) in the main text) from its baseline value ($K_p =$**
 883 **0.489).** In each case, the proportions of infections arising from asymptomatic and symptomatic hosts are
 884 adjusted so that they remain in the same ratio as in the baseline case. Plots are analogous to Fig 3D in the main
 885 text, showing strategies for minimising the surveillance effort required to achieve a pre-specified risk level (an
 886 “acceptable” local outbreak probability). Red dotted lines represent contours along which the probability of a
 887 local outbreak is constant, as labelled; red circles indicate the points along these contours at which the total
 888 surveillance effort $\rho_1 + \rho_2$ is minimised. The white line indicates the optimal strategy to follow if the pre-

889 specified risk level is reduced. Apart from K_p and K_a , as well as α and η (which are changed in each panel to set
 890 the values of K_p and K_a), and β (which is changed in each panel to set $R_0 = 3$), all parameters are held fixed at
 891 their baseline values given in Table 1. A. $K_p = 0.2$. B. $K_p = 0.3$. C. $K_p = 0.4$. D. $K_p = 0.5$. E. $K_p = 0.6$. F.
 892 $K_p = 0.7$.
 893



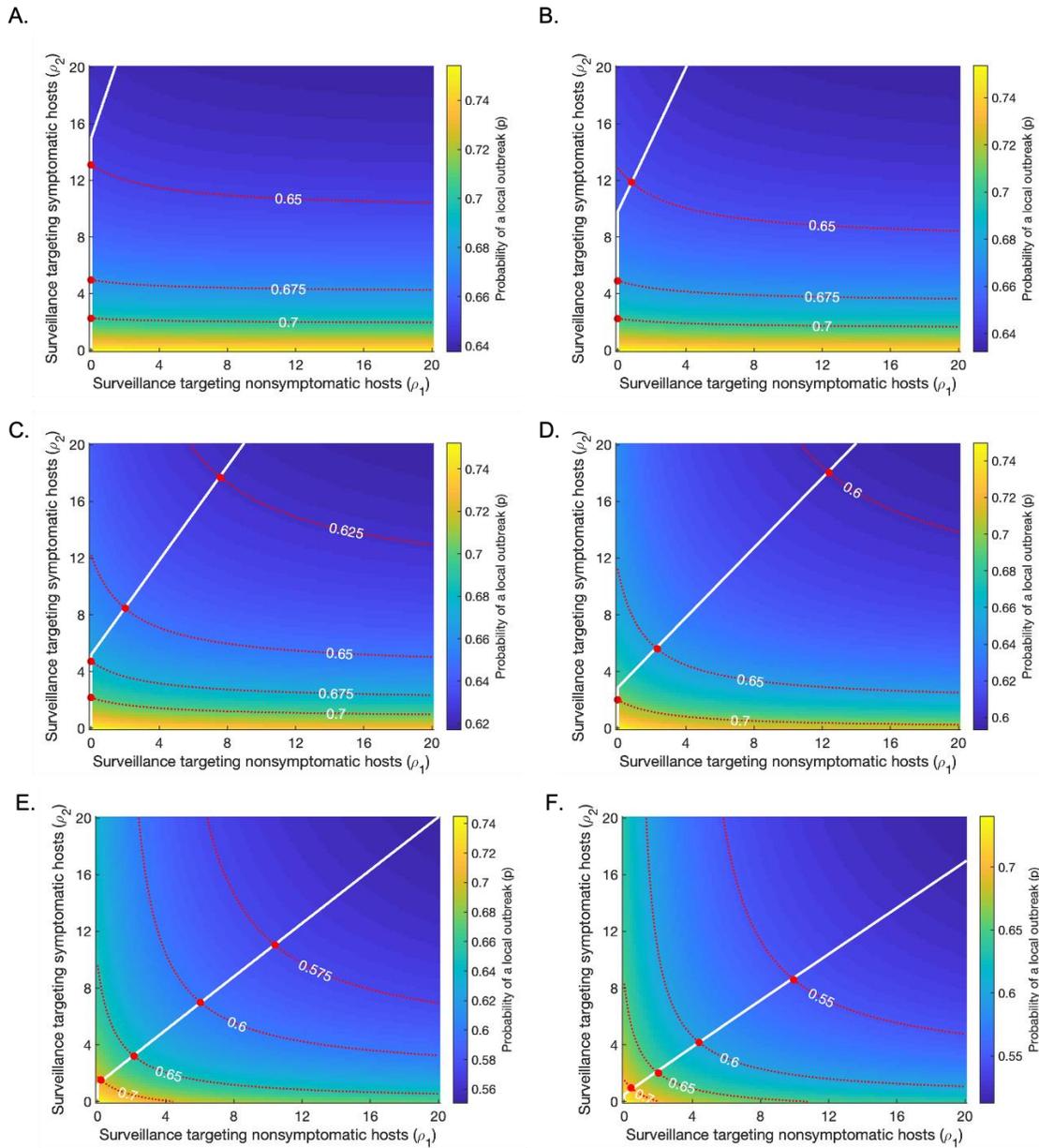
894
 895 **Fig S6. Varying the proportion of infections arising from asymptomatic hosts in the absence of intensified**
 896 **surveillance (K_a , given by expression (2) in the main text) from its baseline value ($K_a = 0.106$). In each**

