

# Effects of individual variation and seasonal vaccination on disease risks

Received: 24 March 2025

Accepted: 15 August 2025

Published online: 26 September 2025

 Check for updates

William S. Hart<sup>1</sup>✉, Jina Amin<sup>1</sup>, Hyeongki Park<sup>2,3</sup>, Kosaku Kitagawa<sup>3</sup>, Yong Dam Jeong<sup>3,4</sup>, Alexander R. Kaye<sup>5,6</sup>, Shingo Iwami<sup>3,7,8,9,10,11</sup> & Robin N. Thompson<sup>1</sup>

Estimates of the risk of a large outbreak resulting from pathogen introduction into a population are valuable for planning interventions. Two key factors affecting outbreak risks are variation in transmission between individuals (e.g., superspreading individuals) and change over time (e.g., through seasonality or changing population immunity due to vaccination). Here, we develop an outbreak risk estimation framework that accounts for both features simultaneously. To demonstrate the real-world application of our framework, we consider the design of annual COVID-19 booster vaccination campaigns, using a multi-scale approach incorporating an individual-level model of vaccine-induced antibody dynamics. Near the start of annual vaccine distribution, when population immunity is low, a high outbreak risk is possible; this can be mitigated by distributing vaccines over a longer period. We show that longer distribution periods are particularly beneficial if vaccine coverage and/or effectiveness is high, and if seasonality in transmission is limited.

Following the introduction of a pathogen into a host population, a key question for public health policy makers is whether a large outbreak is likely to occur or whether the pathogen will fade out after only a few infections<sup>1</sup>. To this end, mathematical models are widely used to estimate the outbreak risk: the probability of a single incident infection resulting in a large outbreak<sup>1</sup>. For example, the risk of sustained transmission in new locations was analysed following the initial emergence of the SARS-CoV-2 virus in China<sup>2–5</sup>. Subsequent studies considered the risk of local COVID-19 outbreaks<sup>6–9</sup>, including assessing the effectiveness of public health measures such as mass testing<sup>6</sup> and vaccination<sup>7</sup>.

A key transmission feature of many pathogens<sup>10,11</sup> including SARS-CoV-2<sup>12–15</sup> and Ebola virus<sup>16</sup> is substantial heterogeneity in transmission between different infected individuals. Because of variation in virus shedding<sup>6,17</sup> and/or contact patterns<sup>13</sup>, a relatively small proportion of

infected individuals may generate a large proportion of transmissions (superspreading; Supplementary Fig. 1). A seminal study by Lloyd-Smith et al.<sup>11</sup> showed how such individual variation can be included in outbreak risk estimates. For a specified reproduction number, the outbreak risk is typically lower when the extent of heterogeneity in infectiousness is greater, since (for example) a randomly chosen initial infected individual is then more likely to generate no transmissions at all (thus causing the outbreak to fade out immediately)<sup>11</sup>. However, when heterogeneity is greater, outbreaks that do occur are often more explosive, since superspreading events can cause rapid outbreak growth<sup>11</sup>.

Another important feature considered in studies estimating outbreak risks is temporal variation in transmission<sup>7,18–21</sup>, for example, arising due to seasonality<sup>18,21</sup> or the rollout of vaccines over time<sup>7</sup>. In the presence of such temporal variation, the naïve approach of estimating

<sup>1</sup>Wolfson Centre for Mathematical Biology, Mathematical Institute, University of Oxford, Oxford, UK. <sup>2</sup>School of Biomedical Convergence Engineering, Pusan National University, Busan, South Korea. <sup>3</sup>interdisciplinary Biology Laboratory (iBLab), Division of Natural Science, Graduate School of Science, Nagoya University, Nagoya, Japan. <sup>4</sup>Department of Mathematics, College of Natural Sciences, Pusan National University, Busan, South Korea. <sup>5</sup>Mathematics Institute, University of Warwick, Coventry, UK. <sup>6</sup>Zeeman Institute for Systems Biology and Infectious Disease Epidemiology Research (SBIDER), University of Warwick, Coventry, UK. <sup>7</sup>Institute of Mathematics for Industry, Kyushu University, Fukuoka, Japan. <sup>8</sup>Institute for the Advanced Study of Human Biology (ASHBi), Kyoto University, Kyoto, Japan. <sup>9</sup>Interdisciplinary Theoretical and Mathematical Sciences Program (iTHEMS), RIKEN, Saitama, Japan. <sup>10</sup>NEXT-Ganken Program, Japanese Foundation for Cancer Research (JFCR), Tokyo, Japan. <sup>11</sup>Science Groove Inc., Fukuoka, Japan. ✉ e-mail: [william.hart@maths.ox.ac.uk](mailto:william.hart@maths.ox.ac.uk)

the outbreak risk based on instantaneous transmission conditions (neglecting changes that occur following pathogen introduction) can lead to inaccurate outbreak risk estimates<sup>18</sup>. Accurate outbreak risk quantification, therefore, requires rigorous approaches that account for changes in transmission soon after the pathogen arrives in the host population. However, individual heterogeneity in transmission has not previously (to our knowledge) been considered alongside temporal variability when calculating the outbreak risk.

Here, we develop a general framework for calculating outbreak risks under time-dependent and heterogeneous transmission. Specifically, we consider a generalised renewal equation model that accounts for heterogeneous individual infectiousness, and derive equations satisfied by the outbreak risk. This framework is intended to be widely applicable, and we start by considering a simple illustrative example of its use to calculate seasonal outbreak risks. To demonstrate the application of our approach in a specific setting, we then consider the risk of localised COVID-19 outbreaks as a case study. Substantial SARS-CoV-2 transmission continues to occur worldwide, driven by viral evolution<sup>22</sup> and waning immunity<sup>23–25</sup>. Seasonal booster vaccinations are therefore offered in many countries, particularly to individuals at a higher clinical risk of severe infection outcomes<sup>26</sup>. We consider the COVID-19 outbreak risk under seasonal transmission<sup>27,28</sup> and annual booster vaccination, utilising a multi-scale approach that

