Interventions targeting nonsymptomatic cases can be important to prevent local outbreaks: COVID-19 as a case-study

AUTHORS

F.A. Lovell-Read[1], S. Funk[2], U. Obolski[3,4], C.A. Donnelly[5,6], R.N. Thompson[1,2,7]

AFFILIATIONS

[1]Mathematical Institute, University of Oxford (Andrew Wiles Building, Radcliffe Observatory Quarter, Woodstock Road, Oxford, OX2 6GG, UK)

[2]Centre for Mathematical Modelling of Infectious Diseases, London School of Hygiene & Tropical Medicine (Keppel Street, London, WC1E 7HT, UK)

[3]Porter School of the Environment and Earth Sciences, Tel Aviv University (Tel Aviv, Israel, 6997801)

[4]School of Public Health, Tel Aviv University (Tel Aviv, Israel, 6997801)

[5]Department of Statistics, University of Oxford (24-29 St. Giles', Oxford, OX1 3LB, UK)

[6]MRC Centre for Global Infectious Disease Analysis, Department of Infectious Disease Epidemiology, Imperial College London (Norfolk Place, London, W2 1PG, UK)

[7]Christ Church, University of Oxford, (St. Aldate's, Oxford, OX1 1DP, UK)

CORRESPONDING AUTHOR

F.A. Lovell-Read. Address: Merton College, Merton Street, Oxford, OX1 4JD.

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

ABSTRACT

Background

During infectious disease epidemics, a key question is whether cases travelling to new locations will trigger local outbreaks. The risk of this occurring depends on a range of factors, such as the transmissibility of the pathogen, the susceptibility of the host population and, crucially, the effectiveness of local surveillance in detecting cases and preventing onward spread. For many pathogens, presymptomatic and/or asymptomatic (together referred to here as nonsymptomatic) transmission can occur, making effective surveillance challenging. In this study, using COVID-19 as a case-study, we show how the risk of local outbreaks can be assessed when nonsymptomatic transmission can occur.

Methods

We construct a branching process model that includes nonsymptomatic transmission, and explore the effects of interventions targeting nonsymptomatic or symptomatic hosts when surveillance resources are limited. Specifically, we consider whether the greatest reductions in local outbreak risks are achieved by increasing surveillance and control targeting nonsymptomatic or symptomatic cases, or a combination of both.

Findings

Seeking to increase surveillance of symptomatic hosts alone is typically not the optimal strategy for reducing outbreak risks. Adopting a strategy that combines an enhancement of surveillance of symptomatic cases with efforts to find and isolate nonsymptomatic hosts leads to the largest reduction in the probability that imported cases will initiate a local outbreak.

Interpretation

During epidemics of COVID-19 and other infectious diseases, effective surveillance for nonsymptomatic hosts can be crucial to prevent local outbreaks.

Funding

UKRI-BBSRC, Wellcome Trust, UKRI-MRC, FCDO, EDCTP2, Christ Church (Oxford).

1. INTRODUCTION

Emerging epidemics represent a substantial challenge to human health worldwide.[1,2] When cases are clustered in specific locations, two key questions are: i) Will exported cases lead to local outbreaks in new locations? and ii) Which surveillance and control strategies in those new locations will reduce the risk of local outbreaks?

Branching process models are used for a range of diseases to assess whether cases that are newly arrived in a host population will generate a local outbreak driven by sustained local transmission.[3-6] These models can also be used to predict the effectiveness of potential control interventions. For example, early in the COVID-19 pandemic, Hellewell *et al.*[7] used simulations of a branching process model to predict whether or not new outbreaks would fade out under different contact tracing strategies. Thompson[8] estimated the probability of local outbreaks analytically using a branching process model and found that effective isolation of infectious hosts leads to a substantial reduction in the outbreak risk.

A factor that can hinder control interventions during any epidemic is the potential for individuals to transmit a pathogen while not showing symptoms. For COVID-19, the incubation period has been estimated to last approximately five or six days on average,[9,10] and presymptomatic transmission can occur during that period.[11-14] Additionally, asymptomatic infected individuals (those who never develop symptoms) are also thought to contribute to transmission.[11,15,16]

Motivated by the need to assess the risk of outbreaks outside China early in the COVID-19 pandemic, we show how the risk that imported cases will lead to local outbreaks can be estimated using a branching process model, including nonsymptomatic individuals in the model explicitly. We explore the effects of interventions that aim to reduce this risk. Under the assumption that detected infected hosts are isolated effectively, we consider whether it is most effective to dedicate resources to enhancing surveillance targeting symptomatic individuals, to instead focus on increasing surveillance for nonsymptomatic individuals, or to use a combination of these approaches.

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

2. METHODS

2.1 Model

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

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

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

$$\frac{\mathrm{d}A}{\mathrm{d}t} = \xi(\eta\beta A + \alpha\beta I_1 + \beta I_2) - \frac{\varepsilon\gamma}{1-\mathrm{f}(\rho_1)}A - \nu A,$$

$$\frac{\mathrm{d}I_1}{\mathrm{d}t} = (1-\xi)(\eta\beta A + \alpha\beta I_1 + \beta I_2) - \frac{\varepsilon\gamma}{1-\mathrm{f}(\rho_1)}I_1 - \lambda I_1,$$

$$\frac{\mathrm{d}I_2}{\mathrm{d}t} = \lambda I_1 - \frac{\gamma}{1-\mathrm{f}(\rho_2)}I_2 - \mu I_2.$$

In our model, the parameter β and its scaled counterparts $\alpha\beta$ and $\eta\beta$ represent the rates at which symptomatic, presymptomatic and asymptomatic hosts generate new infections, respectively.

Since we are modelling the beginning of a potential local outbreak, we assume that the size of the susceptible population remains approximately constant, and do not track depletion of susceptible individuals explicitly. The parameter λ governs the rate at which presymptomatic individuals develop symptoms, so that the expected duration of the presymptomatic period is $1/\lambda$ days in the absence of interventions. Similarly, without interventions, the expected durations of the symptomatic and asymptomatic infectious periods are $1/\mu$ days and $1/\nu$ days, respectively.

The baseline rate at which symptomatic individuals are detected and isolated is determined by the parameter γ. Its scaled counterpart $\epsilon\gamma$ is the analogous quantity for nonsymptomatic individuals, where the scaling factor $\epsilon < 1$ reflects the fact that nonsymptomatic individuals are more challenging to detect than symptomatic individuals. We assume that the sensitivity of surveillance is identical for presymptomatic and asymptomatic individuals, and therefore use the same isolation rate for both of these groups.

The parameters ρ_1 and ρ_2 represent the surveillance intensification effort targeted at nonsymptomatic and symptomatic hosts, respectively. The function $\mathrm{f}(\rho) = \frac{\delta\rho}{1+\rho}$ governs the reduction in the expected time to isolation for a given surveillance effort, ρ. The functional form of $\mathrm{f}(\rho)$ is chosen as it generates a reduced expected time to isolation when the surveillance effort increases, yet the isolation rate saturates and cannot increase indefinitely. The function $\mathrm{f}(\rho)$ is shown in Figure 1B for different values of the parameter δ.

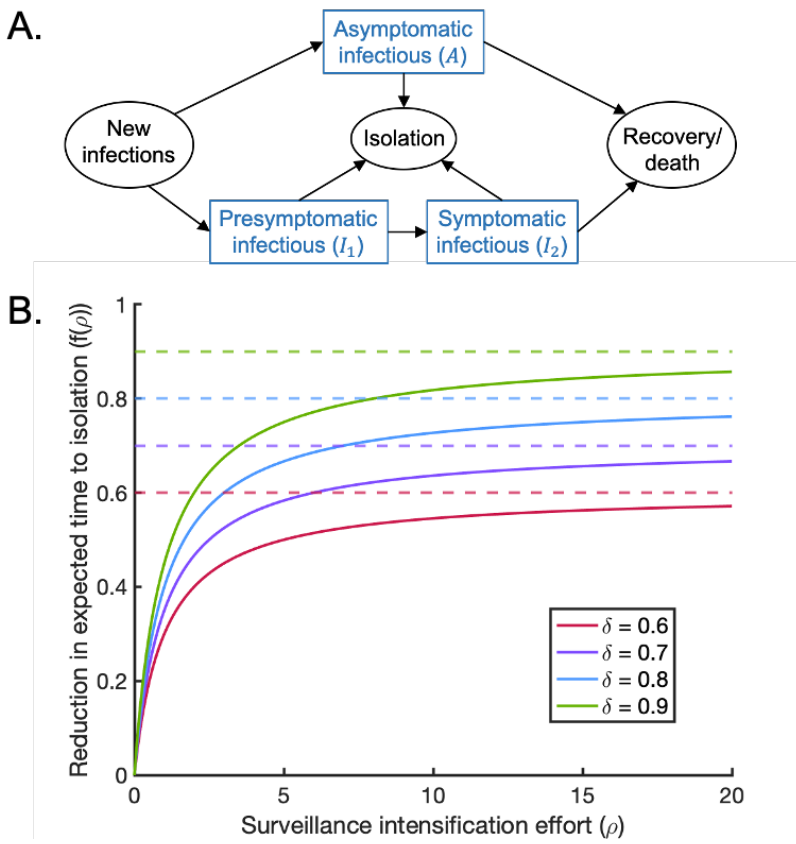


Figure 1: The branching process model used in our analyses. A. Schematic showing the different event types in the branching process model. The parameters of the model are described in the text and in Table 1. B. The relationship between the surveillance intensification effort (ρ) and the reduction in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

2.2 Reproduction number

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

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

This expression is the sum of the expected number of transmissions from a host who begins in the asymptomatic class and from a host who begins in the presymptomatic class, weighted by the respective probabilities ξ and $1 - \xi$ that determine the chance that the host experiences a fully asymptomatic course of infection. The expected number of transmissions from a host who begins in the presymptomatic class comprises transmissions occurring during the incubation period and transmissions occurring during the symptomatic period, accounting for the possibility that the host is isolated prior to developing symptoms.

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

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

and the equivalent quantity for asymptomatic hosts is given by

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

2.3 Baseline values of model parameters

Since this research was motivated by the need to estimate outbreak risks outside China in the initial stages of the COVID-19 pandemic, we used a baseline set of parameter values in our analyses that was informed by studies conducted during this pandemic (Table 1). Where possible, these parameter values were obtained from existing literature. However, we also

171 performed sensitivity analyses to determine how our results varied when the parameter values

172 were changed (see Supplementary Material).

173

174 The value of the parameter governing the baseline rate at which symptomatic individuals are

175 isolated, γ, was chosen to match empirical observations which indicate that individuals who seek

176 medical care prior to recovery or death do so around four to six days after symptom onset.[17]

177 Specifically, we assumed that time to first medical visit could be used a proxy for time to

178 isolation, and chose γ so that the expected time period to isolation conditional on isolation

179 occurring during the symptomatic period was given by $\frac{1}{\gamma+\mu} = 4 \cdot 6$ days.[17] This is different to the

180 time period that we refer to as the expected time to isolation for symptomatic hosts, which is $\frac{1}{\gamma}$

181 days (see Methods).