897 case, the proportions of infections arising from presymptomatic and symptomatic hosts are adjusted so that they
898 remain in the same ratio as in the baseline case. Plots are analogous to Fig 3D in the main text, showing
899 strategies for minimising the surveillance effort required to achieve a pre-specified risk level (an “acceptable”
900 local outbreak probability). Red dotted lines represent contours along which the probability of a local outbreak is
901 constant, as labelled; red circles indicate the points along these contours at which the total surveillance effort
902 $\rho_1 + \rho_2$ is minimised. The white line indicates the optimal strategy to follow if the pre-specified risk level is
903 reduced. Apart from K_p and K_a , as well as α and η (which are changed in each panel to set the values of K_p and
904 K_a), and β (which is changed in each panel to set $R_0 = 3$), all parameters are held fixed at their baseline values
905 given in Table 1. A. $K_a = 0.01$. B. $K_a = 0.05$. C. $K_a = 0.1$. D. $K_a = 0.15$. E. $K_a = 0.2$. F. $K_a = 0.25$.
906



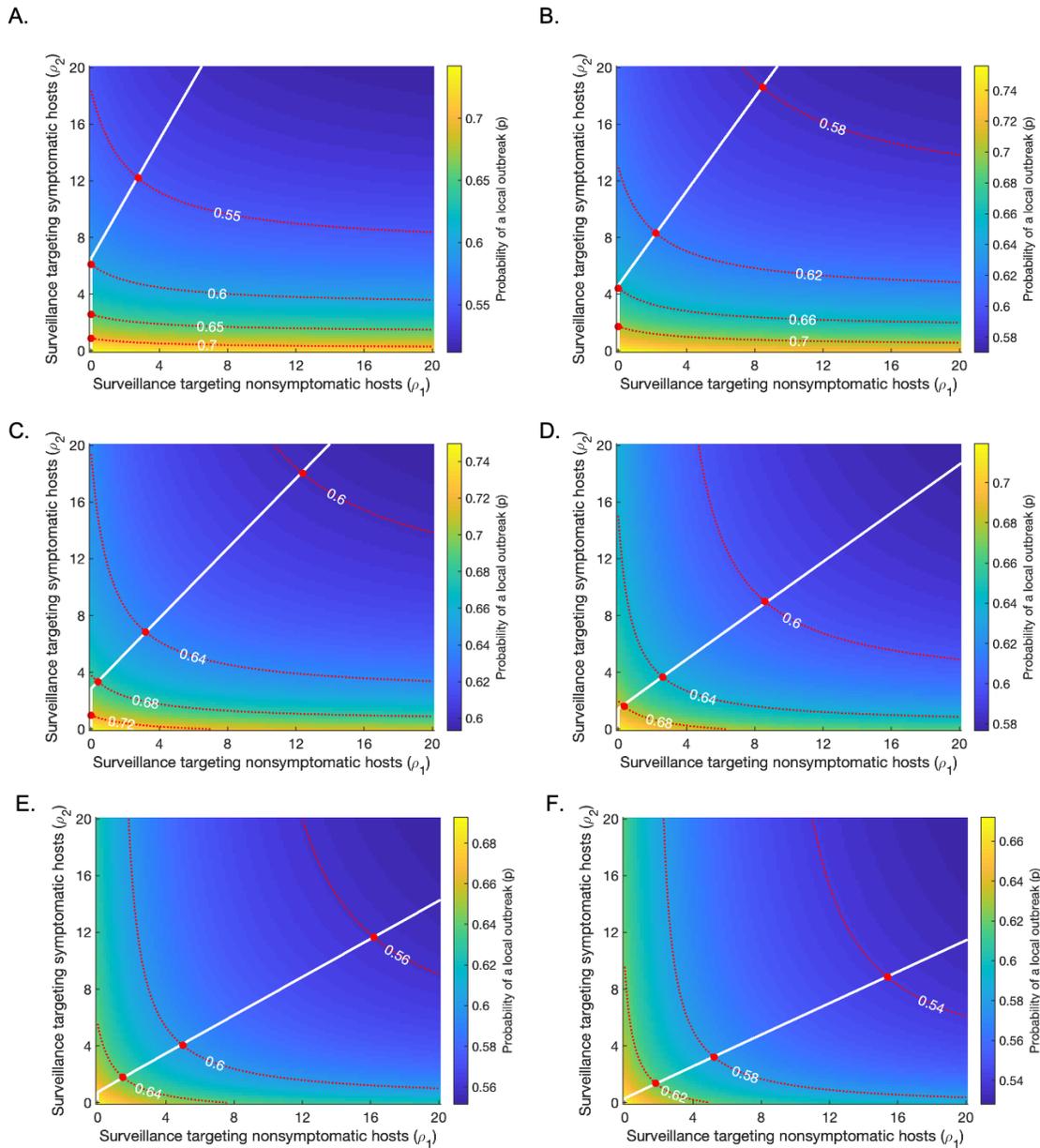
907
 908 **Fig S7. Varying the expected time period to isolation conditional on isolation occurring during the**
 909 **symptomatic period, $1/(\gamma + \mu)$, from its baseline value ($1/(\gamma + \mu) = 4.6$ days).** This is achieved by
 910 varying the parameter γ , whilst holding the recovery rate of symptomatic individuals μ equal to its baseline
 911 value ($\mu = 1/8 \text{ days}^{-1}$). Plots are analogous to Fig 3D in the main text, showing strategies for minimising the
 912 surveillance effort required to achieve a pre-specified risk level (an “acceptable” local outbreak probability). Red
 913 dotted lines represent contours along which the probability of a local outbreak is constant, as labelled; red circles
 914 indicate the points along these contours at which the total surveillance effort $\rho_1 + \rho_2$ is minimised. The white
 915 line indicates the optimal strategy to follow if the pre-specified risk level is reduced. Apart from γ and β (which

916 is changed in each panel to set $R_0 = 3$), all parameters are held fixed at their baseline values given in Table 1. A.
 917 $1/(\gamma + \mu) = 2$ days. B. $1/(\gamma + \mu) = 3$ days. C. $1/(\gamma + \mu) = 4$ days. D. $1/(\gamma + \mu) = 5$ days. E. $1/(\gamma + \mu) = 6$
 918 days. F. $1/(\gamma + \mu) = 7$ days.
 919



920
 921 **Fig S8. Varying ϵ , the relative isolation rate of nonsymptomatic individuals without intensified**
 922 **surveillance (compared to symptomatic individuals), from its baseline value ($\epsilon = 0.1$).** Plots are analogous
 923 to Fig 3D in the main text, showing strategies for minimising the surveillance effort required to achieve a pre-
 924 specified risk level (an “acceptable” local outbreak probability). Red dotted lines represent contours along which