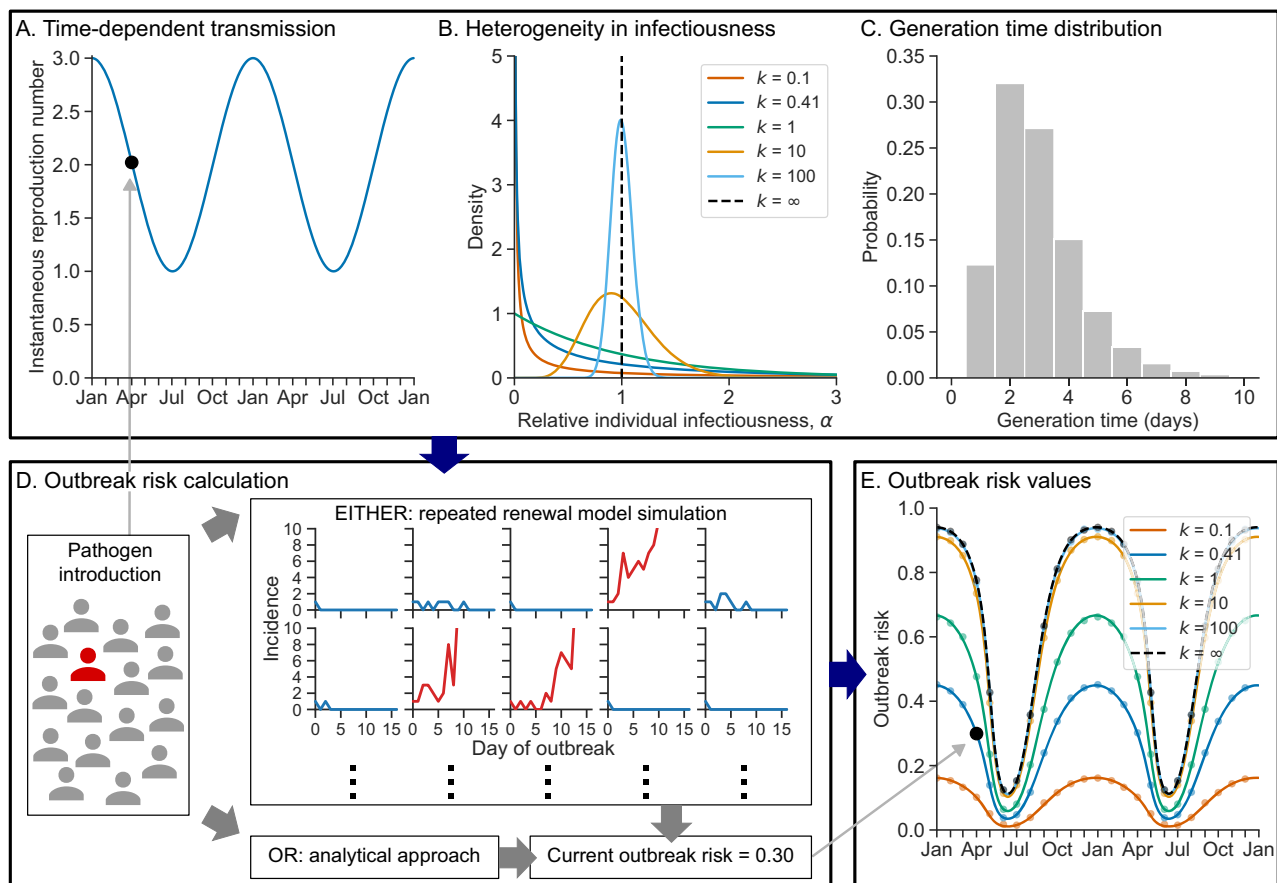
incorporates an individual-level model of antibody dynamics following vaccination<sup>29–31</sup>. We focus on how the timing of annual vaccination distribution can be optimised to minimise the outbreak risk, finding that booster vaccine coverage and effectiveness, as well as the extent of seasonality in transmission, are key determinants of the optimal time of year (and duration of the year) for vaccines to be deployed. This exemplifies the practical conclusions that can be obtained from our general and extensible framework for outbreak risk quantification.

## Results

### Outbreak risks with time-dependent and heterogeneous transmission

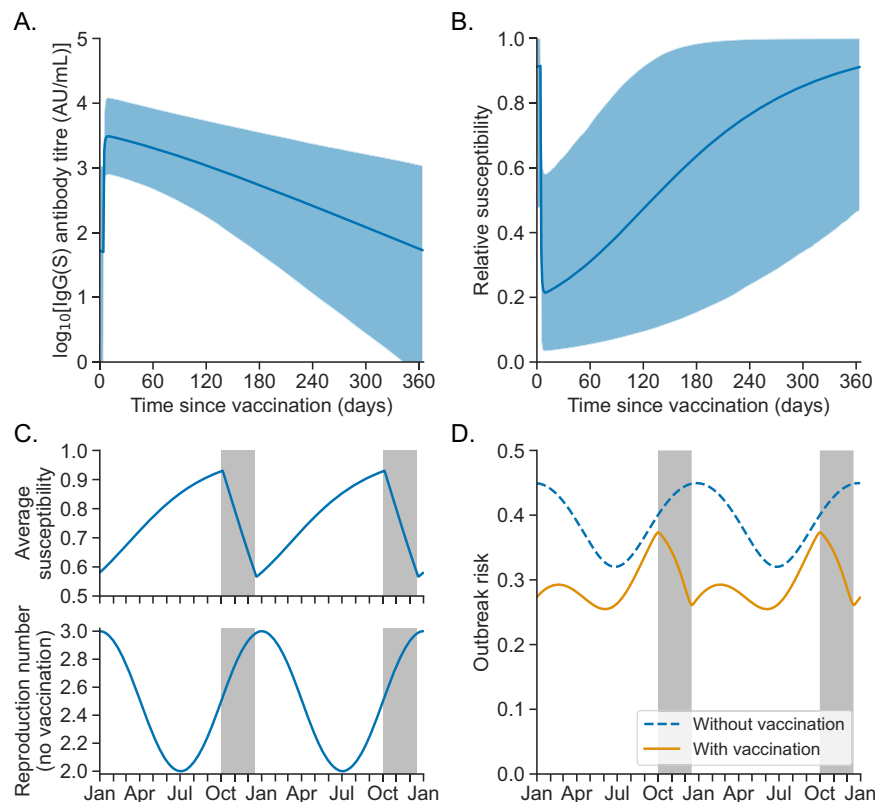
We have developed a framework for calculating the time-dependent outbreak risk (the probability that a large outbreak results from the introduction of a single newly infected individual into the population on a given date), accounting for both temporal variation in transmission and heterogeneity in infectiousness between infected individuals. In Fig. 1, we demonstrate our method for a simple illustrative example.

Our method uses three inputs (Fig. 1A–C). First, time-dependent transmission is characterised by the instantaneous reproduction number,  $R_t$  (Fig. 1A; here,  $R_t$  is a periodic function of time,  $t$ , but other possibilities could be considered in our general framework). Second, similar to the approach of ref. 11, the extent of heterogeneity in



**Fig. 1 | Outbreak risks with time-dependent and heterogeneous transmission.** **A–C** Inputs to the outbreak risk calculation. **A** Instantaneous reproduction number,  $R_t$ , varying periodically over calendar time,  $t$ . **B** Heterogeneity in infectiousness, as characterised by a gamma distribution of relative individual infectiousness factors with mean 1 and dispersion (shape) parameter  $k$  (with a lower value of  $k$  corresponding to a greater degree of heterogeneity). This distribution is shown for a range of values of  $k$  ( $k = 0.1$ : red;  $k = 0.41$ : blue;  $k = 1$ : green;  $k = 10$ : orange;  $k = 100$ : light blue;  $k = \infty$ : black dashed). **C** Generation time distribution<sup>32</sup>. **D** Illustration of how the outbreak risk at a given calendar time can be calculated either through

repeated stochastic simulation of a generalised renewal equation model (in this case, the outbreak risk is defined to be the proportion of simulations in which the daily incidence of new infections ever exceeds a specified threshold, here taken to be 30 daily infections; red curves show example simulations in which incidence exceeds the threshold, and blue curves show simulations in which the threshold is never exceeded), or by solving a system of analytically derived equations numerically. **E** Values of the outbreak risk,  $p_t$ , for the inputs shown in (A–C) (considering the same values of  $k$  as in (B), with the same colour key), for both the analytical approach (lines) and simulation-based approach (dots).