182

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

184

Parameter	Meaning	Baseline value	Justification
R_0	Expected number of secondary infections caused by a single infected individual (when $\rho_1 = \rho_2 = 0$)	$R_0 = 3$	Within estimated range for SARS-CoV-2[18-20]
ξ	Proportion of infections which are asymptomatic	$\xi = 0 \cdot 2$	[21]
β	Rate at which symptomatic individuals generate new infections	$\beta = 0 \cdot 336$ days^{-1} (3 s.f.)	Chosen so that $R_0 = 3$
α	Relative rate at which presymptomatic individuals generate new infections (compared to symptomatic individuals)	$\alpha = 2 \cdot 78$ days^{-1} (3 s.f.)	Chosen so that 48·9% of transmissions arise from presymptomatic hosts (i.e. $K_p = 0 \cdot 489$)[11]
η	Relative rate at which asymptomatic individuals generate new infections (compared to symptomatic individuals)	$\eta = 0 \cdot 519$ days^{-1} (3 s.f.)	Chosen so that 10·6% of transmissions arise from asymptomatic hosts (i.e. $K_a = 0 \cdot 106$)[11]
γ	Isolation rate of symptomatic individuals without intensified surveillance	$\gamma = 0 \cdot 0924$ days^{-1} (3 s.f)	Chosen so that $\frac{1}{\gamma+\mu} = 4 \cdot 6$ days[17]
ε	Relative isolation rate of nonsymptomatic individuals without intensified surveillance (compared to symptomatic individuals)	$\epsilon = 0 \cdot 1$	Assumed (for different values, see Figure S7)
λ	Rate at which presymptomatic individuals develop symptoms	$\lambda = 0 \cdot 5$ days^{-1}	[14]
μ	Recovery rate of symptomatic individuals	$\mu = 1/8$ days^{-1}	[22-24]

ν	Recovery rate of asymptomatic individuals	$\nu = 0 \cdot 1$ days^{-1}	Chosen so that, in the absence of interventions, the expected duration of infection is identical for all infected hosts ($\frac{1}{\nu} = \frac{1}{\lambda} + \frac{1}{\mu}$)
δ	Upper bound on the fractional reduction in the time to isolation (if no other event occurs)	$\delta = 0 \cdot 8$	Assumed (for different values, see Figure 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

2.4 Probability of a local outbreak

The probability that a single imported infectious host initiates a local outbreak was calculated analytically using the branching process model.

The probability of a local outbreak not occurring, starting from i presymptomatic hosts, j symptomatic hosts, and k asymptomatic hosts was denoted by $q_{i,j,k}$. Starting from one presymptomatic host (so that $i = 1$ and $j = k = 0$), there are four possibilities for the next event. That host could: i) generate a new asymptomatic infection (with probability $\frac{\xi\alpha\beta}{\alpha\beta+\lambda+\frac{\varepsilon\gamma}{1-f(\rho_1)}}$); ii) generate a new presymptomatic infection (with probability $\frac{(1-\xi)\alpha\beta}{\alpha\beta+\lambda+\frac{\varepsilon\gamma}{1-f(\rho_1)}}$); iii) develop symptoms (with probability $\frac{\lambda}{\alpha\beta+\lambda+\frac{\varepsilon\gamma}{1-f(\rho_1)}}$), or; iv) be isolated (with probability $\frac{\frac{\varepsilon\gamma}{1-f(\rho_1)}}{\alpha\beta+\lambda+\frac{\varepsilon\gamma}{1-f(\rho_1)}}$). Consequently,

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

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

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

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

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

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

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

$$q_{0,1,0} = c\xi q_{01,0} q_{0,0,1} + c(1-\xi)q_{1,0,0} q_{0,1,0} + (1-c), \qquad (4)$$

where $c = \frac{\beta}{\beta+\frac{\gamma}{1-f(\rho_2)}+\mu}$.

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

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

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

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

$$p_{1,0,0} = 1 - q^*_{1,0,0},$$

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

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

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

2.5 Role of the funding source

The funders had no role in study design, preparation of the manuscript or the decision to publish.

3. RESULTS

3.1 Probability of a local outbreak

We considered the effect of R_0 and the duration of the presymptomatic and asymptomatic periods on the probability of a local outbreak when a nonsymptomatic host enters a new host population (Figure 2). We considered presymptomatic periods of length $1/\lambda = 1$ day, $1/\lambda = 2$ days and $1/\lambda = 4$ days; in each case, the duration of the asymptomatic period ($1/\nu$ days) was adjusted so that the relative proportion of infections arising from asymptomatic hosts compared to presymptomatic hosts remained fixed ($K_a/K_p = 0 \cdot 218$, as in the baseline case). If instead nonsymptomatic infections are not accounted for, the infectious period follows an exponential distribution and the probability of a local outbreak is given by $p = 1 - 1/R_0$ (red dash-dotted

246 line in Figure 2A). Including nonsymptomatic infection in the model therefore led to an
247 increased risk of a local outbreak in the absence of surveillance intensification (Figure 2A).
248
249 We then considered the dependence of the probability of a local outbreak on the intensity of
250 surveillance targeting nonsymptomatic and symptomatic hosts (Figure 2B-D). The maximum
251 value of the surveillance intensification effort that we considered (given by ρ_1 or ρ_2 values of
252 20) corresponded to a 76% reduction in the expected time to isolation (blue line in Figure 1B),
253 i.e. a 76% reduction in $\frac{1}{\varepsilon\gamma}$ or $\frac{1}{\gamma}$.
254
255 The length of the presymptomatic and asymptomatic periods significantly affected the
256 dependence of the probability of a local outbreak on the level of surveillance targeted at
257 nonsymptomatic and symptomatic hosts. In Figure 2B, in which the duration of the
258 presymptomatic period was 1 day, increasing surveillance targeted at nonsymptomatic hosts (ρ_1)
259 had a limited effect on the probability of a local outbreak, while increasing surveillance targeted
260 at symptomatic hosts (ρ_2) had a more significant effect. For example, increasing the surveillance
261 effort targeted at nonsymptomatic hosts to $\rho_1 = 5$ (a 67% reduction in the time to isolation) only
262 reduced the probability of a local outbreak from 0·730 to 0·716 (3 s.f.), whereas the equivalent
263 effort targeted at symptomatic hosts ($\rho_2 = 5$) reduced the probability to 0·630 (3 s.f.). As shown
264 in Figures 3C and D, however, when the presymptomatic and asymptomatic periods were longer,
265 the benefit of directing surveillance resources towards detecting nonsymptomatic individuals
266 increased. This was because longer presymptomatic and asymptomatic periods increased the
267 proportion of infections generated by nonsymptomatic individuals ($K_p + K_a$, see equations (1)

and (2)); a presymptomatic period of 1 day, 2 days and 4 days corresponded to values of $K_p + K_a$ equal to 0·424, 0·595 and 0·746, respectively.

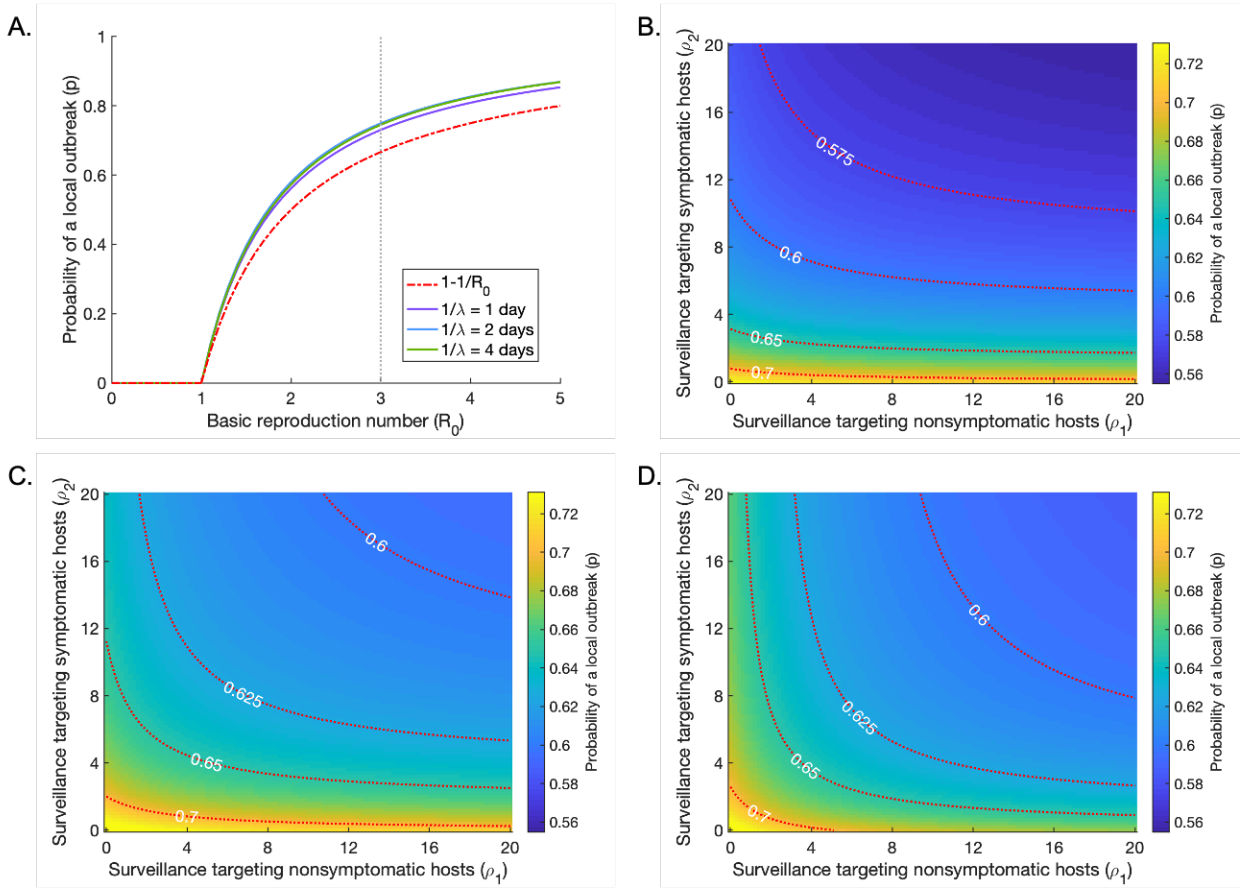


Figure 2. The effect of the duration of the presymptomatic and asymptomatic periods on the probability of a local outbreak (p), starting from a single nonsymptomatic host. A. The probability of a local outbreak as a function of the basic reproduction number R_0, for presymptomatic periods of lengths $1/\lambda = 1$ day (purple), $1/\lambda = 2$ days (blue) and $1/\lambda = 4$ days (green) in the absence of enhanced surveillance ($\rho_1 = \rho_2 = 0$). In each case, the duration of the asymptomatic period ($1/\nu$) is adjusted so that the relative proportion of infections arising from asymptomatic hosts compared to presymptomatic hosts remains constant ($K_a/K_p = 0.218$, as in the baseline case). The red dash-dotted line indicates the probability of a local outbreak in the absence of nonsymptomatic transmission. The vertical grey dotted line indicates $R_0 = 3$, the baseline value used throughout. B. The probability of a local outbreak as a function of the surveillance 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. The analogous figure to B but with $1/\lambda = 4$ days. Red dotted lines indicate contours of constant local outbreak probability (i.e. lines on which the probability of a local outbreak takes the values shown). The value of β is varied in each panel to give the appropriate value of R_0. All other parameter values are held fixed at the values in Table 1 (except where stated).

287 3.2 Optimising surveillance enhancement

288 We next considered in more detail the impact of surveillance targeted at nonsymptomatic hosts
289 (ρ_1) relative to the impact of surveillance targeted at symptomatic hosts (ρ_2). For our baseline
290 parameter values, we considered the probability of a local outbreak starting from a single
291 imported nonsymptomatic individual for a range of values of ρ_1 and ρ_2. We calculated the
292 steepest descent contours (white lines in Figure 3A) numerically using a gradient maximisation
293 approach, in which at each point the contour direction was determined by minimising the local
294 outbreak probability over a fixed search radius (see Supplementary Material). These contours
295 indicate how ρ_1 and ρ_2 should be altered to maximise the reduction in the probability of a local
296 outbreak. In this case, enhancing surveillance targeting both symptomatic and nonsymptomatic
297 hosts is always optimal (the steepest descent contours are neither horizontal nor vertical).
298

299 We then considered a scenario in which, at any time, it is only possible to direct resources
300 towards enhancing surveillance of either nonsymptomatic individuals or symptomatic
301 individuals (e.g. tracing and testing of nonsymptomatic contacts of known infectious individuals,
302 or screening for symptomatic individuals at public events). In Figure 3B, the blue region
303 represents values of ρ_1 and ρ_2 for which enhancing surveillance targeting symptomatic hosts (i.e.
304 increasing ρ_2) leads to a larger reduction in the local outbreak probability than enhancing
305 surveillance targeting nonsymptomatic hosts (i.e. increasing ρ_1). In contrast, in the green region,
306 enhancing surveillance of nonsymptomatic individuals is more effective than enhancing
307 surveillance of symptomatic individuals. The white line represents the steepest descent contour
308 starting from $\rho_1 = \rho_2 = 0$, under the constraint that surveillance can only be enhanced for
309 symptomatic or nonsymptomatic hosts at any time.