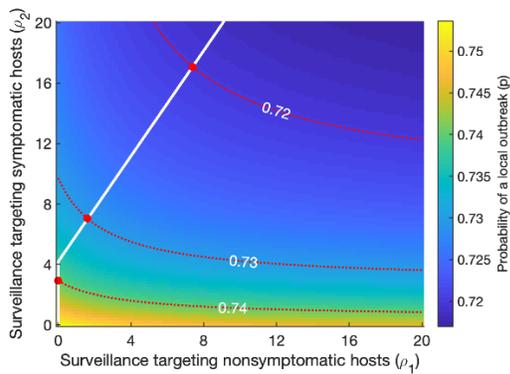
925 the probability of a local outbreak is constant, as labelled; red circles indicate the points along these contours at
 926 which the total surveillance effort $\rho_1 + \rho_2$ is minimised. The white line indicates the optimal strategy to follow if
 927 the pre-specified risk level is reduced. Apart from ϵ and β (which is changed in each panel to set $R_0 = 3$), all
 928 parameters are held fixed at their baseline values given in Table 1. A. $\epsilon = 0.01$. B. $\epsilon = 0.02$. C. $\epsilon = 0.05$. D.
 929 $\epsilon = 0.1$ (baseline). E. $\epsilon = 0.2$. F. $\epsilon = 0.3$.
 930



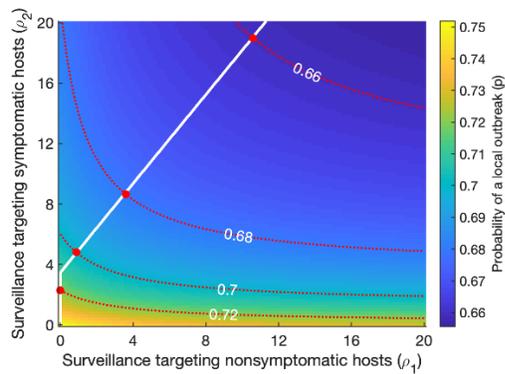
931
 932 **Fig S9. Varying the duration of the presymptomatic period, $1/\lambda$, from its baseline value ($1/\lambda = 0.5$ days).**
 933 Plots are analogous to Fig 3D in the main text, showing strategies for minimising the surveillance effort required

934 to achieve a pre-specified risk level (an “acceptable” local outbreak probability). Red dotted lines represent
 935 contours along which the probability of a local outbreak is constant, as labelled; red circles indicate the points
 936 along these contours at which the total surveillance effort $\rho_1 + \rho_2$ is minimised. The white line indicates the
 937 optimal strategy to follow if the pre-specified risk level is reduced. Apart from λ and β (which is changed in
 938 each panel to set $R_0 = 3$), all parameters are held fixed at their baseline values given in Table 1. A. $1/\lambda = 0.5$
 939 days. B. $1/\lambda = 1$ day. C. $1/\lambda = 2$ days (baseline). D. $1/\lambda = 4$ days. E. $1/\lambda = 6$ days. F. $1/\lambda = 8$ days.
 940

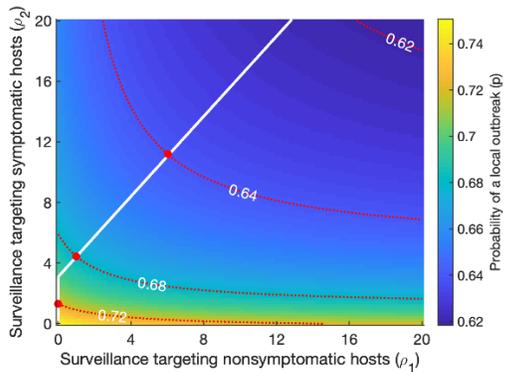
A.



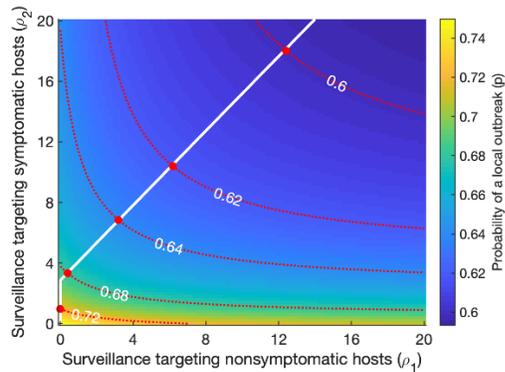
B.