**Fig. 2 | Multi-scale modelling of the COVID-19 outbreak risk under annual booster vaccination.** **A** Log IgG(S) antibody titre relative to time since (most recent) vaccine dose. The curve shows the mean log titre (calculated over model-simulated antibody dynamics profiles for a synthetic cohort of 10,000 individuals), and the shaded region represents pointwise 95% prediction intervals (PIs) for individual antibody titres. **B** Relative susceptibility to SARS-CoV-2 infection following vaccination (compared to unvaccinated individuals, who are assumed to have relative susceptibility 1), which was assumed to be a function of individual antibody titre. The curve shows the mean susceptibility, and the shaded region

represents pointwise 95% PIs for individual susceptibility. **C** Top: population-average susceptibility,  $\eta_t$ , as a function of calendar time,  $t$ , assuming 60% of the population receives booster vaccines between 1 October and 15 December each year (this time period is shown in the grey shaded regions). Bottom: instantaneous reproduction number,  $R_{0,t}$ , in the absence of booster vaccination. The instantaneous reproduction number when booster vaccination takes place,  $R_t = \eta_t R_{0,t}$ , is shown in Supplementary Fig. 2. **D** Outbreak risk,  $p_t$ , following the introduction of a single newly infected individual on day  $t$ , both without (blue dashed curve) and with (orange curve) booster vaccination, calculated using our analytical approach.

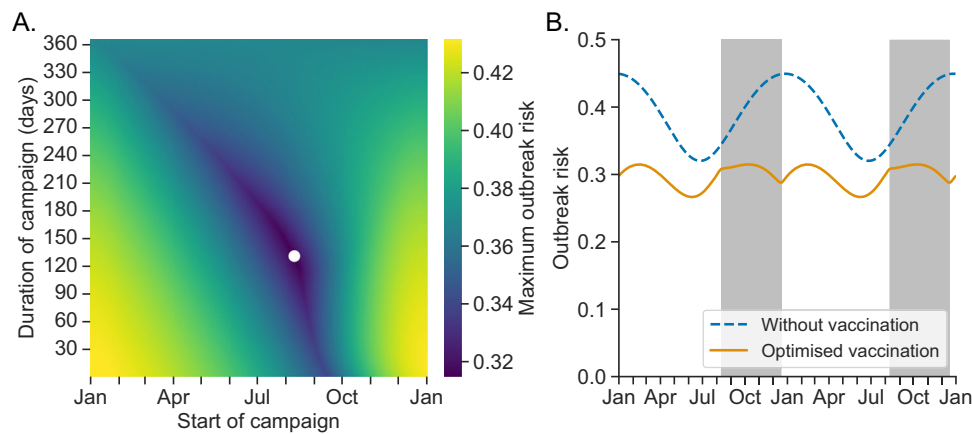
transmission between individuals is characterised by the dispersion parameter,  $k$ , which is the shape parameter of a gamma distribution of relative individual infectiousness factors—an infected individual with infectiousness factor  $\alpha$  is expected (depending on stochasticity in the transmission process) to generate  $\alpha$  times as many transmissions as a typical individual infected on the same day. A lower value of  $k$  corresponds to a greater degree of heterogeneity between individuals (Fig. 1B and Supplementary Fig. 1). This formulation results in a negative binomial distribution for the number of transmissions (offspring) generated by an infected host, with dispersion parameter  $k$ <sup>11</sup>. The individual infectiousness factor distribution is shown for a range of values of  $k$  in Fig. 1B ( $k = 0.41$  is the pooled estimate for COVID-19 obtained in a previous meta-analysis<sup>15</sup>). Finally, the generation time distribution (the distribution of intervals between infection dates of infector–infectee transmission pairs) is also required as an input to our framework. The generation time distribution used throughout our analyses, which is a discretised version of a continuous-time estimate for the omicron SARS-CoV-2 variant<sup>32</sup>, is shown in Fig. 1C.

Based on the specified inputs, the outbreak risk can be estimated via repeated simulation of a generalised renewal equation model (accounting for heterogeneity in infectiousness), calculated as the proportion of simulations in which the daily incidence of new infections ever exceeds a specified threshold (Fig. 1D). Alternatively, a system of equations satisfied by the outbreak risk at different time points can be derived analytically under the same transmission model and then solved numerically.

In Fig. 1E, the time-dependent outbreak risk (calculated using either the simulation-based or analytical approach) is shown for the inputs in Fig. 1A–C, again considering a range of values of the dispersion parameter,  $k$ . The simulation and analytical approaches give very similar outbreak risk values; therefore, only the analytical approach is considered in the remainder of our analyses (however, we note that the simulation-based approach may be needed in scenarios where  $R_t < 1$  for much of the year—see Discussion). Similar to ref. 11 (in which temporal variation in transmission was not considered), we found that the outbreak risk on any given date is lower when there is a greater extent of heterogeneity in infectiousness (i.e., for smaller values of  $k$ ).

### COVID-19 outbreak risk under annual booster vaccination

To show how our outbreak risk estimation approach can be applied in a more complex scenario than the general example shown in Fig. 1, we then considered temporal variation in the risk of localised COVID-19 outbreaks, and the impact of annual booster vaccination campaigns for mitigating this risk (Fig. 2). In our baseline analysis, we considered parameter choices representative of a hypothetical infectious novel SARS-CoV-2 variant and a vulnerable population with relatively high vaccine uptake (sensitivity analyses to model inputs are described later). We took a multi-scale approach, using a previously parameterised<sup>31</sup> individual-level antibody dynamics model<sup>29,30</sup> (Fig. 2A) to infer temporal variation in relative susceptibility to infection (compared to an individual who does not receive booster vaccines) following booster vaccination (Fig. 2B). This enabled us to estimate the



**Fig. 3 | Optimisation of the timing of booster vaccination campaigns.** **A** Annual maximum outbreak risk, plotted against the start date and duration of annual booster vaccine distribution (assuming 60% of the population is vaccinated each year). The white dot indicates the start date and duration at which the annual

maximum outbreak risk is minimised. **B** Outbreak risk,  $p_t$ , without booster vaccination (blue dashed curve) and with optimised booster vaccine distribution timing (orange curve); the optimal distribution period each year is shown in the grey shaded regions).

population-averaged susceptibility at each calendar time, assuming in our baseline analysis that 60% of individuals (randomly chosen in the population) receive a vaccine dose each autumn/winter (Fig. 2C, top panel), with an equal number of individuals vaccinated each day of the booster campaign (1 October to 15 December each year). We combined these susceptibility values with a profile of seasonal transmissibility in the absence of booster vaccination (Fig. 2C, bottom panel) to obtain the instantaneous reproduction number,  $R_t$ , under annual booster vaccination (Supplementary Fig. 2). We then estimated the time-dependent outbreak risk (using our analytical approach) both without and with booster vaccination (Fig. 2D), assuming  $k = 0.41^{15}$  and the generation time distribution shown in Fig. 1C.

Under our default assumptions, we found that booster vaccination reduces the outbreak risk, but does not eliminate it entirely (Fig. 2D). For example, the outbreak risk on 1 January each year (the assumed date of maximum transmissibility) is reduced by 39% by booster vaccination (from 0.45 to 0.27). On the other hand, the peak outbreak risk under booster vaccination (0.37) occurs on 2 October (one day after the start of annual vaccination) and is only 7% lower than the outbreak risk on the same date with no booster vaccination (0.40) because of limited population immunity on that date (since most individuals included in the booster vaccination campaign have not yet received their dose for the year in question).