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

Objective 1: Minimise the probability of a local outbreak for a fixed total surveillance effort. First, we considered the question: given a fixed maximum surveillance effort ($\rho_1 + \rho_2 = C$), how should surveillance be targeted at nonsymptomatic and symptomatic hosts? This involves setting the values of ρ_1 and ρ_2 to minimise the local outbreak probability. The optimal strategies in this case are shown in Figure 3C. The red dotted lines represent contours along which the total surveillance effort $\rho_1 + \rho_2$ is held constant (i.e. different values of C). On each contour, the red circle indicates the point at which the local outbreak probability is minimised.

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

These results indicate that, if surveillance resources are such that C is greater than 2·8 (corresponding to a 59% reduction in time to isolation of symptomatic hosts), the optimal surveillance strategy involves both enhanced surveillance of symptomatic individuals and

nonsymptomatic individuals (the red dots correspond to positive values of ρ_1, unless C is less than 2·8).

Objective 2: Minimise the total surveillance effort to achieve a pre-specified reduction in the probability of a local outbreak

Second, we considered the question: given a pre-specified acceptable risk level (i.e. probability of a local outbreak), how should the surveillance level targeted at nonsymptomatic and symptomatic hosts be chosen? This involves choosing ρ_1 and ρ_2 to minimise $\rho_1 + \rho_2$ along a given contour corresponding to a fixed local outbreak probability (red dotted lines in Figure 3D). On each contour, the red circle indicates the point along that contour at which the total surveillance effort $\rho_1 + \rho_2$ is minimised. These optimal points also lie exactly along the line on which enhancing surveillance targeted at symptomatic hosts is equally effective compared to enhancing surveillance targeted at nonsymptomatic hosts.

As long as the target local outbreak probability was less than 0.69, optimal surveillance involved enhanced surveillance of nonsymptomatic individuals as well as symptomatic individuals. For example, in order to reduce the local outbreak probability to 0.6, the optimal approach was to deploy resources such that $\rho_1 = 12 \cdot 4$ (a 74% reduction in time to isolation of nonsymptomatic individuals) and $\rho_2 = 18 \cdot 0$ (a 76% reduction in time to isolation of symptomatic individuals).

Plots analogous to Figure 3D in which the parameters were varied from their baseline values are shown in the Supplementary Material. In each case we considered, our main message was unchanged. There always existed a threshold local outbreak probability such that, below this

threshold, the optimal strategy for further reduction in the local outbreak probability involved enhancing surveillance targeting both nonsymptomatic and symptomatic individuals.

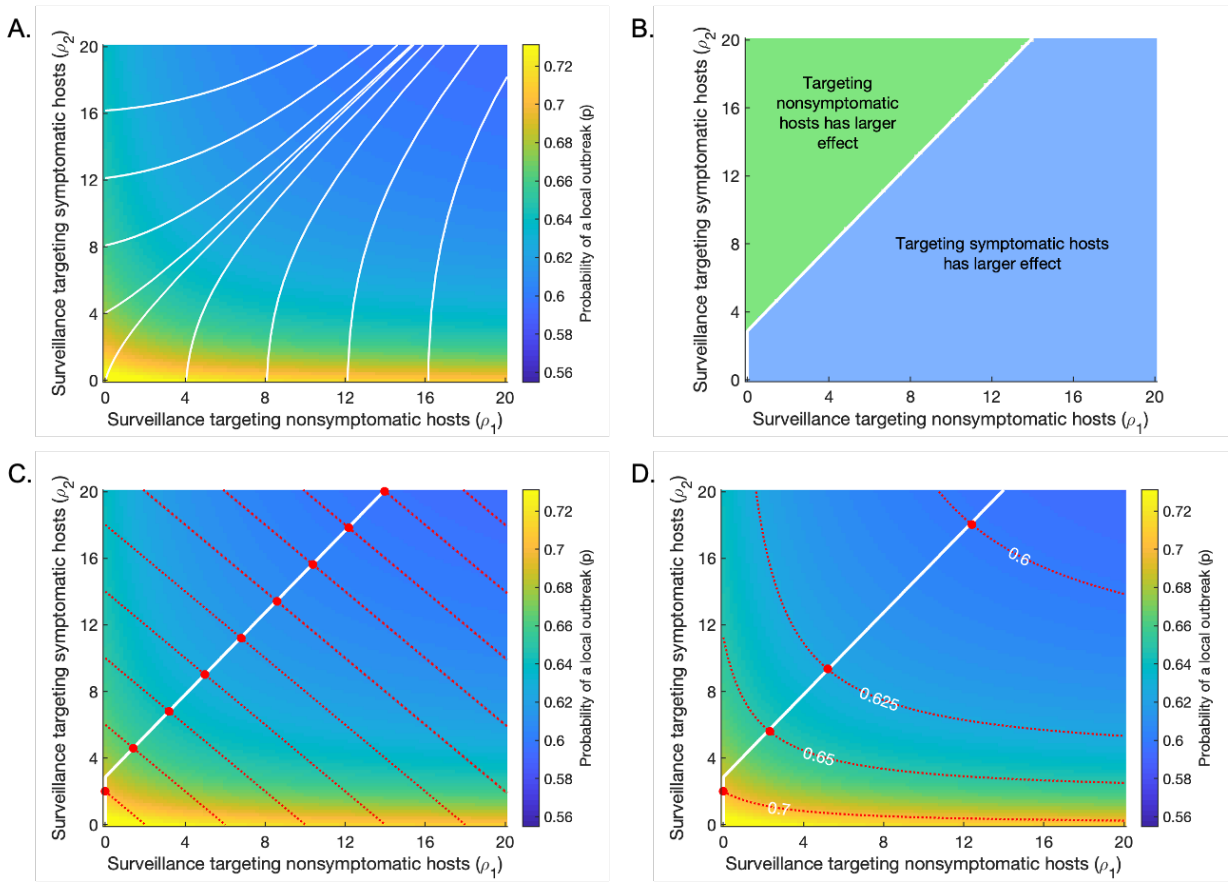


Figure 3. Optimal surveillance strategies to reduce the probability of a local outbreak (p) starting from a single nonsymptomatic host. A. The local outbreak probability for different values of ρ_1 and ρ_2, with the steepest descent contours overlaid (white lines). For the maximum reduction in the probability of a local outbreak at each point, surveillance must be enhanced for both nonsymptomatic and symptomatic individuals. B. Values of ρ_1 and ρ_2 for which increasing surveillance for nonsymptomatic hosts (i.e. increasing ρ_1) is more effective at reducing the local outbreak probability than increasing surveillance for symptomatic hosts (i.e. increasing ρ_2) (green region) and vice versa (blue region). The white line represents the steepest descent contour starting from $\rho_1 = \rho_2 = 0$, under the constraint that surveillance can only be enhanced for symptomatic or nonsymptomatic hosts at any time. The diagonal section of the steepest descent contour is made up of small horizontal and vertical sections. C. Strategies for minimising the local outbreak probability for a given fixed total surveillance effort ($\rho_1 + \rho_2 = C$). Red dotted lines indicate contours on which $\rho_1 + \rho_2$ is constant (i.e. lines on which $\rho_1 + \rho_2$ takes the values shown); red circles indicate the points along these contours at which the local outbreak probability is minimised. The white line indicates the optimal surveillance enhancement strategy if the maximum possible surveillance level (i.e. the maximum value of $\rho_1 + \rho_2 = C$) is increased. D. Strategies for minimising the surveillance

effort required to achieve a pre-specific risk level (an "acceptable" local outbreak probability). Red dotted lines indicate contours of constant local outbreak probability (i.e. lines on which the probability of a local outbreak takes the values shown); 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 increased or reduced.

4. DISCUSSION

A key component of infectious disease epidemic management is inferring the risk of outbreaks in different locations.[3-6,25] Different surveillance and control strategies can be introduced to reduce the risk that imported cases will lead to local outbreaks.[7,8,26] However, for a range of pathogens, public health measures are hindered by nonsymptomatic hosts who can transmit the pathogen yet are challenging to detect.[11,25,26]

Here, we showed how the probability of a local outbreak can be estimated using a branching process model that accounts for nonsymptomatic transmission (Figure 1). The model can be used to assess the local outbreak probability for different surveillance strategies that target nonsymptomatic or symptomatic individuals (Figure 2). Previous studies have shown that detection of nonsymptomatic infections can be a key component of epidemic forecasting[25] and containment,[26] and have demonstrated the benefits of identifying and isolating infectious nonsymptomatic hosts to reduce transmission.[11,12] We focused instead on investigating how surveillance should be targeted at nonsymptomatic or symptomatic hosts in order to reduce the probability that cases imported to new locations will trigger a local outbreak (Figure 3A,B). We also showed how the optimal surveillance level targeting these two groups can be assessed when surveillance resources are limited and policy-makers have specific objectives (Figure 3C,D). In each case, our main conclusion was that enhanced surveillance of nonsymptomatic hosts ($\rho_1 > 0$) can be an important component of reducing the local outbreak risk during epidemics.

401

402 Our goal here was to use the simplest possible model to explore the effects of surveillance of
403 nonsymptomatic and symptomatic individuals on the risk of local outbreaks. However, for our
404 modelling approach to be used to make precise quantitative predictions during epidemics, it
405 would be necessary to update the model to include the range of different specific surveillance
406 and control interventions that are in place. Detection of nonsymptomatic individuals is facilitated
407 by contact tracing and testing, which are carried out routinely during epidemics and can be
408 included in models explicitly.[7,26,27] Reductions in contacts due to social distancing strategies and
409 school or workplace closures could also be accounted for, although such interventions are often
410 introduced after a local outbreak has begun rather than in the initial phase of a potential local
411 outbreak as considered here. We modelled the level of surveillance targeted at nonsymptomatic
412 and symptomatic hosts in a simple way, using a function describing the relationship between
413 surveillance effort and effectiveness (Figure 1B). If different public health measures are included
414 in the model explicitly, then it would be possible to increase the accuracy of assessments of the
415 relative public health benefits of specific interventions that only target symptomatic individuals
416 (e.g. screening for passengers with heightened temperatures at airports) compared to
417 interventions that also target nonsymptomatic hosts (e.g. travel bans or quarantine of all inbound
418 passengers). Of course, this would require data from which the relative effectiveness of different
419 measures could be inferred.

420

421 In this article, we chose a baseline set of parameter values that were consistent with findings of
422 studies conducted during the COVID-19 pandemic, although constructing a detailed transmission
423 model for this pandemic was not our main focus. For example, we set the relative rates at which

presymptomatic and asymptomatic individuals generate new infections compared to symptomatic individuals so that 48·9% of transmissions arise from presymptomatic infectors, and 10·6% arise from asymptomatic infectors.[11] While this is in line with other reported estimates,[28,29] there is substantial variation between studies. We therefore also conducted sensitivity analyses in which we explored a range of different values of model parameters (Supplementary Material). In each case we considered, our main conclusion was unchanged: surveillance of nonsymptomatic individuals can contribute to reducing the risk of local outbreaks. This result is therefore expected to hold for epidemics of any pathogen for which nonsymptomatic individuals contribute significantly to transmission.

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

DECLARATION OF INTERESTS

We declare no competing interests.

AUTHORS' CONTRIBUTIONS

Conceptualisation: All authors. Methodology: FALR, UO, RNT. Investigation: FALR. Writing – original draft: FALR, RNT. Writing – review and editing: All authors. Supervision: RNT.

FUNDING

FALR acknowledges funding from the Biotechnology and Biological Sciences Research Council (UKRI-BBSRC), grant number BB/M011224/1. SF and RNT acknowledge funding from the Wellcome Trust, grant number 210758/Z/18/Z. RNT also acknowledges funding from Christ Church (University of Oxford) via a Junior Research Fellowship. CAD acknowledges funding from the MRC Centre for Global Infectious Disease Analysis (reference MR/R015600/1), jointly funded by the UK Medical Research Council (MRC) and the UK Foreign, Commonwealth & Development Office (FCDO) under the MRC/FCDO Concordat agreement, and is also part of the EDCTP2 programme supported by the European Union.