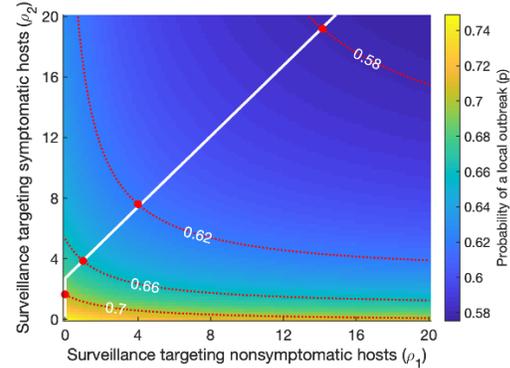
C.



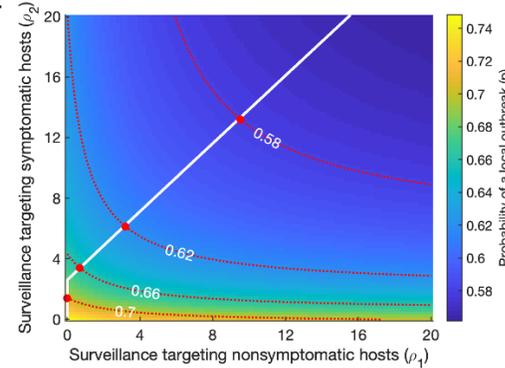
D.



E.