### Optimisation of the timing of booster vaccination campaigns

We then considered how our multi-scale approach can be used to inform the timing of annual vaccine distribution. Specifically, we calculated the annual maximum outbreak risk for different possible start dates and durations of vaccine distribution (Fig. 3A), again under our baseline assumption that vaccine doses are received by 60% of the population each year. We found that the annual peak outbreak risk is minimised to 0.31 under a distribution window of 10 August to 18 December (compared to annual maximum values of 0.45 with no booster vaccination and 0.37 under the default distribution timing)—an earlier start date and similar end date to the default window considered in Fig. 2 (1 October to 15 December). The earlier start date ensures a lower outbreak risk when immunity is lowest (Fig. 3B) than under the default distribution timing (Fig. 2D), at the cost of a slightly higher outbreak risk near the winter transmission peak (due to waning immunity in those vaccinated earlier in the booster vaccination campaign).

### Factors determining optimal vaccine deployment timing

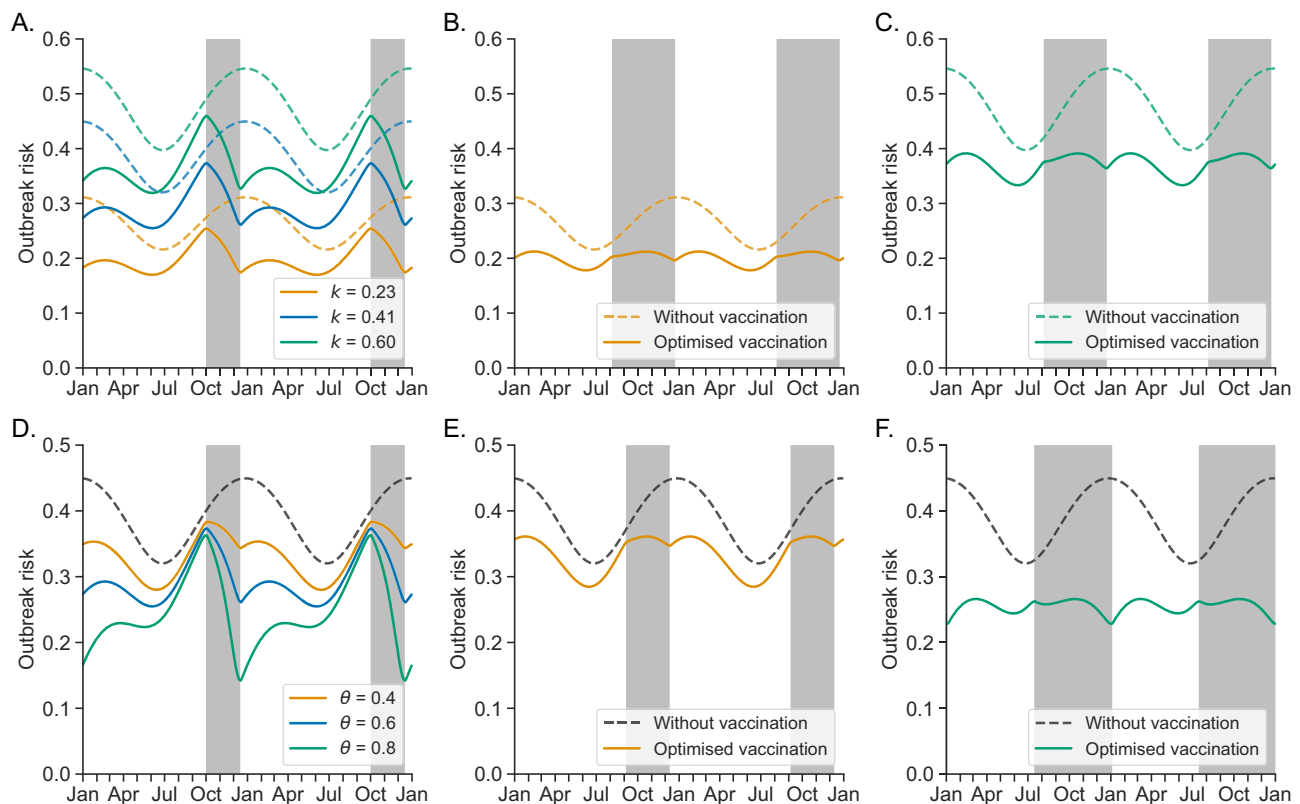
We analysed the impacts of numerous transmission features, which may vary based on factors such as SARS-CoV-2 variant and population

characteristics, on outbreak risk estimates under booster vaccination and optimal vaccine distribution timing (Fig. 4 and Supplementary Figs. 3–5). While the extent of heterogeneity in infectiousness between individuals (Fig. 4A–C) and the (mean) level of transmissibility in the absence of booster vaccination (Supplementary Fig. 3) both substantially affect quantitative outbreak risk estimates, we found that the optimal vaccine distribution timing is identical (10 August to 18 December) for all values of these parameters considered. In other words, under a policy goal of minimising the peak outbreak risk through an annual vaccination campaign with a fixed number of doses, the practical decision of when to deploy the vaccine doses is not affected by either variant transmissibility or the extent of individual heterogeneity. This is likely because, for the range of parameter values considered, these quantities do not substantially affect the extent to which the outbreak risk varies between different times of year.

On the other hand, we found the proportion of the population who receive annual booster vaccines,  $\theta$ , to be an important determinant of the optimal distribution strategy (Fig. 4D–F). Under our baseline distribution timing scenario, because of waning immunity, the outbreak risk at the start of the annual vaccine distribution period is only slightly reduced when even a large proportion of individuals are vaccinated (Fig. 4D). The optimal vaccine distribution window is longer when more individuals are vaccinated, since avoiding low immunity at times of moderate transmission becomes relatively more important compared to increasing (already high) immunity levels at times of peak transmission (Fig. 4E, F). Similarly, we found the optimal distribution window to be longer in scenarios where the extent of seasonality in transmission is smaller (Supplementary Fig. 4) or vaccines confer greater protection against infection (Supplementary Fig. 5).

## Discussion

Estimates of the risk of infectious disease cases introduced into a population leading to a large outbreak are useful for planning pathogen surveillance and interventions. Building on previous work, we have developed a general outbreak risk estimation framework that accounts for two key details: time-dependent transmission<sup>18</sup>, and heterogeneity in the transmission potential between infected individuals<sup>11</sup>. To demonstrate how our framework could be applied to inform public health policy, we considered the risk of localised COVID-19 outbreaks under seasonal transmission and booster vaccination, using a multi-scale approach that incorporates an individual-level antibody dynamics model<sup>29–31</sup>. Typically, vaccines are distributed ahead of an expected winter peak in transmission, pre-empting the time of year at which



**Fig. 4 | Factors determining the impact of booster vaccination and the optimal timing of distribution.** **A–C** Effect of the dispersion parameter,  $k$ . **A** Outbreak risk without booster vaccination (dashed curves) and under the default assumed timing of booster vaccine distribution (solid curves), for values of  $k = 0.23$  (orange),  $0.41$  (our default assumed value; blue) and  $0.60$  (green)—these values represent the lower 95% confidence interval bound, mean estimate and upper 95% confidence interval bound obtained for COVID-19 in ref. 15, respectively. **B, C** Outbreak risk without booster vaccination (dashed curves) and with optimised booster vaccine

distribution timing (solid curves; the optimal distribution period each year is shown in the grey shaded regions) for  $k = 0.23$  (**B**) and  $k = 0.60$  (**C**). **D–F** Effect of the proportion of individuals vaccinated each year,  $\theta$ . **D** Outbreak risk without booster vaccination (black dashed curve) and under the default assumed timing of booster vaccine distribution (solid curves), for values of  $\theta = 0.4$  (orange),  $0.6$  (our default assumed value; blue) and  $0.8$  (green). **E, F** Outbreak risk without booster vaccination (black dashed curves) and with optimised booster vaccine distribution timing (solid curves) for  $\theta = 0.4$  (**E**) and  $\theta = 0.8$  (**F**).

healthcare resources are under the highest strain<sup>26</sup>. Our analyses highlight that under a range of conditions, it may be beneficial to distribute vaccines over a relatively long period starting in late summer or early autumn (even aside from any operational constraints on the speed of distribution). This is because moderate autumn transmissibility, combined with waning immunity from the previous vaccination campaign, may otherwise generate a high outbreak risk near the start of annual vaccination.