ACKNOWLEDGEMENTS

Thank you to Andrew Wood for proofreading.

REFERENCES

1. Bloom DE, Cadarette D. Infectious Disease Threats in the Twenty-First Century: Strengthening the Global Response. *Front. Immunol.* 2019; **10**(549).

2. Morens DM, Fauci AS. Emerging Infectious Diseases: Threats to Human Health and Global Stability. *PLos Pathog.* 2013; **9**(7): e1003467.

3. Boldog P, Tekeli T, Vizi Z, Denes A, Bartha FA, Rost G. Risk Assessment of Novel Coronavirus COVID-19 Outbreaks Outside China. *J. Clin. Med.* 2020; **9**(2): 571.

470 4. Daon Y, Thompson RN, Obolski U. Estimating COVID-19 outbreak risk through air
471 travel. *J. Travel Med.* 2020; **27**(5).
472 5. Lashari AA, Trapman P. Branching process approach for epidemics in dynamic
473 partnership network. *J. Math. Biol.* 2018; **76**(1): 265-94.
474 6. Thompson RN, Thompson CP, Pelerman O, Gupta S, Obolski U. Increased frequency of
475 travel in the presence of cross-immunity may act to decrease the chance of a global pandemic.
476 *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 2019; **374**(1775): 20180274.
477 7. Hellewell J, Abbott S, Gimma A, et al. Feasibility of controlling COVID-19 outbreaks by
478 isolation of cases and contacts. *Lancet Glob. Health* 2020; **8**(4): e488-e96.
479 8. Thompson RN. Novel Coronavirus Outbreak in Wuhan, China, 2020: Intense
480 Surveillance Is Vital for Preventing Sustained Transmission in New Locations. *J. Clin. Med.*
481 2020; **9**(2): 498.
482 9. Lauer SA, Grantz KH, Bi Q, et al. The Incubation Period of Coronavirus Disease 2019
483 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. *Ann.*
484 *Intern. Med.* 2020; **172**(9): 577-82.
485 10. McAloon C, Collins A, Hunt K, et al. Incubation period of COVID-19: a rapid systematic
486 review and meta-analysis of observational research. *BMJ Open* 2020; **10**(8): e039652.
487 11. Ferretti L, Wymant C, Kendall M, et al. Quantifying SARS-CoV-2 transmission suggests
488 epidemic control with digital contact tracing. *Science* 2020; **368**(6491): eabb6936.
489 12. Moghadas SM, Fitzpatrick MC, Sah P, et al. The implications of silent transmission for
490 the control of COVID-19 outbreaks. *Proc. Natl. Acad. Sci. U. S. A.* 2020; **117**(30): 17513-5.
491 13. Tindale LC, Stockdale JE, Coombe M, et al. Evidence for transmission of COVID-19
492 prior to symptom onset. *eLife* 2020; **9**: e57149.

493 14. Wei WE, Li Z, Chiew CJ, Yong SE, Toh MP, Lee VJ. Presymptomatic Transmission of
494 SARS-CoV-2 — Singapore, January 23–March 16, 2020. *Morb. Mortal. Weekly Rep.* 2020;
495 **69**(14): 411-5.
496 15. Bai Y, Yao L, Wei T, et al. Presumed Asymptomatic Carrier Transmission of COVID-19.
497 *J. Am. Med. Assoc.* 2020; **323**(14): 1406-7.
498 16. Furukawa NW, Brooks JT, Sobel J. Evidence Supporting Transmission of Severe Acute
499 Respiratory Syndrome Coronavirus 2 While Presymptomatic or Asymptomatic. *Emerging*
500 *Infectious Diseases* 2020; **26**(7): e201595.
501 17. Li Q, Guan X, Wu P, et al. Early Transmission Dynamics in Wuhan, China, of Novel
502 Coronavirus–Infected Pneumonia. *N. Engl. J. Med.* 2020; **382**(13): 1199-207.
503 18. Lai C-C, Shih T-P, Ko W-C, Tang H-J, Hsueh P-R. Severe acute respiratory syndrome
504 coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): The epidemic and the
505 challenges. *Int. J. Antimicrob. Agents* 2020; **55**(3): 105924.
506 19. Liu Y, Gayle AA, Wilder-Smith A, Rocklov J. The reproductive number of COVID-19 is
507 higher compared to SARS coronavirus. *J. Travel Med.* 2020; **27**(2): taaa021.
508 20. Zhao S, Lin Q, Ran J, et al. Preliminary estimation of the basic reproduction number of
509 novel coronavirus (2019-nCoV) in China, from 2019 to 2020: A data-driven analysis in the early
510 phase of the outbreak. *Int. J. Infect. Dis.* 2020; **92**: 214-7.
511 21. Buitrago-Garcia D, Egli-Gany D, Counotte MJ, et al. Occurrence and transmission
512 potential of asymptomatic and presymptomatic SARS-CoV-2 infections: A living systematic
513 review and meta-analysis. *PLoS Med.* 2020; **17**(9): e1003346.
514 22. Arons MM, Hatfield KM, Reddy SC, et al. Presymptomatic SARS-CoV-2 Infections and
515 Transmission in a Skilled Nursing Facility. *N. Engl. J. Med.* 2020; **382**(22): 2081-90.

516 23. Bullard J, Dust K, Funk D, et al. Predicting Infectious Severe Acute Respiratory
517 Syndrome Coronavirus 2 From Diagnostic Samples. *Clin. Infect. Dis.* 2020: ciaa638.
518 24. Wolfel R, Corman VM, Guggemos W, et al. Virological assessment of hospitalized
519 patients with COVID-2019. *Nature* 2020; **581**(7809): 465-9.
520 25. Thompson RN, Gilligan CA, Cunniffe NJ. Detecting Presymptomatic Infection Is
521 Necessary to Forecast Major Epidemics in the Earliest Stages of Infectious Disease Outbreaks.
522 *PLoS Comput. Biol.* 2016; **12**(4): e1004836.
523 26. Fraser C, Riley S, Anderson RM, Ferguson NM. Factors that make an infectious disease
524 outbreak controllable. *Proc. Natl. Acad. Sci. U. S. A.* 2004; **101**(16): 6146-51.
525 27. Webb G, Browne C, Huo X, Seydi O, Seydi M, Magal P. A Model of the 2014 Ebola
526 Epidemic in West Africa with Contact Tracing. *PLoS Curr.* 2015; **7**.
527 28. He X, Lau EHY, Wu P, et al. Temporal dynamics in viral shedding and transmissibility
528 of COVID-19. *Nat. Med.* 2020; **26**(5): 672-5.
529 29. Ren X, Li Y, Yang X, et al. Evidence for pre-symptomatic transmission of coronavirus
530 disease 2019 (COVID-19) in China. *Influenza Other Respi. Viruses* 2020: 1-9.
531 30. Norris JR. Markov Chains (Cambridge Series in Statistical and Probabilistic
532 Mathematics). Cambridge: Cambridge University Press; 1997.
533

534 **SUPPLEMENTARY MATERIAL**

535 **1. Probability of a local outbreak**

536 In the Methods section, we outlined an approach for deriving the probability of a local outbreak starting from a
537 single infectious host in either the presymptomatic, symptomatic or asymptomatic classes. Here, we provide
538 more details about that derivation. The probability of a local outbreak not occurring, starting from i
539 presymptomatic hosts, j symptomatic hosts and k asymptomatic hosts, is denoted by $q_{i,j,k}$. If we consider the

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 hitting probability of the state $(0,0,0)$ starting from the state (i, j, k). The vector of hitting probabilities $q^* = \{q^*_{i,j,k} \mid (i, j, k) \in M\}$ is therefore the minimal non-negative (real) solution to the following system of equations:

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

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

where $p_{(i,j,k),(l,m,n)}$ is the transition probability from state (i, j, k) to state (l, m, n).[30] 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 $q^*_{i,j,k} \leq \hat{q}_{i,j,k}$ for all $(i, j, k) \in M$.

From this, equations (3), (4) and (5) in the main text are obtained. These equations may be reduced to a single quartic equation for $q_{0,0,1}$, yielding 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 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, 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 probability of a local outbreak occurring starting from i presymptomatic hosts, j symptomatic hosts and k asymptomatic hosts is simply $1 - q_{i,j,k}$, giving the result stated in the main text. If required, this result can be confirmed for specific model parameter values via repeated simulation of the branching process model, where the probability of a local outbreak corresponds to the proportion of simulations that progress beyond the initial stochastic phase.

2. Computation of steepest descent contours

The steepest descent contours shown in Figure 3A were computed using a gradient maximisation approach, in which at each point the contour direction was determined by minimising the local outbreak probability over a fixed search radius (Figure S1A,B). Starting from ρ_1, ρ_2, at each step we considered increasing ρ_1 by an amount $\Delta\rho_1$ and increasing ρ_2 by an amount $\Delta\rho_2$ subject to the constraint $(\Delta\rho_1)^2 + (\Delta\rho_2)^2 = r^2$, where r is a small pre-specified constant. To achieve this, we scanned over the circular arc $\Delta\rho_1 = r\cos\theta, \Delta\rho_2 = r\sin\theta$, for $\theta \in [0, \pi/2]$ (Figure S1A). In practice, this range was discretised into 33 search 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 outbreak probability evaluated at $\rho_1 +$

$\Delta\rho_1$, $\rho_2 + \Delta\rho_2$ was minimised. The process was then repeated beginning from $\rho_{1_{new}} = \rho_1 + \Delta\rho_1$, $\rho_{2_{new}} = \rho_2 + \Delta\rho_2$ (Figure S1B). The white line in Figure 3B, which divides the region in which increasing ρ_1 has a greater effect on the local outbreak probability from the region in which increasing ρ_2 has a greater effect on the local outbreak probability, was computed in an analogous way, with the additional restriction that we only considered the search directions $\theta = 0$ and $\theta = \pi/2$ (i.e. intensifying only surveillance of nonsymptomatic or symptomatic hosts).

In Figure 3C, the white line represents the contour of steepest descent under the constraint that the total change in surveillance effort ($\Delta\rho_1 + \Delta\rho_2 = S$) is held constant at each step, rather than fixing the search radius $(\Delta\rho_1)^2 + (\Delta\rho_2)^2 = r^2$ as in Figure 3A. Therefore, instead of scanning over a circular arc, at each step we scan along the line $\Delta\rho_1 = s$, $\Delta\rho_2 = S - s$, where c varies in the range $[0, S]$ (Figure S1D). Otherwise, the process is completely analogous to that described above.