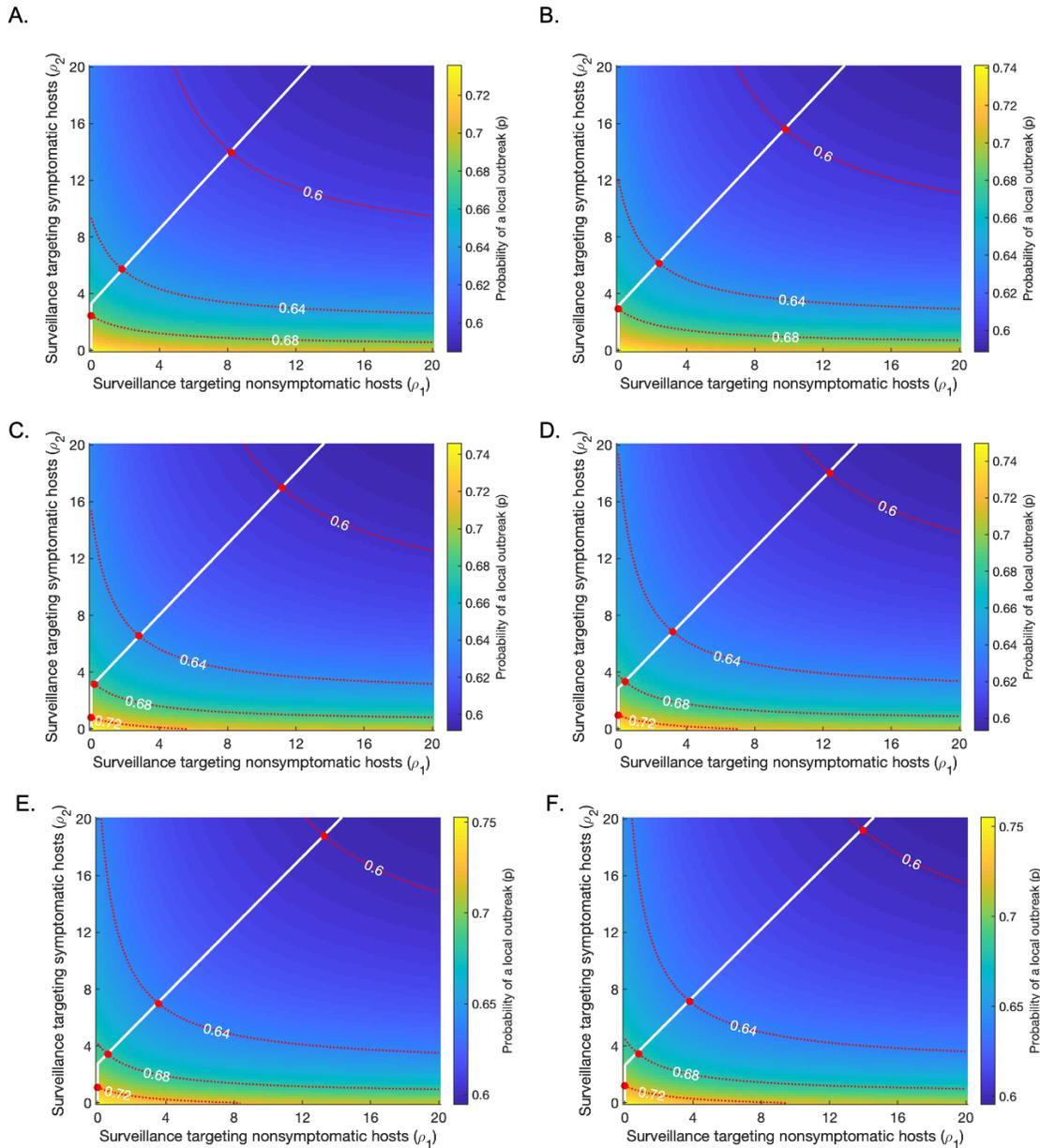


F.