In our baseline analysis, we considered parameter values representative of an infectious novel SARS-CoV-2 variant (with similar transmission characteristics to the omicron variant upon its initial emergence) in a population with relatively high booster vaccine uptake (as may be the case in a vulnerable population). However, even for COVID-19, the inputs to our outbreak risk estimation framework may vary depending on factors such as SARS-CoV-2 variant, vaccine availability and uptake, vaccine type(s) used, and characteristics of the local population. Therefore, we conducted several sensitivity analyses, finding that vaccine coverage and effectiveness, as well as the extent of seasonality in transmission, are important factors for determining optimal vaccination deployment timing. In general, a longer vaccine distribution window, starting earlier, decreases the outbreak risk near the time at which immunity is lowest (by bringing this time earlier in the year), at the expense of reduced immunity levels and so a higher outbreak risk around the winter transmission peak—the optimised distribution window considered here balances these factors to minimise the annual peak outbreak risk. If vaccine coverage is lower,

vaccine effectiveness is lower, and/or transmission seasonality is stronger, the optimal autumn vaccination campaign length is shorter, since the importance of maximising population immunity around the transmission peak is then increased.

Even in a scenario with higher vaccine uptake and optimised distribution timing, a non-zero outbreak risk may remain (Fig. 4F). While for simplicity we considered annual booster vaccination here, this finding suggests that more frequent vaccination may be required to prevent outbreaks of an infectious variant entirely in highly vulnerable populations such as care homes. In countries such as the UK, spring booster vaccinations have been offered to the most vulnerable members of the population, in addition to the annual autumn vaccinations. We expect that the optimal timing of autumn vaccination would be later and of shorter duration under twice- than once-yearly vaccination, since then the issue of low immunity near the start of vaccine distribution would be reduced.

In populations in which preventing outbreaks is impossible, the primary objective of booster vaccination campaigns would likely be to protect individuals from severe outcomes while the pathogen spreads. Decisions regarding which groups should receive booster vaccines, and how often, would then be based on individual risk of severe disease (considering waning immunity), alongside other factors such as economic cost. While we focussed on how the timing of annual booster vaccination (for a specified coverage level) can be optimised to reduce the outbreak risk, this is likely to correspond to a lower incidence of severe outcomes, since fewer outbreaks (particularly in vulnerable

populations) will usually lead to fewer severe cases (provided that the extent of individual heterogeneity remains unchanged).

We assumed that an optimal booster vaccine deployment strategy is one that minimises the annual peak of the outbreak risk (calculated each day assuming a single incident infection), but other possibilities could be considered. One option would be to consider the total annual outbreak risk, accounting for temporal variation in the likelihood of incident infections. If infectious introductions are most likely to occur when transmissibility (in the absence of vaccination) is highest, we expect that a shorter and later vaccination distribution window would be preferable to increase population immunity at that time. More generally, other modelling studies have explored how booster vaccination campaigns can be optimised, considering quantities such as peak hospitalisations<sup>33</sup>, total deaths<sup>34</sup> and measures accounting for the extent of premature mortality<sup>34</sup>. Our approach here should be used alongside other sources of evidence (depending on policymakers' objectives).

While we considered COVID-19 as a case study, we expect both our outbreak risk estimation framework and insights about seasonal vaccination to be applicable more widely. We caution, however, that our analytical approach may not be suitable in scenarios in which transmission is impossible or unlikely at some times of year (for example, this is often the case for mosquito-borne diseases in temperate climates). Specifically, if outbreaks are always extinguished when transmission becomes low, even if a substantial number of cases have occurred already<sup>35</sup>, the outbreak risk as calculated by numerically solving our analytically derived equations may be zero. In this scenario, we recommend using a simulation-based approach to estimate the outbreak risk (as used in Fig. 1E to verify the analytical outbreak risk estimates); careful consideration should be given to the specified incidence threshold for a major outbreak, since outbreak risk estimates may then be sensitive to the choice of threshold<sup>35</sup>.

As with any modelling analysis, our results required us to make assumptions and simplifications. In each of our results, we were required to set parameter values characterising transmission features such as the mean transmission rate, the extent of seasonality, and the timing of peak transmission in the absence of vaccination (assuming transmission varies sinusoidally without vaccination), as well as vaccine coverage (which depends on both availability and uptake) and effectiveness. While these quantities are not known precisely, particularly for future SARS-CoV-2 variants, we ran a range of sensitivity analyses to draw general principles from our results, including identifying the parameters that affect the optimal time of year at which to deploy vaccines. If precise quantitative outbreak risk estimates are required, in principle a Bayesian approach could be adopted in which (expert-ascertained) uncertainty in future transmission conditions is incorporated into outbreak risk estimates and vaccination planning. We expect that uncertainty in the timing of peak transmission may lengthen the optimal duration of annual vaccine distribution to maintain population immunity over a range of possible times of high transmission.

We assumed that antibody dynamics only affect susceptibility to SARS-CoV-2 infection, but future work may extend our general framework to include a link between antibody titre (prior to breakthrough infection) and subsequent infectiousness. Specifically, our approach here could be combined with our previously developed framework for including individual-level virus dynamics in outbreak risk estimates<sup>6</sup>, incorporating a relationship between prior immunity and virus dynamics during infection<sup>36</sup>. Because incorporating virus dynamics allows for detailed modelling of interventions such as antigen test screening<sup>6</sup>, this extension would enable simultaneous modelling of vaccination and other interventions. Another target for future work is to relax our assumption of random vaccination by incorporating age structure; for COVID-19, older individuals have typically been prioritised for vaccination. Additionally, contact rates<sup>34,37</sup> and

susceptibility<sup>34</sup> may vary between age groups, and these features could be included in an age-structured model. We note that non-random vaccination may alter the extent of individual heterogeneity in transmission—for example, targeting high-contact individuals would reduce heterogeneity. As a result, in addition to the outbreak risk, non-random vaccination strategies could affect the explosiveness of outbreaks that do occur (by altering the likelihood of superspreading); such impacts could be explored via model simulations.

Despite these simplifications, we have shown how outbreak risk estimates can be obtained in scenarios in which transmission is both heterogeneous between infected individuals and temporally varying. Applying our approach to the design of annual COVID-19 vaccination campaigns, we have identified vaccine coverage and effectiveness, and the extent of seasonality in transmission, as key determinants of the optimal vaccine distribution timing for minimising the annual peak outbreak risk. We expect these insights to be applicable to routine vaccination campaigns against other diseases (e.g., seasonal influenza) in addition to COVID-19. Our general modelling framework can be adapted to quantify outbreak risks in a wide range of different scenarios as an aid to public health policy-making.