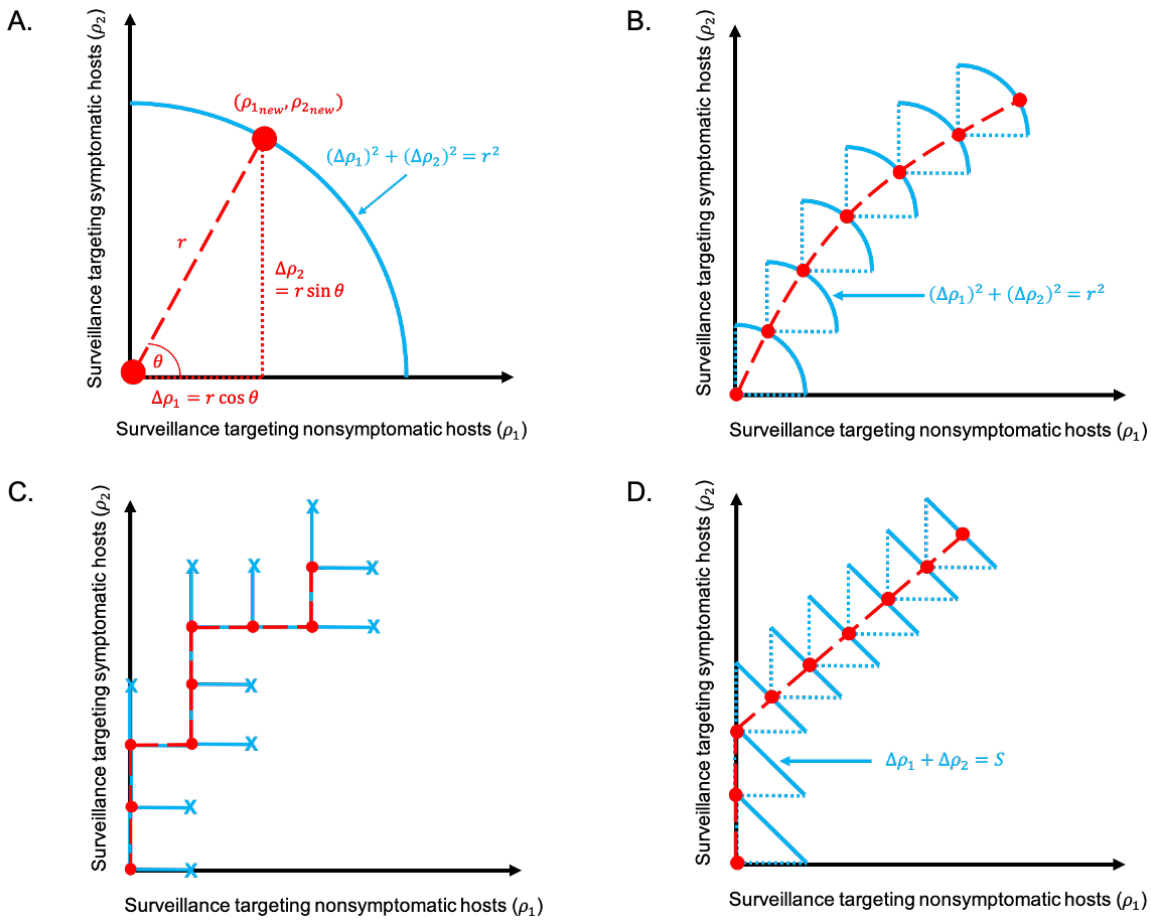


Figure S1. Computation of the steepest descent contours shown in the main text. A. To compute the steepest descent contours shown in Figure 3A of the main text, we increment ρ_1 and ρ_2 by scanning over a constant search radius $(\Delta\rho_1)^2 + (\Delta\rho_2)^2 = r^2$ (blue arc), and moving to the point $(\rho_{1_{new}}, \rho_{2_{new}})$ along that arc at which the local outbreak probability is minimised. B. The process shown in A is repeated to generate the complete contour (red dashed line). C. The analogous figure to B, in which the search direction is limited to directly to the right ($\theta = 0$) or directly upwards ($\theta = \pi/2$). This procedure is used to generate the contour in Figure 3B in the main text. D. The analogous figure to B, in which the total change in surveillance effort ($\Delta\rho_1 + \Delta\rho_2 = S$) is held constant at each step, rather than the search radius. This procedure is used to generate the contour in Figures 3C and D in the main text.

3. Robustness of results to parameter values used

In this section, we demonstrate how our results are affected by varying the parameters from their baseline values given in Table 1. We performed sensitivity analyses on the values of $R_0, \xi, K_p, K_a, \gamma + \mu, \epsilon, \lambda, \nu$ and δ. For each of these, we present plots analogous to Figure 3D for six different values of the relevant parameter. In each case we considered, our qualitative message was unchanged – for any reduction in the local outbreak probability below a particular threshold value, the optimal strategy involved surveillance targeting both nonsymptomatic and symptomatic individuals.

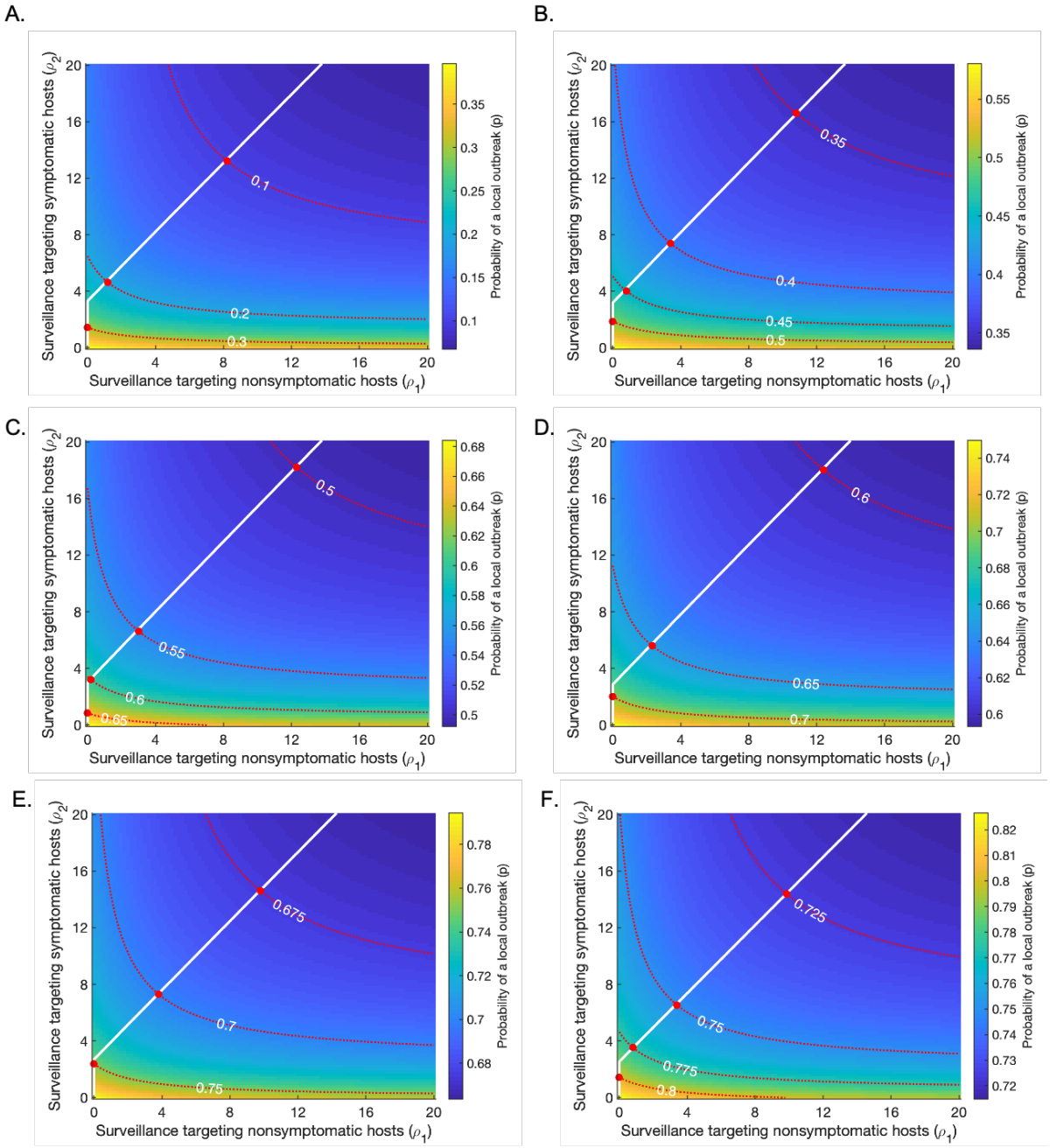


Figure S2. Varying the basic reproduction number R_0 from its baseline value ($R_0 = 3$). Plots are analogous to Figure 3D, showing strategies for minimising the surveillance effort required to achieve a pre-specified risk level (an "acceptable" local outbreak probability). Red dotted lines represent 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. $R_0 = 1 \cdot 5$. B. $R_0 = 2$. C. $R_0 = 2 \cdot 5$. D. $R_0 = 3$ (baseline). E. $R_0 = 3 \cdot 5$. F. $R_0 = 4$.

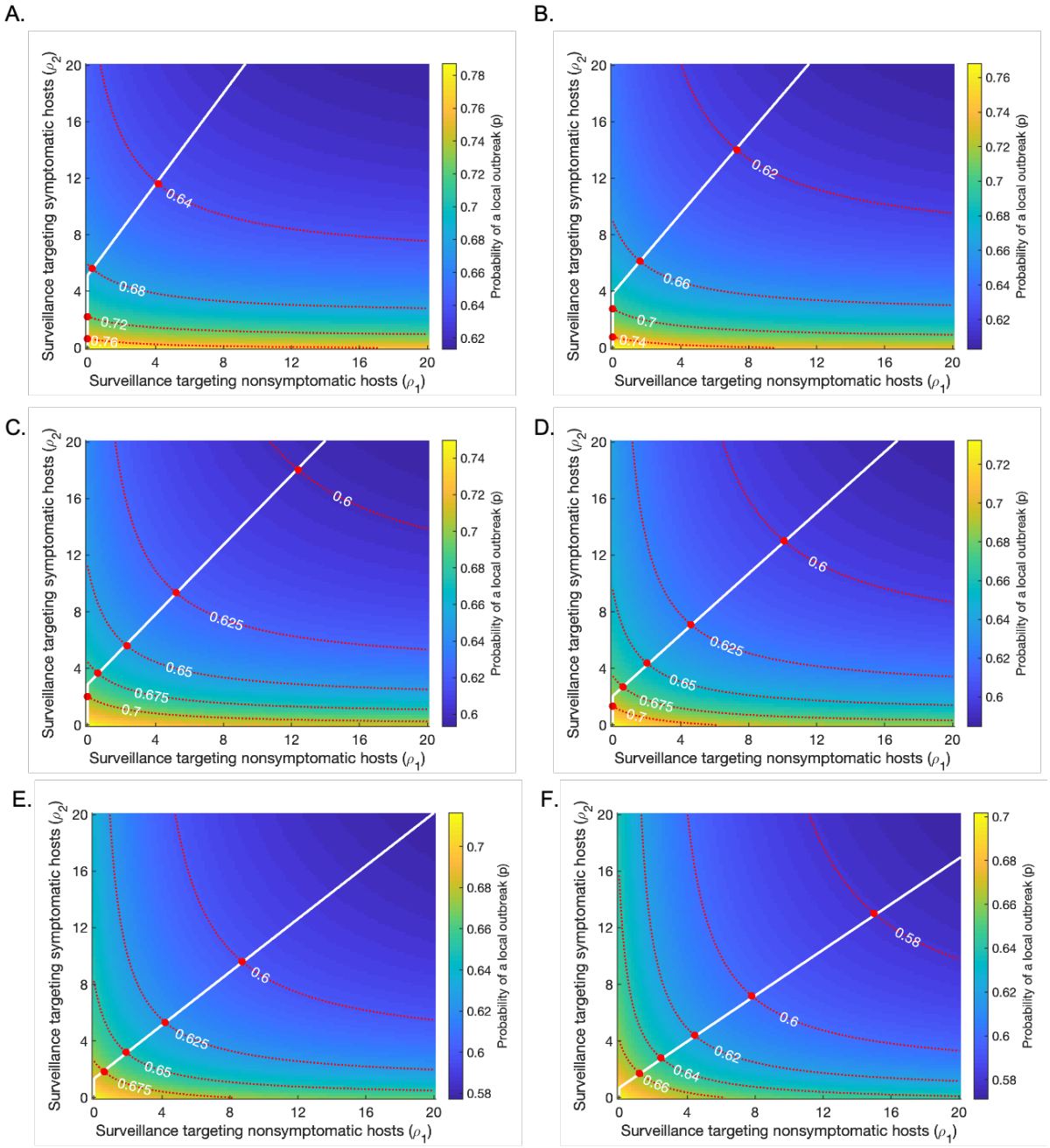


Figure S3. Varying the proportion of infections from asymptomatic infectors, ξ, from its baseline value ($\xi = 0.2$). Plots are analogous to Figure 3D, showing strategies for minimising the surveillance effort required to achieve a pre-specified risk level (an "acceptable" local outbreak probability). 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 ξ and β (which is changed in each panel to set $R_0 = 3$), all parameters are held fixed at their baseline values given in Table 1. A. $\xi = 0$. B. $\xi = 0 \cdot 1$. C. $\xi = 0 \cdot 2$ (baseline). D. $\xi = 0 \cdot 3$. E. $\xi = 0 \cdot 4$. F. $\xi = 0 \cdot 5$.