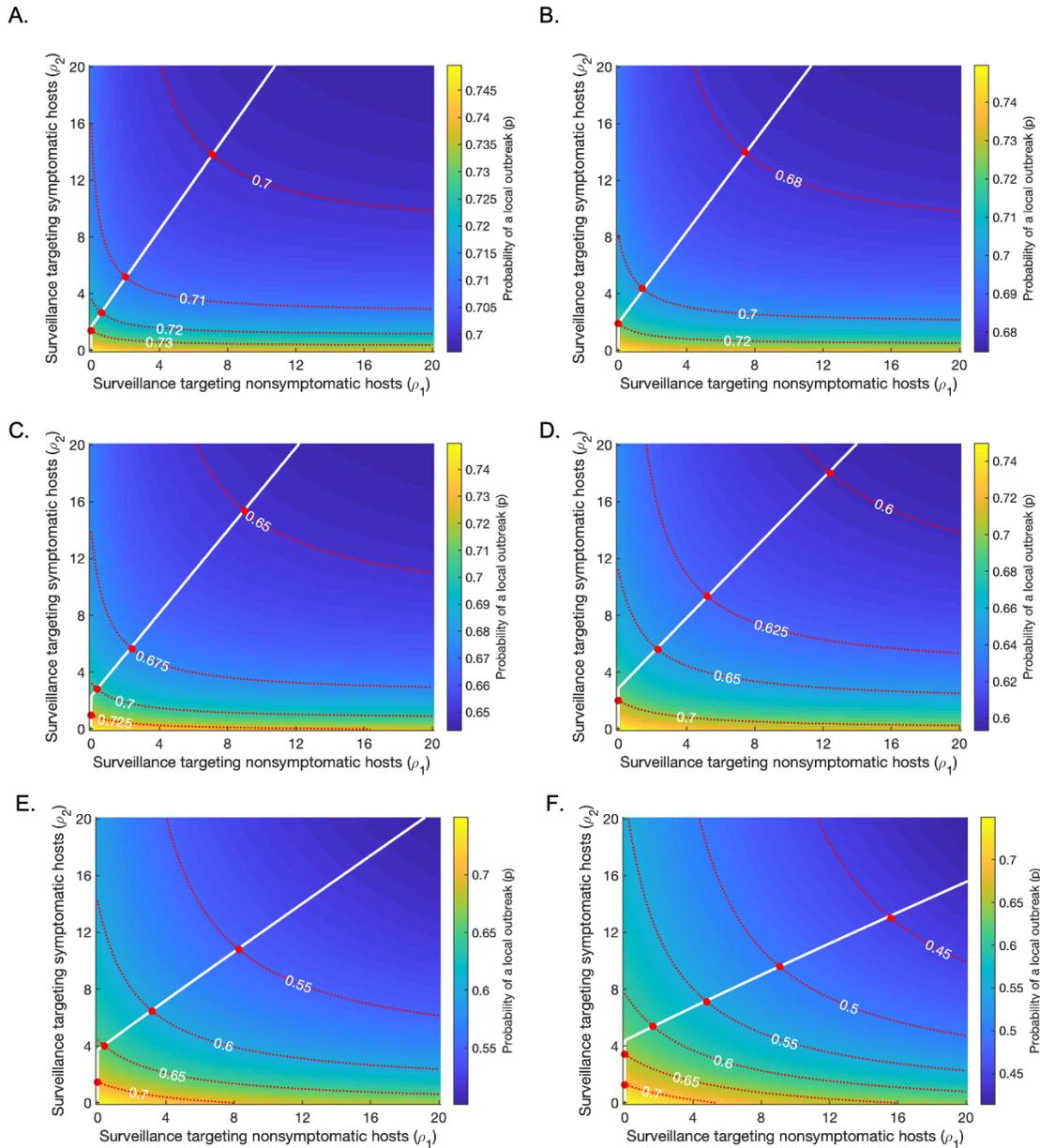


941

942 **Fig S10. Varying the duration of the symptomatic period, $1/\mu$, from its baseline value ($1/\mu = 8$ days).** The
943 parameter γ is varied simultaneously such that $1/(\gamma + \mu)$, the expected time period to isolation conditional on
944 isolation occurring during the symptomatic period, remains at its baseline value ($1/(\gamma + \mu) = 4.6$ days). Plots
945 are analogous to Fig 3D in the main text, showing strategies for minimising the surveillance effort required to
946 achieve a pre-specified risk level (an “acceptable” local outbreak probability). Red dotted lines represent
947 contours along which the probability of a local outbreak is constant, as labelled; red circles indicate the points
948 along these contours at which the total surveillance effort $\rho_1 + \rho_2$ is minimised. The white line indicates the
949 optimal strategy to follow if the pre-specified risk level is reduced. Apart from μ and γ , and β (which is changed
950 in each panel to set $R_0 = 3$), all parameters are held fixed at their baseline values given in Table 1. A. $1/\mu =$
951 5 days. B. $1/\mu = 6$ days. C. $1/\mu = 7$ days. D. $1/\mu = 8$ days (baseline). E. $1/\mu = 9$ days. F. $1/\mu = 10$ days.
952



953
 954 **Fig S11. Varying $1/\nu$, the duration of the asymptomatic period, from its baseline value ($1/\nu = 10$ days).**
 955 Plots are analogous to Fig 3D in the main text, showing strategies for minimising the surveillance effort required
 956 to achieve a pre-specified risk level (an “acceptable” local outbreak probability). Red dotted lines represent