## Methods

### Transmission model

We considered a discrete-time generalised renewal equation model that accounts for heterogeneity in infectiousness between different infected individuals. In the transmission model, each infected individual is assigned an individual infectiousness factor,  $\alpha$ , from a gamma distribution with shape parameter  $k$  and mean one (although, in principle, other distributions could be used). If the reproduction number,  $R$ , remains constant during that individual's infection, the realised number of secondary infections generated by the individual is then assumed to follow a Poisson distribution with mean  $\alpha R$ . This Poisson-gamma formulation is equivalent to assuming that the realised number of secondary infections from a randomly chosen infected individual follows a negative binomial offspring distribution with mean  $R$  and dispersion parameter  $k$ <sup>11</sup>. Smaller values of  $k$  correspond to a greater degree of superspreading (Supplementary Fig. 1). The offspring distribution is a geometric distribution if  $k = 1$ , while a Poisson distribution is obtained in the limit  $k \rightarrow \infty$ <sup>11</sup>.

More generally, accounting for temporal changes in the reproduction number,  $R_t$ , between different times,  $t$ , the model-simulated incidence of newly infected individuals,  $I_t$ , is generated as follows:

$$I_t \sim \text{Poisson} \left( \text{mean} = R_t \sum_{\tau=1}^{\infty} w_{\tau} J_{t-\tau} \right) \quad (1)$$

$$J_t \sim \text{Gamma} \left( \text{shape} = k I_t, \text{scale} = \frac{1}{k} \right) \quad (2)$$

Here,  $w_{\tau}$  is the probability mass function of the discrete generation time distribution (the probability that the interval between the dates of infection of an infector-infected pair is  $\tau$  days), and  $J_t = \sum_{i=1}^{I_t} \alpha_{t,i}$  represents the sum of the individual infectivity factors ( $\alpha_{t,i}$ ) of the individuals infected on day  $t$  (labelled  $i=1, \dots, I_t$ ). We note that this formulation is valid because the sum of independent gamma-distributed random variables with identical scale parameters also follows a gamma distribution (with the same scale parameter).

### Outbreak risk

**Analytical approach.** The outbreak risk,  $p_t$ , is defined to be the probability (under the above transmission model, and assuming that the reproduction number is known at all times) that an outbreak, initiated with a single incident infection at time  $t$  (i.e., taking  $I_t = 1$ , with  $I_s = 0$  for  $s < t$ ), does not become extinct. As shown in the

Supplementary Methods, the outbreak risk satisfies the following equation:

$$p_t = 1 - \left( 1 + \frac{1}{k} \sum_{\tau=1}^{\infty} R_{t+\tau} w_{\tau} p_{t+\tau} \right)^{-k} \quad (3)$$

In our numerical analyses, we assumed that  $R_t$  is a periodic function of time (so that  $p_t$  is also periodic with the same period). In that scenario, a closed system of equations for the values of  $p_t$  (across all values of  $t$ ), which can be solved numerically, can be obtained from the above expression (see the Supplementary Methods for details).

If  $R_t = R$  is constant in time, then the implicit equation for the outbreak risk,  $p_t = p$ , derived in ref. 11 is recovered:

$$p = 1 - \left( 1 + \frac{Rp}{k} \right)^{-k} \quad (4)$$

The outbreak risk is then given by the largest solution less than one of Eq. (4)<sup>11</sup>.

**Simulation-based approach.** An alternative method for calculating the outbreak risk at time  $t$  involves simulating the renewal model a large number of times, initiating each simulation with a single incident infection at time  $t$ . The outbreak risk is then calculated as the proportion of model simulations in which the daily incidence of new infections ever exceeds a specified threshold. Each simulation is run until either the outbreak ends (i.e., there are no new infections for a period exceeding the maximal generation time) or the incidence threshold is exceeded. When we used the simulation-based approach in Fig. 1E, we conducted 20,000 model simulations for each time of pathogen introduction and used an incidence threshold of 30 infections.

**Preliminary analysis of seasonal transmission**

We initially considered a simple scenario in which the reproduction number varies periodically with period 365 days:

$$R_t = 2 + \cos\left(\frac{2\pi t}{365}\right) \quad (5)$$

Over the course of each year,  $R_t$  therefore varies between one and three. We compared the outbreak risk, calculated using either the analytical or simulation-based approach, between different values of the dispersion parameter,  $k$  (Fig. 1E), under the generation time distribution for COVID-19 described below and shown in Fig. 1C.

**COVID-19 booster vaccination case study**

**Antibody dynamics model.** We considered a model of individual-level antibody dynamics following a dose of an mRNA COVID-19 vaccine<sup>29–31</sup>. In this model, the IgG(S) antibody titre,  $A(\tau)$ , of an individual at time since (most recent) vaccination  $\tau$  evolves according to the ordinary differential equation,

$$\frac{dA}{d\tau} = \frac{HM(\tau)^m}{K^m + M(\tau)^m} - \mu A(\tau), \quad \tau \geq \tau_d \quad (6)$$

where

$$M(\tau) = D e^{-\delta\tau} \quad (7)$$

is the quantity of mRNA inoculated by the vaccine dose remaining. Here,  $\tau_d$  represents a delay before the vaccine elicits an antibody response; for  $0 \leq \tau < \tau_d$ , the antibody titre of a previously vaccinated individual is assumed to continue to vary based on the remaining

inoculated mRNA and induced antibody titre from their previous vaccination dose (for a previously unvaccinated individual, the antibody titre is assumed to be zero until time  $\tau_d$ ). Interpretations of the model parameters  $H$ ,  $m$ ,  $K$ ,  $\mu$ ,  $D$  and  $\delta$  are given in Supplementary Table 1. We note that we here slightly simplified the model compared to previous studies<sup>29–31</sup> by assuming that residual mRNA from previous vaccination doses is negligible (after the delay,  $\tau_d$ ).

In a previous study<sup>31</sup>, the antibody dynamics model was fitted to longitudinal data collected following a booster (third overall) dose of the BNT162b2 or mRNA-1273 vaccine. In that study<sup>31</sup>, densely sampled blood data from 12 healthcare workers were considered first, indicating that only the parameters  $H$  and  $m$  exhibit substantial variability between individuals. Then, sparsely sampled data from a larger cohort of 1618 individuals were used to obtain individual estimates of  $H$  and  $m$ , with remaining parameters fixed based on model fits to the healthcare worker data. Here, we used the mean and standard deviation of logarithms of individual parameter estimates for the 1618 individuals (Supplementary Table 1) to generate synthetic antibody profiles (see below and the Supplementary Methods), assuming each parameter value varies between individuals according to a lognormal distribution.

**Individual susceptibility model.** We assumed that individual relative susceptibility (i.e. one minus the level of vaccine protection against infection),  $S(\tau)$ , is given in terms of the antibody titre by

$$S(\tau) = \frac{1}{1 + \exp\left(\kappa \log_{10}\left(\frac{A(\tau)}{A_{1/2}}\right)\right)} \quad (8)$$

This functional relationship has previously been applied to relate susceptibility to neutralising antibody titres<sup>38–41</sup>, and the application of the same formula here is justified by the high correlation between neutralising antibodies and the IgG(S) antibody titre<sup>31</sup>. We assumed a value of  $\kappa = 3.1$  as estimated in ref. 40. The parameter  $A_{1/2}$ , representing the antibody titre at which 50% protection is conferred, is likely to depend on both vaccine type and the SARS-CoV-2 variant under consideration<sup>25,38,39,41</sup>. Therefore, we took a default value of  $A_{1/2} = 1,000$  AU/mL—this gave a maximum average vaccine protection against infection of 79% (Fig. 2B), which is comparable to previous model estimates of the effectiveness of a fourth vaccine dose against symptomatic disease due to the omicron SARS-CoV-2 variant<sup>25</sup>. We also considered alternative values of  $A_{1/2}$  in Supplementary Fig. 5.