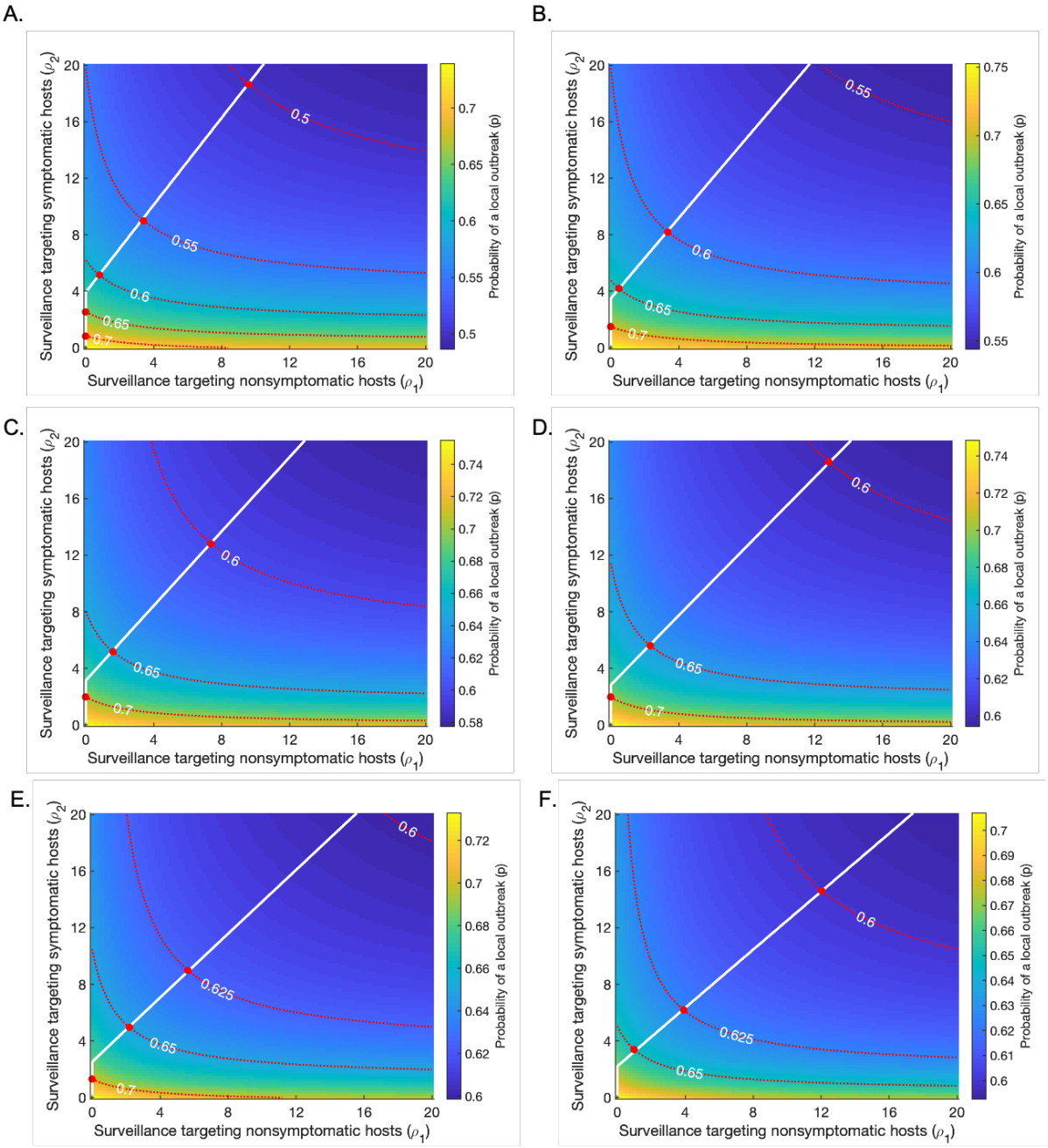


Figure S4. Varying the proportion of infections arising from presymptomatic hosts in the absence of intensified surveillance (K_p, given by expression (1) in the main text) from its baseline value ($K_p = 0.489$). In each case, the proportions of infections arising from asymptomatic and symptomatic hosts are adjusted so that they remain in the same ratio as in the baseline case. Plots are analogous to Figure 3D, 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 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 K_a), and β (which is

changed in each panel to set $R_0 = 3$), all parameters are held fixed at their baseline values given in Table 1. A. $K_p = 0 \cdot 2$. B. $K_p = 0 \cdot 3$. C. $K_p = 0 \cdot 4$. D. $K_p = 0 \cdot 5$. E. $K_p = 0 \cdot 6$. F. $K_p = 0 \cdot 7$.

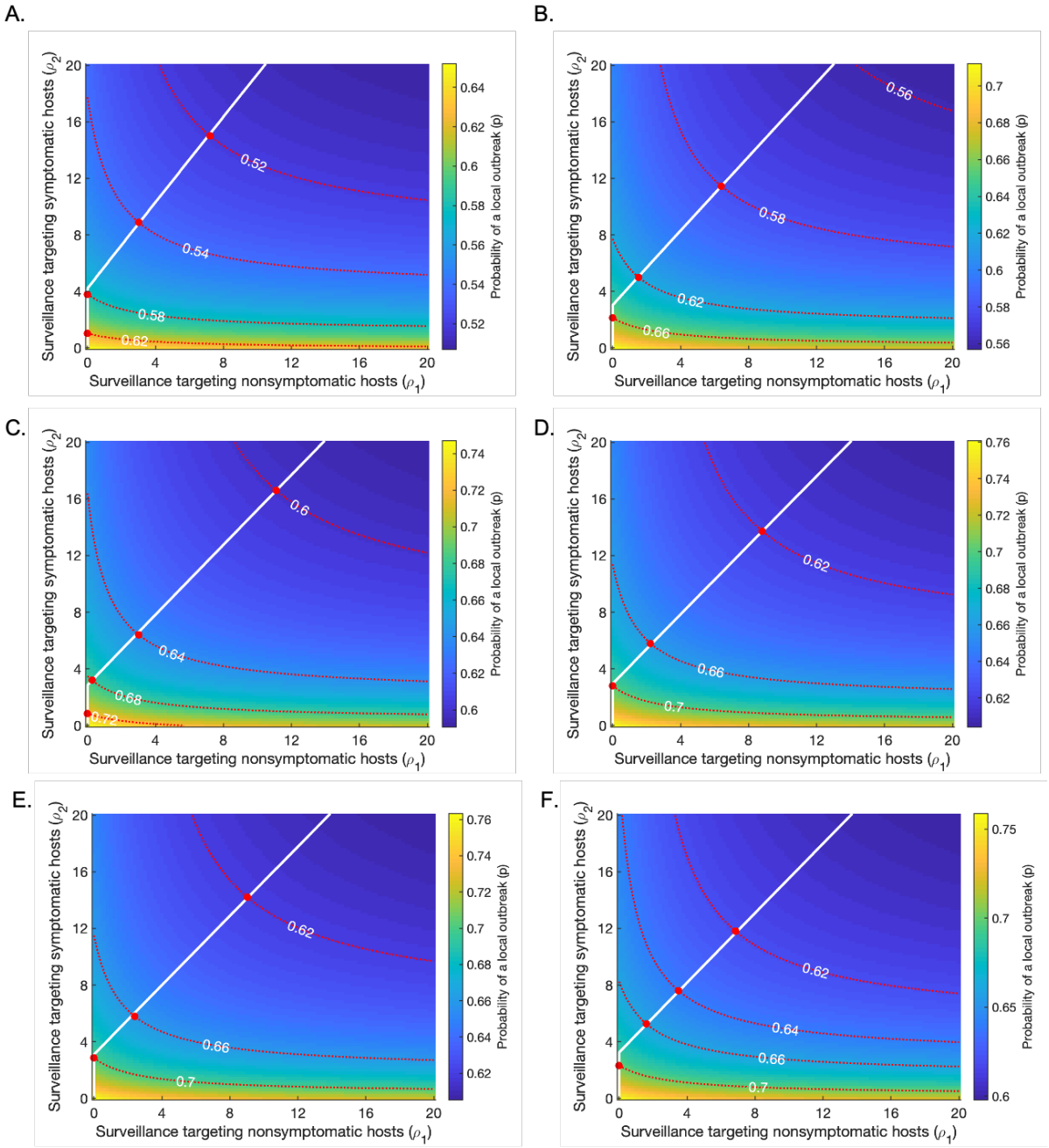


Figure S5. Varying the proportion of infections arising from asymptomatic hosts in the absence of intensified surveillance (K_a, given by expression (2) in the main text) from its baseline value ($K_a = 0.106$). In each case, the proportions of infections arising from presymptomatic and symptomatic hosts are adjusted so that they remain in the same ratio as in the baseline case. Plots are analogous to Figure 3D, 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 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 K_a), and β (which is changed in each panel to set $R_0 = 3$), all parameters are held fixed at their baseline values given in Table 1. A. $K_a = 0 \cdot 01$. B. $K_a = 0 \cdot 05$. C. $K_a = 0 \cdot 1$. D. $K_a = 0 \cdot 15$. E. $K_a = 0 \cdot 2$. F. $K_a = 0 \cdot 25$.