 957 contours along which the probability of a local outbreak is constant, as labelled; red circles indicate the points
 958 along these contours at which the total surveillance effort $\rho_1 + \rho_2$ is minimised. The white line indicates the
 959 optimal strategy to follow if the pre-specified risk level is reduced. Apart from ν and β (which is changed in
 960 each panel to set $R_0 = 3$), all parameters are held fixed at their baseline values given in Table 1. A. $1/\nu = 7$
 961 days. B. $1/\nu = 8$ days. C. $1/\nu = 9$ days. D. $1/\nu = 10$ days (baseline). E. $1/\nu = 11$ days. F. $1/\nu = 12$ days.



962
 963 **Fig S12. Varying the upper bound on the fractional reduction in the time to isolation (if no other event**
 964 **occurs), δ , from its baseline value ($\delta = 0.8$).** Plots are analogous to Fig 3D in the main text, showing
 965 strategies for minimising the surveillance effort required to achieve a pre-specified risk level (an “acceptable”
 966 local outbreak probability). Red dotted lines represent contours along which the probability of a local outbreak is
 967 constant, as labelled; red circles indicate the points along these contours at which the total surveillance effort
 968 $\rho_1 + \rho_2$ is minimised. The white line indicates the optimal strategy to follow if the pre-specified risk level is
 969 reduced. Apart from δ , all parameters are held fixed at their baseline values given in Table 1. A. $\delta = 0.5$. B. $\delta =$
 970 0.6. C. $\delta = 0.7$. D. $\delta = 0.8$ (baseline). E. $\delta = 0.9$. F. $\delta = 0.95$