**Annual booster vaccination.** We suppose that booster vaccines are received by a proportion,  $\theta$ , of the population, each year between calendar times  $t_s$  and  $t_e$  (assuming that vaccinations are spread equally between each day in this time window).

To calculate the mean population susceptibility,  $\eta_t$ , on each calendar day  $t$ , we considered a synthetic cohort of 10,000 individuals. Considering individual variation in antibody dynamics model parameters and assuming an interval of one year between successive vaccine doses, we calculated periodic solutions of the antibody dynamics model for each individual. This allowed us to calculate the expected susceptibility,  $\bar{S}(\tau)$ , of individuals included in booster vaccination campaigns, as a function of time since most recent vaccine dose,  $\tau$ . Details of this calculation are given in the Supplementary Methods. Assuming individuals who do not receive booster vaccinations have relative susceptibility one, and for simplicity neglecting variation in susceptibility within each day, we then calculated

$$\eta_t = (1 - \theta) + \theta \sum_{t_v = t_s}^{t_e} \bar{S}(t - t_v) \quad (9)$$

In our main analyses, we took  $\theta = 0.6$  to represent a vulnerable population with relatively high booster uptake. Alternative values of  $\theta$  are considered in Fig. 4D–F. Initially, we considered booster vaccination rollout between  $t_s = 1$  October and  $t_e = 15$  December, which is similar to the original planned timetable for 2023 in England<sup>42</sup> (the 2023 booster vaccination start date was later moved forwards in response to the emergence of the BA.2.86 omicron subvariant<sup>43</sup>). We then explored how the rollout timing could be optimised to minimise the annual peak outbreak risk.

**Seasonal transmissibility.** In the absence of booster vaccination (i.e., assuming all individuals are entirely susceptible), we assumed the instantaneous basic reproduction number to be given by<sup>27,44</sup>

$$R_t = R_{0,t} = \bar{R}_0 \left( 1 + \varepsilon \cos \left( \frac{2\pi t}{365} \right) \right) \quad (10)$$

We assumed that the basic reproduction number is highest on 1 January each year, which is consistent with the respiratory infectious disease burden in the UK typically being highest in December and January<sup>42</sup>.

We took default values of  $\varepsilon = 0.2$ , as previously assumed for COVID-19 based on estimates for other coronaviruses in ref. 27, and  $\bar{R}_0 = 2.5$ , giving a maximum annual basic reproduction number ( $R_{0,t} = 3$ ) consistent with estimates of the reproduction number of the omicron SARS-CoV-2 variant (BA.1 subvariant) shortly after it emerged in late 2021<sup>45</sup>. Different values of  $\bar{R}_0$  and  $\varepsilon$  are considered in Supplementary Figs. 3, 4, respectively.

The instantaneous reproduction number accounting for booster vaccination was then calculated as  $R_t = \eta_t R_{0,t}$ .

**Heterogeneity in infectiousness.** In our initial analysis, we assumed a dispersion parameter of  $k = 0.41$ , as estimated for COVID-19 in ref. 15. Alternative values of  $k$  are considered in Fig. 4A–C.

**Generation time distribution.** Throughout our analyses, we assumed a lognormal continuous generation time distribution (i.e., distribution of the interval between the exact times of infection of an infector–infectee pair) with mean 3.0 days and standard deviation 1.5 days (this corresponds to lognormal distribution parameters, describing the mean and standard deviation of the natural logarithm of the continuous generation time, of  $\mu = 0.98$  and  $\sigma = 0.47$ , respectively), as estimated for household SARS-CoV-2 omicron variant transmissions in ref. 32. We then discretised this distribution using the method described in refs. 46,47, truncating the discretised distribution at 14 days (after which the remaining probability mass is negligible). The discretised generation time distribution is shown in Fig. 1C.

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

All data are available at <https://github.com/will-s-hart/covid-boosters> (archived at <https://doi.org/10.5281/zenodo.15800215>). Individual parameter estimates for the antibody dynamics model were originally obtained by ref. 31.

### Code availability

All code used in the analyses is available at <https://github.com/will-s-hart/covid-boosters> (archived at <https://doi.org/10.5281/zenodo.15800215>). Computer code was written in Python (compatible with version 3.12).

## References

- Southall, E. et al. A practical guide to mathematical methods for estimating infectious disease outbreak risks. *J. Theor. Biol.* **562**, 111417 (2023).
- Thompson, R. N. Novel coronavirus outbreak in Wuhan, China, 2020: intense surveillance is vital for preventing sustained transmission in new locations. *J. Clin. Med.* **9**, 498 (2020).
- Anzai, A. et al. Assessing the impact of reduced travel on exportation dynamics of novel coronavirus Infection (COVID-19). *J. Clin. Med.* **9**, 601 (2020).
- Boldog, P. et al. Risk assessment of novel coronavirus COVID-19 outbreaks outside China. *J. Clin. Med.* **9**, 571 (2020).
- Hellewell, J. et al. Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. *Lancet Glob. Health* **8**, e488–e496 (2020).
- Hart, W. S. et al. Analysis of the risk and pre-emptive control of viral outbreaks accounting for within-host dynamics: SARS-CoV-2 as a case study. *Proc. Natl Acad. Sci. USA* **120**, e2305451120 (2023).
- Sachak-Patwa, R., Byrne, H. M., Dyson, L. & Thompson, R. N. The risk of SARS-CoV-2 outbreaks in low prevalence settings following the removal of travel restrictions. *Commun. Med.* **1**, 39 (2021).
- Lovell-Read, F. A., Shen, S. & Thompson, R. N. Estimating local outbreak risks and the effects of non-pharmaceutical interventions in age-structured populations: SARS-CoV-2 as a case study. *J. Theor. Biol.* **535**, 110983 (2022).
- Lovell-Read, F. A., Funk, S., Obolski, U., Donnelly, C. A. & Thompson, R. N. Interventions targeting non-symptomatic cases can be important to prevent local outbreaks: SARS-CoV-2 as a case study. *J. R. Soc. Interface* **18**, 20201014 (2021).
- Woolhouse, M. E. J. et al. Heterogeneities in the transmission of infectious agents: Implications for the design of control programs. *Proc. Natl Acad. Sci. USA* **94**, 338–342 (1997).
- Lloyd-Smith, J. O., Schreiber, S. J., Kopp, P. E. & Getz, W. M. Superspreading and the effect of individual variation on disease emergence. *Nature* **438**, 355–359 (2005).
- Endo, A., Abbott, S., Kucharski, A. J. & Funk, S. Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China. *Wellcome Open Res* **5**, 67 (2020).
- Sun, K. et al. Transmission heterogeneities, kinetics, and controllability of SARS-CoV-2. *Science* **371**, eabe2424 (2021).
- Wang, L. et al. Inference of person-to-person transmission of COVID-19 reveals hidden super-spreading events during the early outbreak phase. *Nat. Commun.* **11**, 5006 (2020).
- Wegehaupt, O., Endo, A. & Vassall, A. Superspreading, overdispersion and their implications in the SARS-CoV-2 (COVID-19) pandemic: a systematic review and meta-analysis of the literature. *BMC Public Health* **23**, 1–22 (2023).
- Althaus, C. L. Ebola superspreading. *Lancet Infect. Dis.* **15**, 507–508 (2015).
- Ke, R. et al. Daily longitudinal sampling of SARS-CoV-2 infection reveals substantial heterogeneity in infectiousness. *Nat. Microbiol.* **7**, 640–652 (2022).
- Kaye, A. R., Hart, W. S., Bromiley, J., Iwami, S. & Thompson, R. N. A direct comparison of methods for assessing the threat from emerging infectious diseases in seasonally varying environments. *J. Theor. Biol.* **548**, 111195 (2022).
- Bacaër, N. Deux modèles de population dans un environnement périodique lent ou rapide. *J. Math. Biol.* **80**, 1021–1037 (2020).
- Carmona, P. & Gandon, S. Winter is coming: Pathogen emergence in seasonal environments. *PLoS Comput. Biol.* **16**, e1007954 (2020).