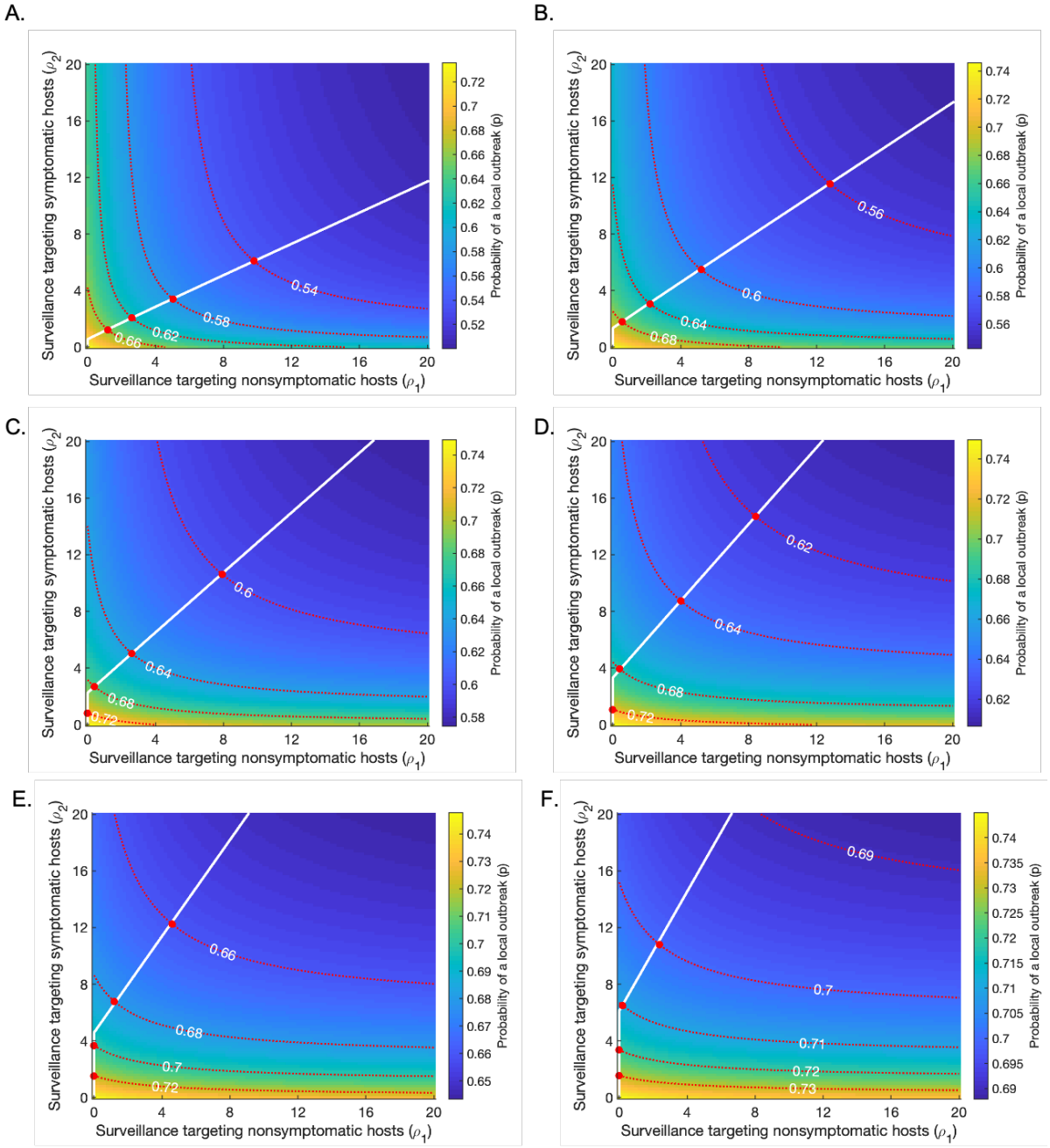
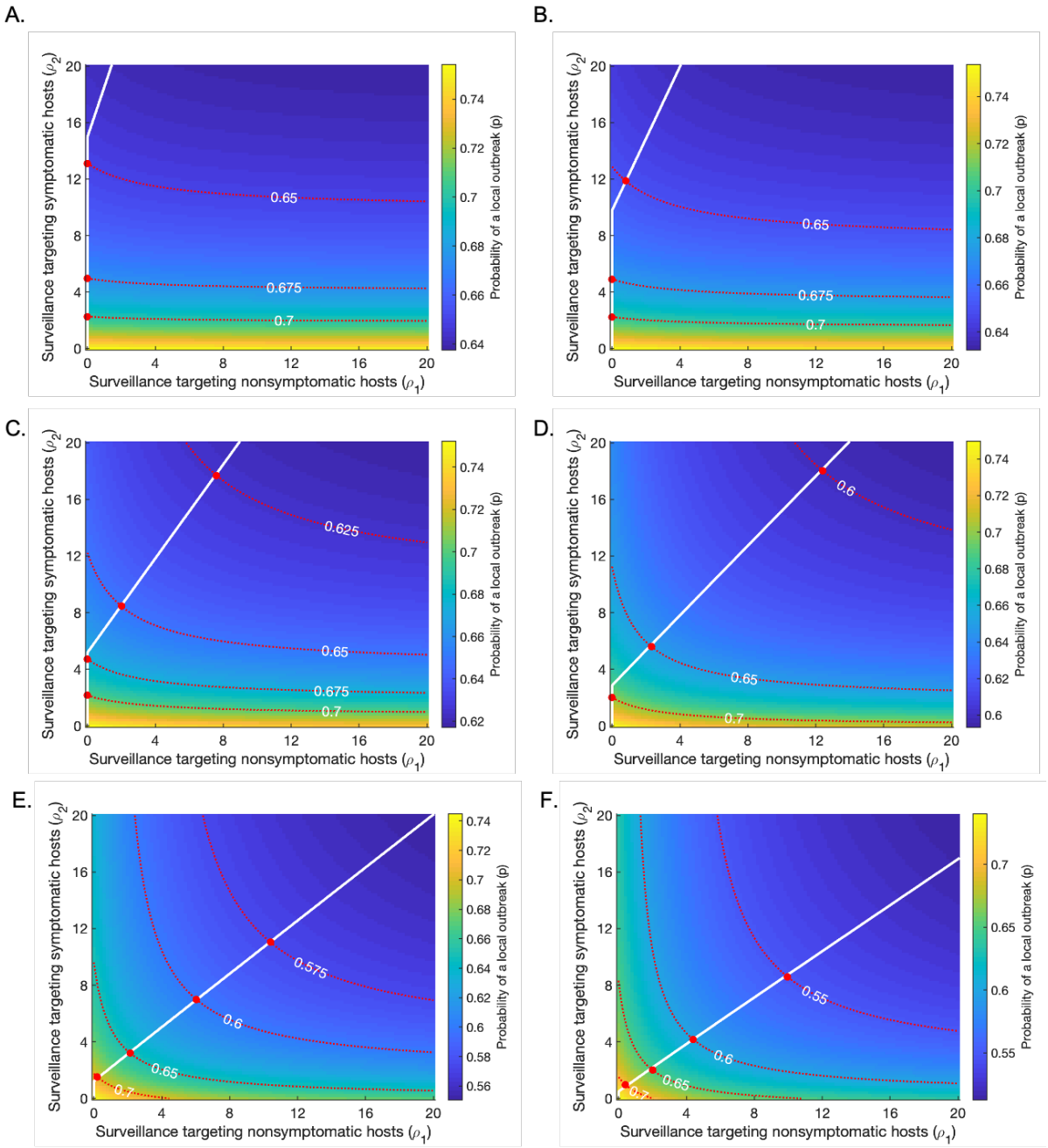


Figure S6. Varying the expected time period to isolation conditional on isolation occurring during the symptomatic period, $1/(\gamma + \mu)$, from its baseline value ($1/(\gamma + \mu) = 4.6$ days). This is achieved by varying the parameter γ, whilst holding the recovery rate of symptomatic individuals μ equal to its baseline value ($\mu = 1/8$ days^{-1}). Plots are analogous to Figure 3D, showing strategies for minimising the surveillance effort required to achieve a pre-specified risk level (an "acceptable" local outbreak probability). Red dotted lines

652 represent contours along which the probability of a local outbreak is constant, as labelled; red circles indicate
653 the points along these contours at which the total surveillance effort $\rho_1 + \rho_2$ is minimised. The white line
654 indicates the optimal strategy to follow if the pre-specified risk level is reduced. Apart from γ and β (which is
655 changed in each panel to set $R_0 = 3$), all parameters are held fixed at their baseline values given in Table 1. A.
656 $1/(\gamma + \mu) = 2$ days. B. $1/(\gamma + \mu) = 3$ days. C. $1/(\gamma + \mu) = 4$ days. D. $1/(\gamma + \mu) = 5$ days. E. $1/(\gamma + \mu) =$
657 6 days. F. $1/(\gamma + \mu) = 7$ days.
658



659

660 Figure S7. Varying ϵ, the relative isolation rate of nonsymptomatic individuals without intensified surveillance
661 (compared to symptomatic individuals), from its baseline value ($\epsilon = 0.1$). Plots are analogous to Figure 3D,
662 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 ϵ and β (which is changed in each panel to set $R_0 = 3$), all parameters are held fixed at their baseline values given in Table 1. A. $\epsilon = 0 \cdot 01$. B. $\epsilon = 0 \cdot 02$. C. $\epsilon = 0 \cdot 05$. D. $\epsilon = 0 \cdot 1$ (baseline). E. $\epsilon = 0 \cdot 2$. F. $\epsilon = 0 \cdot 3$.

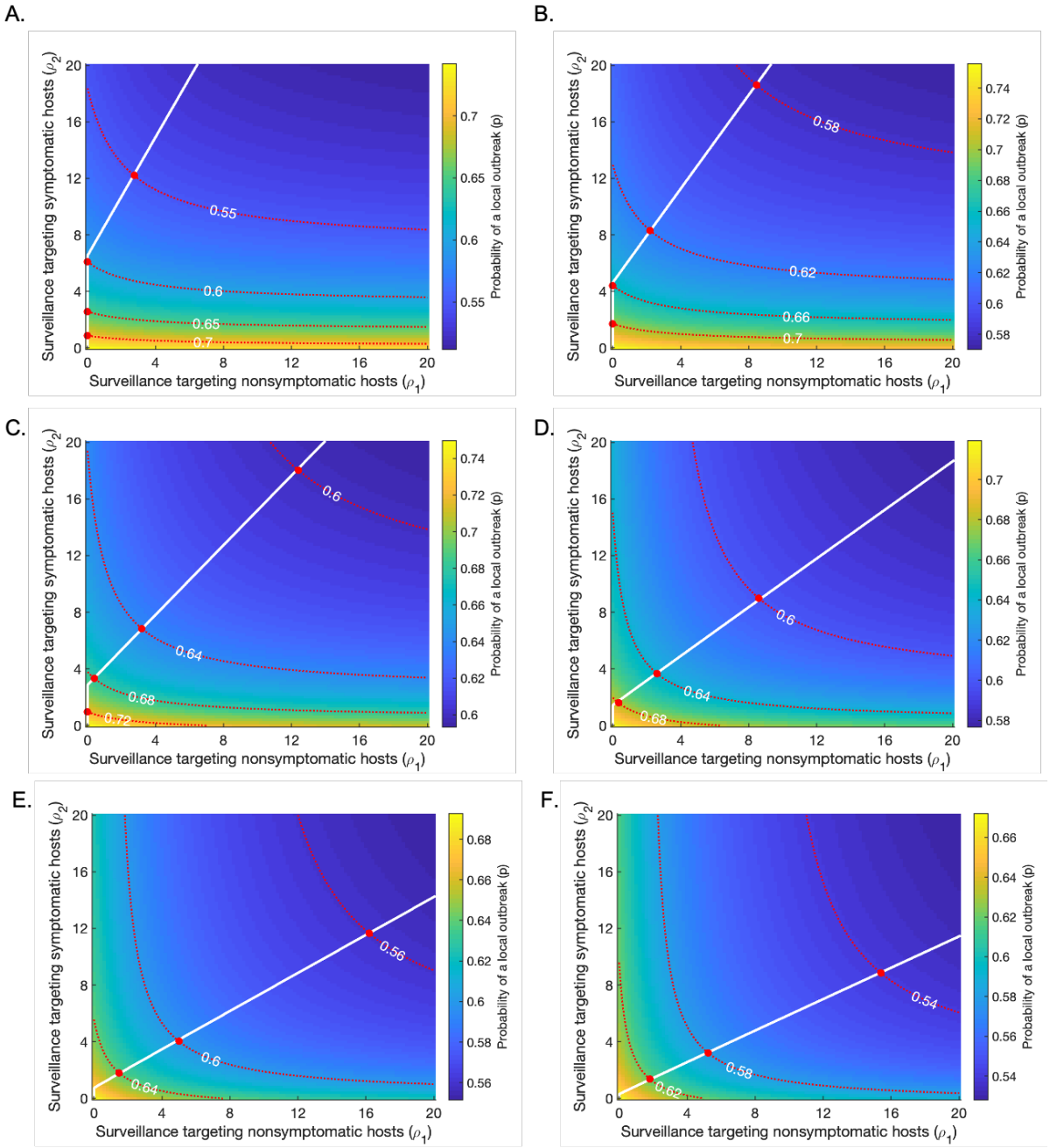


Figure S8. Varying the duration of the presymptomatic period, $1/\lambda$, from its baseline value ($1/\lambda = 0.5$ days). Plots are analogous to Figure 3D, showing strategies for minimising the surveillance effort required to achieve a pre-specified risk level (an "acceptable" local outbreak probability). 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 λ and β (which is changed in each panel to set $R_0 = 3$), all parameters are held fixed at their baseline values given in Table 1. A. $1/\lambda = 0.5$ 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.

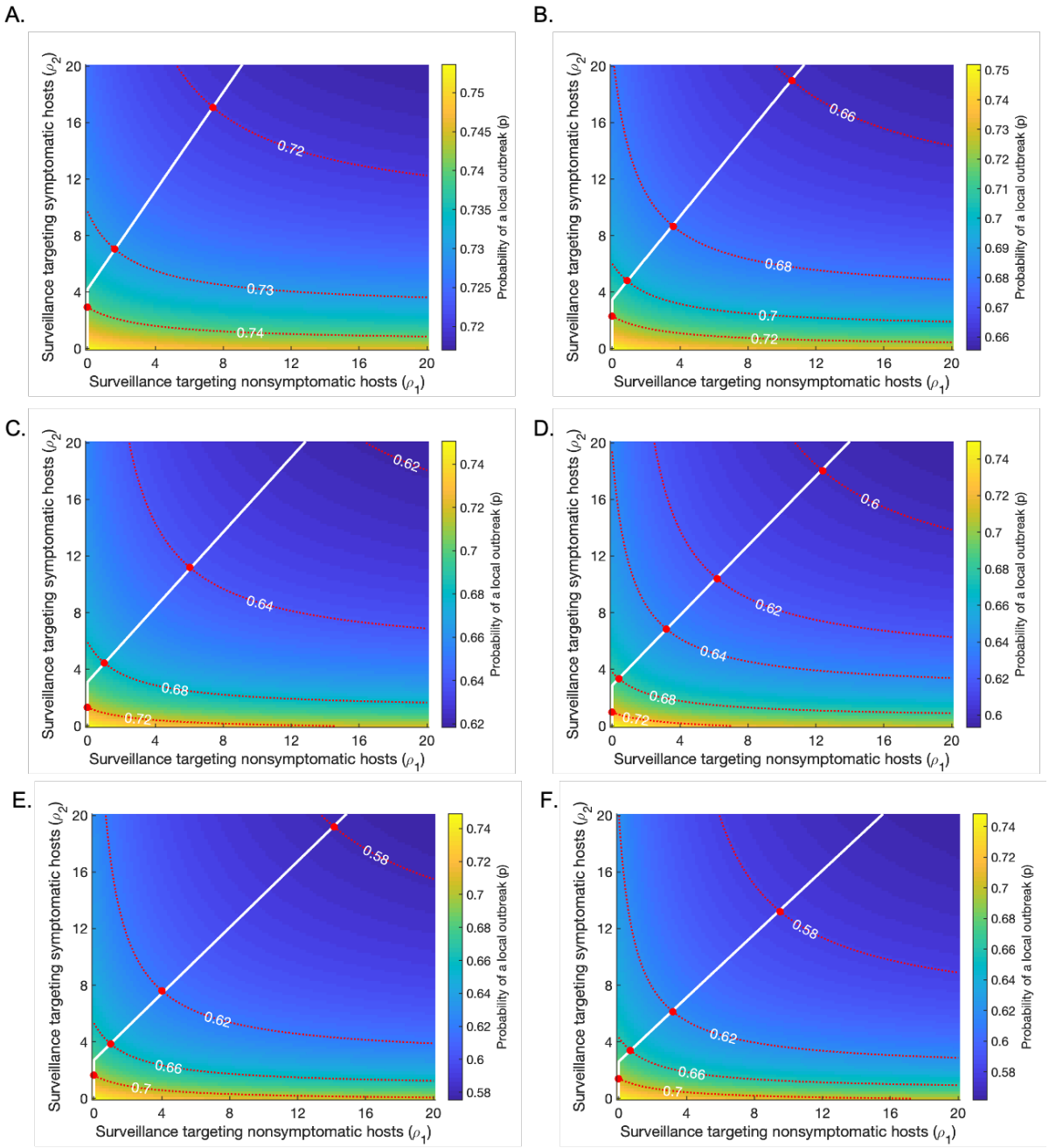


Figure S9. Varying the duration of the symptomatic period, $1/\mu$, from its baseline value ($1/\mu = 8$ days). The parameter γ is varied simultaneously such that $1/(\gamma + \mu)$, the expected time period to isolation conditional on isolation occurring during the symptomatic period, remains at its baseline value ($1/(\gamma + \mu) = 4.6$ days). Plots are analogous to Figure 3D, 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 μ and γ, and β (which is changed in each panel to set $R_0 = 3$), all parameters are held fixed at their baseline values given in Table 1. A. $1/\mu = 5$ 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.

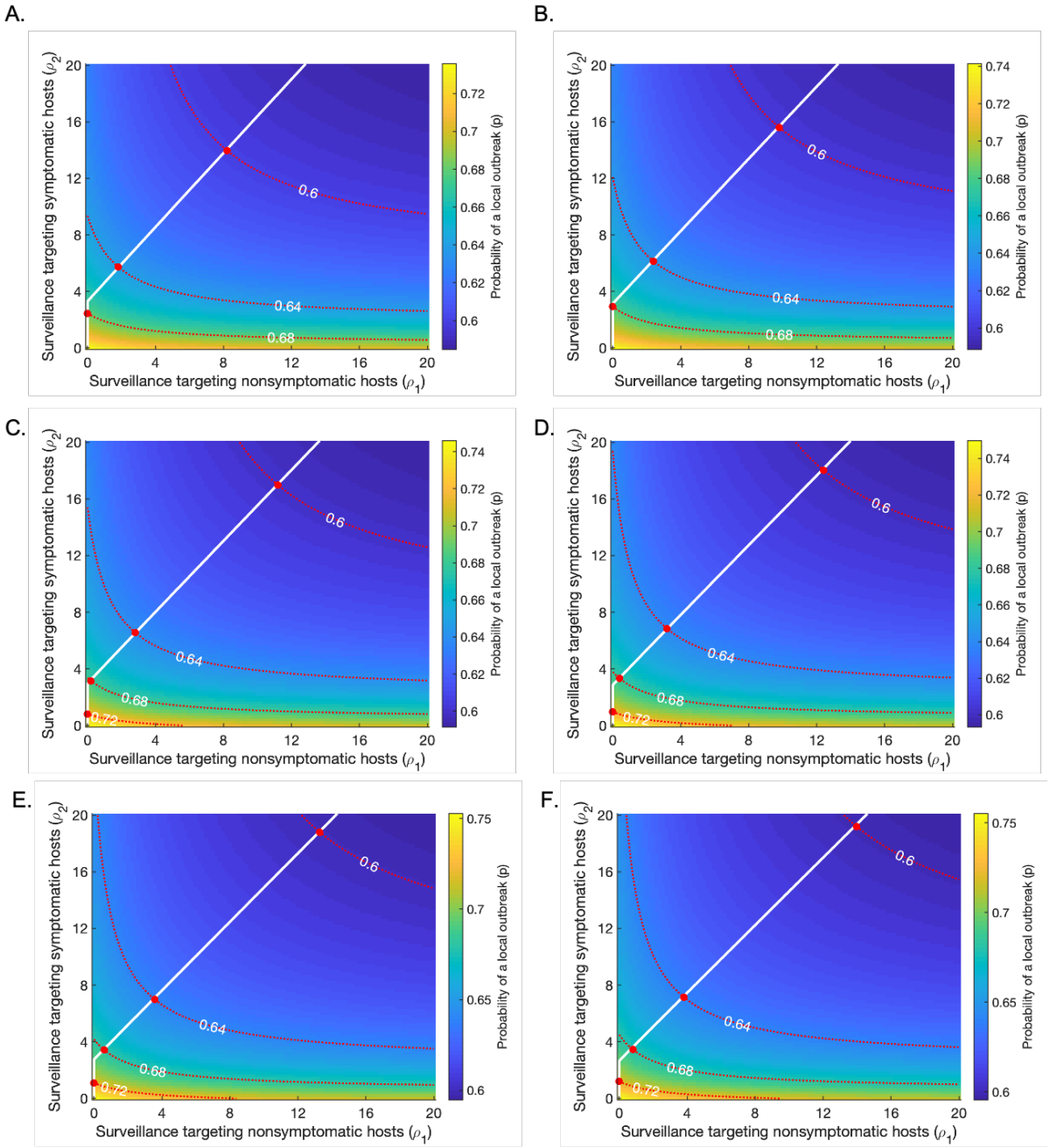


Figure S10. Varying $1/\nu$, the duration of the asymptomatic period, from its baseline value ($1/\nu = 10$ days). Plots are analogous to Figure 3D, showing strategies for minimising the surveillance effort required to achieve a pre-specified risk level (an "acceptable" local outbreak probability). Red 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 ν and β (which is changed in each panel to set $R_0 = 3$), all parameters are held fixed at their baseline values given in Table 1. A. $1/\nu = 7$ 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.

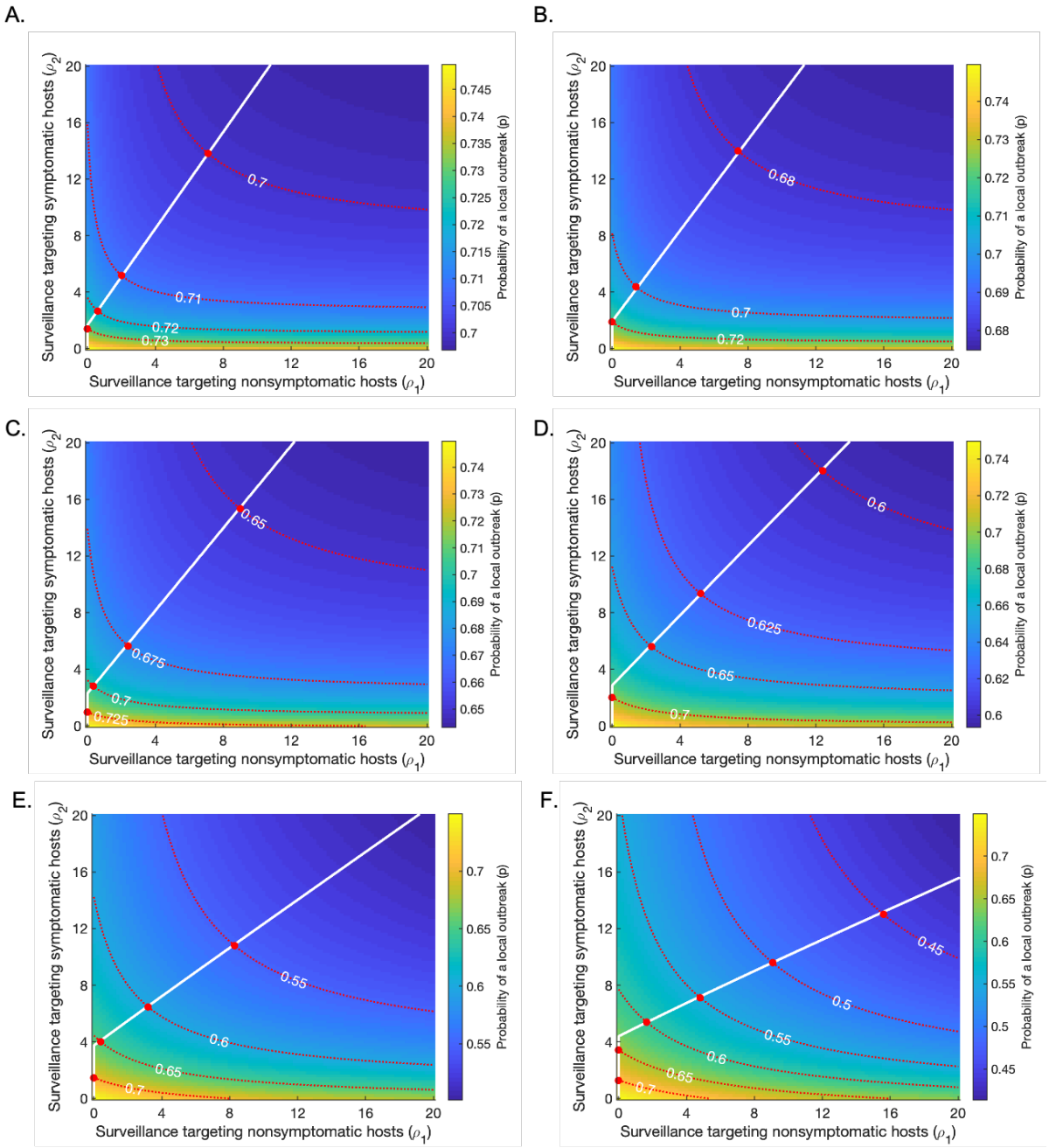


Figure S11. Varying the upper bound on the fractional reduction in the time to isolation (if no other event occurs), δ, from its baseline value ($\delta = 0.8$). Plots are analogous to Figure 3D, showing strategies for minimising the surveillance effort required to achieve a pre-specified risk level (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 δ, all parameters are held fixed at their baseline values given in Table 1. A. $\delta = 0 \cdot 5$. B. $\delta = 0 \cdot 6$. C. $\delta = 0 \cdot 7$. D. $\delta = 0 \cdot 8$ (baseline). E. $\delta = 0 \cdot 9$. F. $\delta = 0 \cdot 95$.