21. Hridoy, M. B. & Allen, L. J. S. Investigating seasonal disease emergence and extinction in stochastic epidemic models. *Math. Biosci.* **381**, 109383 (2025).
22. Markov, P. V. et al. The evolution of SARS-CoV-2. *Nat. Rev. Microbiol.* **21**, 361–379 (2023).
23. Goldberg, Y. et al. Protection and waning of natural and hybrid immunity to SARS-CoV-2. *N. Engl. J. Med.* **386**, 2201–2212 (2022).
24. Menegale, F. et al. Evaluation of waning of SARS-CoV-2 vaccine-induced immunity: a systematic review and meta-analysis. *JAMA Netw. Open* **6**, e2310650 (2023).
25. Hogan, A. B. et al. Estimating long-term vaccine effectiveness against SARS-CoV-2 variants: a model-based approach. *Nat. Commun.* **14**, 4325 (2023).
26. Joint Committee on Vaccination and Immunisation. JCVI statement on the COVID-19 vaccination programme for autumn 2024. <https://www.gov.uk/government/publications/covid-19-autumn-2024-vaccination-programme-jcvi-advice-8-april-2024> (2024).
27. Engelbrecht, F. A. & Scholes, R. J. Test for Covid-19 seasonality and the risk of second waves. *One Health* **12**, 100202 (2021).
28. Liu, X. et al. The role of seasonality in the spread of COVID-19 pandemic. *Environ. Res.* **195**, 110874 (2021).
29. Nishiyama, T. et al. Modeling COVID-19 vaccine booster-elicited antibody response and impact of infection history. *Vaccine* **41**, 7655–7662 (2023).
30. Nakamura, N. et al. Modeling and predicting individual variation in COVID-19 vaccine-elicited antibody response in the general population. *PLoS Digit. Health* **3**, e0000497 (2024).
31. Park, H. et al. Longitudinal analysis of antibody titers after primary and booster mRNA COVID-19 vaccination can identify individuals at risk for breakthrough infection. Preprint at *bioRxiv* <https://www.biorxiv.org/content/10.1101/2025.03.01.639360v1> (2025).
32. Park, S. W. et al. Inferring the differences in incubation-period and generation-interval distributions of the Delta and Omicron variants of SARS-CoV-2. *Proc. Natl Acad. Sci. USA* **120**, e2221887120 (2023).
33. Ko, Y., Peck, K. R., Kim, Y. J., Kim, D. H. & Jung, E. Effective vaccination strategies to control COVID-19 in Korea: a modeling study. *Epidemiol. Health* **45**, e2023084 (2023).
34. Bourros, I., Hill, E. M., Keeling, M. J., Moore, S. & Thompson, R. N. Prioritising older individuals for COVID-19 booster vaccination leads to optimal public health outcomes in a range of socio-economic settings. *PLoS Comput. Biol.* **20**, e1012309 (2024).
35. Kaye, A. R., Guzzetta, G., Tildesley, M. J. & Thompson, R. N. Quantifying infectious disease epidemic risks: a practical approach for seasonal pathogens. *PLoS Comput. Biol.* **21**, e1012364 (2025).
36. Lingas, G. et al. Neutralizing antibody levels as a correlate of protection against SARS-CoV-2 infection: a modeling analysis. *Clin. Pharm. Ther.* **115**, 86–94 (2024).
37. Prem, K. et al. Projecting contact matrices in 177 geographical regions: an update and comparison with empirical data for the COVID-19 era. *PLoS Comput. Biol.* **17**, e1009098 (2021).
38. Khoury, D. S. et al. Predicting the efficacy of variant-modified COVID-19 vaccine boosters. *Nat. Med.* **29**, 574–578 (2023).
39. Cromer, D. et al. Neutralising antibody titres as predictors of protection against SARS-CoV-2 variants and the impact of boosting: a meta-analysis. *Lancet Microbe* **3**, e52–e61 (2022).
40. Khoury, D. S. et al. Neutralizing antibody levels are highly predictive of immune protection from symptomatic SARS-CoV-2 infection. *Nat. Med.* **27**, 1205–1211 (2021).
41. Cromer, D. et al. Predicting vaccine effectiveness against severe COVID-19 over time and against variants: a meta-analysis. *Nat. Commun.* **14**, 1633 (2023).
42. NHS England. Autumn/winter (AW) 2023-24 flu and COVID-19 seasonal campaign. <https://www.england.nhs.uk/long-read/autumn-winter-aw-2023-24-flu-and-covid-19-seasonal-campaign/> (2023).
43. NHS England. NHS flu and covid vaccine programmes brought forward due to risk of new covid variant. <https://www.england.nhs.uk/2023/08/nhs-flu-and-covid-vaccine-programmes-brought-forward-due-to-risk-of-new-covid-variant/> (2023).
44. Kelly, S. L., Le Rutte, E. A., Richter, M., Penny, M. A. & Shattock, A. J. COVID-19 vaccine booster strategies in light of emerging viral variants: frequency, timing, and target groups. *Infect. Dis. Ther.* **11**, 2045–2061 (2022).
45. Wang, S., Zhang, F., Wang, Z., Du, Z. & Gao, C. Reproduction numbers of SARS-CoV-2 Omicron subvariants. *J. Travel Med.* **29**, taac108 (2022).
46. Cori, A., Ferguson, N. M., Fraser, C. & Cauchemez, S. A new framework and software to estimate time-varying reproduction numbers during epidemics. *Am. J. Epidemiol.* **178**, 1505–1512 (2013).
47. Thompson, R. et al. Using real-time modelling to inform the 2017 Ebola outbreak response in DR Congo. *Nat. Commun.* **15**, 5667 (2024).

## Acknowledgements

Thanks to members of the Wolfson Centre for Mathematical Biology at the University of Oxford, particularly the Infectious Disease Modelling group, for useful discussions about this work. This research was funded by the EPSRC through a Vacation Internship (to J.A., grant number EP/W524311/1). W.S.H., A.R.K. and R.N.T. acknowledge the support of the JUNIPER partnership (grant number MR/X018598/1). The collaboration between R.N.T. and S.I. was supported by a Royal Society International Exchange award (grant number IES-R3-193037). For the purpose of Open Access, the authors have applied a CC BY public copyright licence to any Author Accepted Manuscript (AAM) version arising from this submission.

## Author contributions

W.S.H.: conceptualisation, methodology, formal analysis, software, investigation, visualisation, supervision, writing (original draft) and writing (review and editing). J.A.: methodology, formal analysis, software, investigation, visualisation and writing (review and editing). H.P., K.K., Y.D.J. and A.R.K.: methodology and writing (review and editing). S.I.: conceptualisation, methodology and writing (review and editing). R.N.T.: methodology, supervision and writing (review and editing).

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-025-63375-5>.

**Correspondence** and requests for materials should be addressed to William S. Hart.

**Peer review information** *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025